"Driver or passenger" : an integrated epidemiological and experimental perspective on the association between nontyphoidal salmonella infection and colon cancer

Duijster, J.W.

“DRIVER OR PASSENGER?”

An integrated epidemiological and experimental perspective on the association between nontyphoidal Salmonella infection and colon cancer

Janneke W. Duijster

The research presented in this thesis was supported by the Dutch Cancer Society (KWF) (grant number 2017-1-11001) and the Dutch Ministry for Public Health, Welfare and Sport via 'ZonMw' (grant number 522004001).

ISBN:	978-94-6458-393-9
Cover design:	The picture originates from Shutterstock
Lay-out:	Publiss | www.publiss.nl
Print:	Ridderprint | www.ridderprint.nl

“DRIVER OR PASSENGER?”

An integrated epidemiological and experimental perspective on the association between nontyphoidal Salmonella infection and colon cancer

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 12 oktober 2022
klokke 11.15 uur

door

Janneke Willemina Duijster

geboren te Ede
in 1991

Promotor

Prof. dr. J.J.C. Neefjes	

Co-promotores

Dr. L. Mughini-Gras	RIVM/UU
Dr. E. Franz	RIVM

Leden promotiecommissie

Prof. dr. R.C. Hoeben	
Prof. dr. E.J. Kuijper	
Prof. dr. F.R. Rosendaal	
Prof. dr. D. van Baarle	UMCG/RIVM
Dr. M. Schaapveld	NKI-AVL

CONTENTS

Chapter 1

GENERAL INTRODUCTION

We live a life in close interaction with microorganisms. In fact, the number of bacteria we carry in and on our body exceeds our own cells by 10-fold. Many microorganisms are commensal and play a crucial role in the maintenance of homeostasis, host nutrient metabolism, education and regulation of the immune system and protection against invasion by pathogenic microorganisms. Some bacteria use cells in their life cycle. In this thesis, we focus on one intracellular pathogenic microorganism in particular: nontyphoidal *Salmonella*, a bacterium that invades host cells and lives and propagates there before the release of many daughter Salmonellae. During the intracellular part of the life cycle, *Salmonella* partially controls the biochemical pathways in host cells. These pathways can also be involved in cell growth and transformation, which led to the question whether *Salmonella* can induce cancer as the results of its control and manipulation of the host cell biology. This question and how to investigate this beyond laboratory experiments, is topic of this thesis.

Pathogenesis of gastrointestinal bacterial infections

Gastroenteritis is caused by an inflammation of the stomach and/or intestines and is generally characterized by diarrhea and/or vomiting, which is commonly accompanied by other symptoms, including nausea, abdominal pain, fever, fatigue and headache. Leading infectious causes of gastroenteritis in Europe include norovirus, rotavirus (particularly in children), *Campylobacter* spp., nontyphoidal *Salmonella*, and pathogenic *Escherichia coli* [1, 2]. Particularly for bacteria, infection with these pathogens mainly occurs through ingestion of contaminated food or water or contact with infected animals [3]. Gastroenteritis can also be the result of non-infectious causes such as food intolerance and side effects of medication. Usually, symptoms are mild and self-limiting not requiring medical attention, hence, the etiological agent is only diagnosed in a small fraction of infections [4].

Salmonella enterica, is a genus of Gram-negative facultative anaerobe intracellular bacteria belonging to the Enterobacteriaceae family. The genus consists of six subspecies of which the subspecies *enterica* is the only one relevant for public health as virtually all human *Salmonella* infections are caused by bacteria within this subspecies [5]. Based on differences in (expression of) somatic (O), flagellar (H) and capsular (K) antigens, over 2600 distinct serovars have been identified within the genus *Salmonella* [6]. While most *Salmonella* serovars are considered generalists, able to colonize a broad range of hosts including mammals, birds, reptiles and humans; *S.* Typhi and *S.* Paratyphi (A, B, and C) are specialists and restricted to humans [7, 8]. Serovars other than *S*. Typhi and *S*. Paratyphi are referred to as nontyphoidal *Salmonella*. In contrast to nontyphoidal *Salmonella* which is a causative agent of gastroenteritis in both low- and high-income countries, typhoidal *Salmonella* is

nowadays endemic only in developing countries where it causes typhoid fever characterized by a severe systemic infection affecting multiple organs (spleen, liver, gallbladder) [7].

Nontyphoidal *Salmonella* is able to infect several cell types, including epithelial cells, fibroblasts, macrophages and dendritic cells [9]. Successful cellular uptake, survival and proliferation in the host are achieved by deploying various mechanisms [10]. These include the production of toxins, alteration of (cell) signaling pathways and the induction of inflammatory responses [10]. Collectively, these mechanisms can lead to loss of the barrier function of the epithelial cell layer due to disruption of the cells and epithelial junctions, consequently causing an imbalance in absorption and secretion of fluids and electrolytes (i.e. diarrhea) [10, 11]. Both the pathogenesis and the degree of pathogenicity differ strongly between subspecies or serotypes, owing to differences in presence and expression of so-called virulence factors [12]. One such virulence factor for *Salmonella* is its type three secretion system (T3SS). This complex syringe-like structure (analogues are used by many Gram-negative bacteria), enables bacteria to translocate its proteins across the host cell membrane into the host cell cytoplasm [13, 14]. After attachment to epithelial cells in the terminal ileum, *Salmonella* injects a range of effector proteins into the host cell. Part of these proteins (the *Salmonella* pathogenicity island [SPI]-1 proteins), including SopB, SopE, SopE2, SipA and SipC have a proinflammatory effect. They mediate the invasion of *Salmonella* into the host cell by inducing membrane ruffling, followed by endocytosis/pinocytosis of the bacterium and biogenesis of the *Salmonella*-containing vacuole (SCV) (Figure 1). After engulfment into the cell, Salmonella also secretes proteins with an anti-inflammatory function (SPI1 and SPI2 proteins), including SptP, SspH1, AvrA, and SipA. These proteins restore the host cell surface and enable Salmonella to survive and replicate within the host cell [13, 14, 15, 16] (Figure 1).

While *Salmonella* is traditionally considered a vacuolar pathogen living inside the SCV, a portion of the bacteria can survive and exert their effects in the host cell cytosol [9]. The ratio of intra-vacuolar and cytosolic bacteria in the host cell is assumed to be cell type dependent, yet, the triggers that direct the bacterium into an intra-vacuolar or cytosolic pathway are not fully understood as is the question how the co-existence of two intracellular niches impacts the pathogenicity of *Salmonella* [9]. Several effector proteins contribute to the inflammatory response by activating the NF-κB signaling pathway, resulting in secretion of proinflammatory cytokines by epithelial cells and activation of the host immune system [16, 17]. The subsequent influx of neutrophils into the inflamed intestinal tissue is suggested to be one of the explanations for the induction of diarrhea, although the exact mechanisms are not yet fully understood [17]. In addition to gastroenteritis, *Salmonella* is able to invade tissues

beyond the intestinal epithelium, thereby causing an invasive infection, characterized by infection of normally sterile sites, sepsis and bacteremia [18]. Invasive *Salmonella* infections more often occur at older age and are associated with a higher risk of hospitalization.

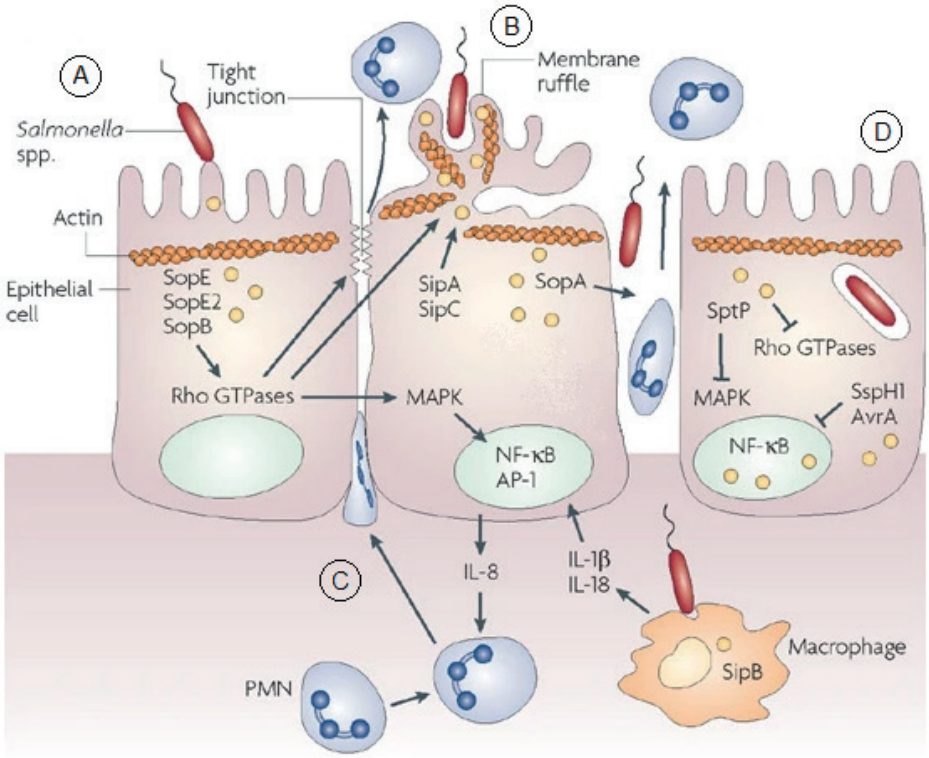


Figure 1. Schematic illustration of nontyphoidal *Salmonella* invasion of epithelial cells. [A]. Attachment of *Salmonella* to apical surface of the epithelial cell and secretion of SPI1 effector proteins including SopE, SopE2 and SopB. The effector proteins activate the host Rho GTPases and induce [B] actin cytoskeletal rearrangements, membrane ruffling and uptake of the bacterium through pinocytosis. [C]. The effector proteins induce an inflammatory response by activating the NF-κB pathway, resulting in the production of proinflammatory cytokines such as IL-8 and the subsequent transepithelial migration of polymorphonuclear neutrophils (PMNs). [D] After uptake of *Salmonella* into the *Salmonella*-containing vacuole (SCV), SPI1 and SPI2 effector proteins (including SptP, SspH1 and AvrA) restore the cell surface, and inhibit Rho GTPase activation, the MAPK pathway and its downstream pathways. Adapted from Haraga, 2008 [16].

Epidemiology of gastrointestinal infections

Enteric infections accounted for an estimated 6.6 billion human infections and 1.7 million associated deaths worldwide in 2019, thereby remaining a leading cause of morbidity

and mortality, particularly among young children and in low-income countries [19]. In the Netherlands, the National Institute for Public Health and the Environment (RIVM) monitors the annual number of human infections and deaths caused by 14 foodborne pathogens. These pathogens collectively caused ~4.9 million incident cases of gastroenteritis in 2019. Two of the main causative agents, *Campylobacter* spp. and Shiga toxin-producing *E. coli* O157, accounted for ~73,000 and ~2100 infections respectively [2]. In contrast to several other European countries, incident cases of salmonellosis are not notifiable in the Netherlands [20]. Instead, monitoring of (trends in) *Salmonella* infections is based on data obtained from a laboratory network with an estimated population coverage of 64%. *Salmonella* isolates, mostly originating from people with severe symptoms (as testing for *Salmonella* is usually performed after 1-2 weeks of illness), are submitted by peripheral laboratories to the RIVM for subtyping/serotyping [20]. Most people acquire multiple *Salmonella* infections throughout life [21]. However, given the usually mild course of illness, not requiring medical attention, it is estimated that for each reported case of salmonellosis, 57 *Salmonella* infections go unreported in the European Union [22]. While about 1500-2000 laboratory-confirmed *Salmonella* infections are reported annually, the true number of symptomatic *Salmonella* infections in the Dutch population is estimated at 26,000 cases annually with 24 associated deaths [2, 18]. Globally, almost 85 thousands deaths were attributable to *Salmonella* infections in 2016, the majority of these were children <5 years (44.1%) and people aged ≥70 years (22.5%) [1]. In the Netherlands, ~30% and ~25% of the infections are reported among people aged <20 years and above 60 years respectively. Moreover, ~4.6% of the reported salmonellosis cases are invasive *Salmonella* infections [18].

The serotypes Enteritidis and Typhimurium (including the monophasic variant 1,4,[5],12:i:-) are responsible for the majority of human *Salmonella* infections in high-income countries, up to 70% in the European Union in 2019 [8, 23, 24]. Enteritidis is a highly poultry-adapted serovar with human infections almost exclusively attributable to eggs and poultry meat [8, 24, 25]. On the other hand, Typhimurium is virtually ubiquitous with human cases predominantly associated with the consumption of beef and pork, though other reported risk factors include dairy products, vegetables and fruits [24-27]. Other risk factors, not related to consumption, include contact with canine puppies, improper kitchen hygiene, playing in sandboxes, contact with people suffering gastroenteritis and the use of proton pump inhibitors allowing bacteria to pass the stomach [28].

Biology, etiology and risk factors of cancers in the gastrointestinal tract

The development of cancer results from a series of genetic modifications dysregulating normal cell growth and survival of cells [29]. Tumor formation is initiated by the abnormal proliferation of a single cell due to a genetic alteration, followed by accumulation of clonally derived tumor cells into a tumor cell population. As a result of continuous mutations within the tumor cells, some cells obtain selective advantages related to for instance more rapid growth, invasion or survival. During clonal selection, cells with the most beneficial properties become dominant, leading to growth of a tumor with increasing levels of malignancy and metastatic capacity [29]. Cancer is characterized by different developmental stages with distinct treatment options and prognoses. Definition of cancer stage is done according to several indicators, including the size of the primary tumor and whether it has spread to nearby tissues/organs (T0-T4) and/or lymph nodes (N0-N3) and/or metastasis to other organs (M0-M1) [30]. Moreover, the degree of tumor cell differentiation (low versus high grade) is often a proxy for the behavior and growth rate of tumors [30].

Gastrointestinal tract cancers affect the organs in the digestive tract, including the esophagus, stomach, small intestine, colon, rectum, anus, pancreas, liver, gallbladder and the biliary tract. Within most of these organs different subsites/regions exist with their own distinct risk factors for developing a tumor, diversity in clinical symptoms and prognosis [31-33]. Traditionally, cancers are classified according to the tissue and cell type from where they originate. Many gastrointestinal tract cancers are adenocarcinomas originating from glandular cells. These glandular cells are situated in the inner lining of the organs and have a secretory function (i.e. mucus production) [31]. The process and duration of tumor development until the first clinical signs is dependent on a variety of factors, including the type and location of the tumor, age and underlying medical conditions [34]. Estimates of the duration of the pre-clinical phase between onset of cancer development and the first symptoms (i.e. the sojourn time) remain elusive as outcomes are inconsistent between studies [35]. Nonetheless, the estimated sojourn times for esophageal, liver, gallbladder and pancreatic cancer tend to be lower as compared to gastric and colon cancer [34, 35]. Colon cancer can develop from an initially harmless polyp into dysplasia and ultimately a malignant polyp or invasive adenocarcinoma during a process that takes several years (Figure 2a) [36, 37]. The colon can be divided into the proximal colon, including the cecum and ascending and transverse parts, and the distal colon, consisting of the descending and sigmoid parts (Figure 2b). Early signs of colon cancer often include abdominal pain, weight loss, change in bowel habits (including diarrhea and blood in the feces), weakness and anemia [38]. Despite differences in tumor

growth behavior, treatment options and prognosis, colon and rectum cancer are often considered as one in cancer research, collectively named colorectal cancer (CRC) [39].

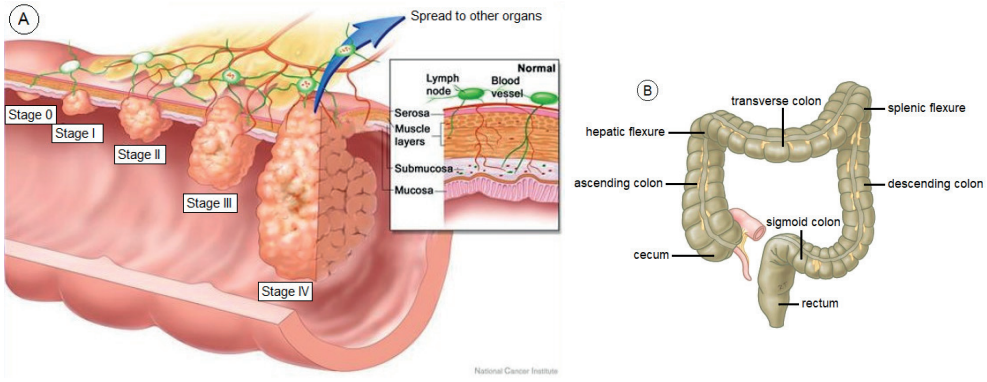


Figure 2. a) colon cancer stages. b)anatomical subsites of the colon.
a. retrieved from Colorectal Cancer Alliance, 2021 [40]; b. retrieved from Bredenoord, 2010 [41].

The risk of developing gastrointestinal tract cancer is partially defined by non-modifiable risk factors including hereditary causes, age and gender, with incidence generally being positively correlated to age and the male gender for most gastrointestinal tract cancers [42]. One such hereditary cause includes the familial adenomatous polyposis (FAP) syndrome, which is an autosomal dominant disease characterized by a mutation in the adenomatous polyposis coli (APC) gene. FAP patients develop hundreds to thousands of colonic polyps, already from childhood, with a consequent dramatic increase in cancer risk. Less than one percent of the colorectal cancer diagnoses is attributable to FAP, though the syndrome is also associated with gastric and duodenal cancer [43]. In addition, the presence of an (inflammatory) disease acts as predisposition for some cancers, as is the case for inflammatory bowel disease (IBD) and colon cancer and Barrett's esophagus in esophageal cancer [44]. IBDs such as Crohn's disease and ulcerative colitis, often present with abdominal pain, diarrhea, weight loss and fatigue due to chronic inflammations in the digestive tract, most frequently occurring in the small intestines, colon and rectum [45]. The cumulative probability of being diagnosed with colorectal cancer increases from around 2% for both conditions after 10 years of disease to 8% and 18% for Crohn's disease and ulcerative colitis respectively after 30 years of disease [45]. Amongst the risk factors considered modifiable, such as environmental or occupational exposures, dietary habits and factors related to lifestyle, differences in nature and magnitude of the risk factors exist between the distinct gastrointestinal malignancies. Whilst tobacco

use, alcohol consumption and obesity increase the risk of practically all gastrointestinal tract cancers, (occupational) exposure to chemical substances (lead, arsenic, pesticides etc.) or pathogens have been reported risk factors for esophageal, stomach and liver cancer [46]. Moreover, consumption of hot drinks, pickled or spicy food, salty foods and red and processed meat are associated with higher risk of esophageal, stomach, liver and colorectal cancer respectively [46]. Conversely, physical activity and a diet characterized by high levels of fruits, vegetables, whole grain and fibers decrease the cancer risk [46].

Epidemiology of cancers in the gastrointestinal tract

Globally, an estimated over 5 million new diagnoses of gastrointestinal tract cancers and 3.6 million associated deaths were reported in 2020 **(Table 1)** [47]. Colorectal (10.0%), gastric (5.6%), liver (4.7%) and esophageal cancer (3.1%) together accounted for 23.4% of all cancers [47]. Overall survival is generally low for most gastrointestinal tract cancers as the malignancy is often detected at a late stage. The age-standardized 5-year net survival for gastric, liver and pancreatic cancer is on average less than 30% in many countries, whilst the prognosis is better in case of colon cancer (up to 70% 5-year net survival) and may further increase with the National Colon Screening programs [48].

Table 1. Number of new diagnoses of gastrointestinal tract cancers in 2020, worldwide and in the Netherlands

Malignancy	Worldwide				The Netherlands
	Diagnoses	**Deaths**	**IR***	**Males (%)**	**Diagnoses**
Esophagus	604,100	544,076	6.3	69.3	2476
Stomach	1,089,103	768,793	11.1	66.1	984
Colon	1,148,515	576,858	11.4	52.3	8400
Rectum	732,210	339,022	7.6	60.6	3300
Anus					
Liver	905,677	830,180	9.5	69.8	800
Gallbladder & biliary tract	115,949	84,695	1.2	35.4	918
Pancreas	495,773	466,003	4.9	53.0	2668

* Age-standardized incidence rate per 100,000 person-years. Cancers of the small intestine are rare, hence, not shown in the table.
Data retrieved from WHO, 2021 [49] and IKNL, 2021 [50]

Substantial global diversity is observed in the incidence of the different gastrointestinal tract cancers, mainly due to differences in genetic susceptibility and the occurrence or magnitude of risk factors [51]. Colon, rectum and pancreatic cancer are most frequently diagnosed in Europe,

North America and Australia/New Zealand, reflecting the average higher levels of sedentary behavior in high-income countries. On the other hand, the incidence of esophageal cancer (particularly squamous cell carcinoma), gastric cancer and liver cancer is high in Asian countries resulting from a combination of lifestyle related factors (alcohol consumption and tobacco use), genetic predisposition, sanitary conditions as well as infectious causes (hepatitis B or C virus infection in case of liver cancer and*Helicobacter pylori* for gastric cancer) [33, 47, 51, 52].

Worldwide, colorectal cancer is the third most common cancer, with over a million new diagnoses annually (Table 1). In the Netherlands, almost 12000 people were diagnosed with colorectal cancer in 2020, about one-third of these represent rectal cancers (Table 1, Figure 3) [50]. More than half of the colorectal cancer diagnoses occur after the age of 70 year, whilst 6% of the people diagnosed were under the age of 50 years (based on 2020 data) [53]. The age distribution of the colorectal cancer diagnoses in the Netherlands largely corresponds to the age distributions in other countries, albeit the incidence in people before the age of 50 is increasing in a multitude of countries (including the Netherlands) for yet unknown reasons [54].

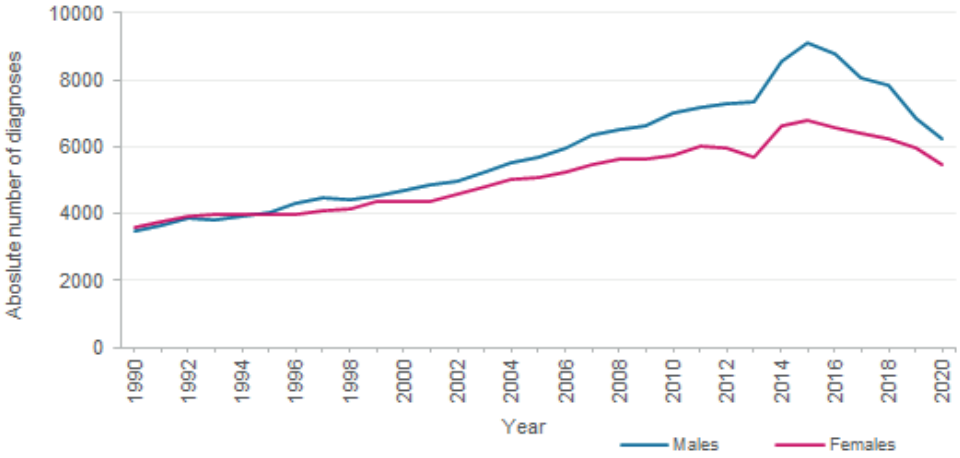


Figure 3. number of new diagnoses of colon and rectum cancer in 1990-2020. Rectal cancers represent ~30% of the colorectal cancer diagnoses.
Data retrieved from Atlasvolksgezondheid.info, 2021 [53].

In most Western/developed countries, a colorectal screening program is being executed by means of a non-invasive stool test (fecal occult blood test [FOBT] or fecal immunochemical test [FIT]) or a colonoscopy, which is offered regularly to adults above the age of 50 to 60 years [36, 55]. FOBT and FIT are based on detection of traces of blood in stool samples, as early sign of colorectal cancer. In the Netherlands, this screening program was introduced in

2014, which led to an increased number of diagnosis for several years (Figure 3). Nowadays, individuals aged between 55-75 years are invited every two years for a FIT, which is in case of a positive outcome followed by a colonoscopy [56]. The participation rate of the screening was 71.5% in 2019 [56]. Among people diagnosed with colorectal cancer by means of the screening program in 2019, almost 50% present with stage I cancer and <10% with stage IV cancer, whereas among people diagnosed outside the screening program, about 15% and 30% are diagnosed with stage I and stage IV respectively [50]. Worldwide, mortality rates have decreased up to 60% after introduction of the screening programs [57].

Link between bacterial infection and cancer - the experimental perspective

An increasing number of bacteria is being recognized as another external risk factor contributing to the onset and progression of cancer. During its infection cycle, bacteria can manipulate their hosts in various ways including the elicitation of an inflammatory response, manipulation of cell signaling pathways, and the induction of DNA damage by toxins [58]. For bacteria, the fueling of cancer is a collateral damage of the induced host cell manipulations which are essential for successful infection. The best established causal links between bacterial infection and gastrointestinal cancers are the development of gastric cancer and gallbladder carcinoma resulting from chronic *Helicobacter pylori* infection and *S*. Typhi respectively. Several *H. pylori* proteins, including CagA, CagL and VacA, stimulate host cell signaling pathways such as the ERK, Akt and Wnt/β-catenin pathways [58, 59]. These pathways play crucial roles in (embryonic) development and homeostasis, upregulation or continuous activation can lead to a perturbation in cell proliferation, malignant cell transformation and the inhibition of apoptosis [59, 60]. Part of the oncogenic potential of *S*. Typhi can be attributed to the action of its CdtB toxin. CdtB causes DNA damage in host cells and activates the host MAPK signaling pathway through a cascade of intermediate steps, thereby supporting cell transformation, the accumulation of genomic instability and prolonged survival of damaged cells [58, 61, 62].

Although nontyphoidal *Salmonella* lacks the production of CdtB toxin, various mechanisms have been proposed by which nontyphoidal *Salmonella* supports the progression of cancer. The effector protein acetyltransferase AvrA secreted by *S*. Typhimurium induces epithelial cell proliferation through activation of the β-catenin/Wnt signaling pathway, possibly in cross-talk with the STAT3 signaling pathway [58, 63]. AvrA inhibits phosphorylation of β-catenin, leading to cytosolic accumulation of β-catenin and translocation to the nucleus where it activates the transcription of oncogenes including c-MYC and IL-8 (Figure 4a) [63, 64].

Moreover, AvrA represses both an inflammatory response as well as apoptosis by blocking the MAPK pathway and the downstream JNK and NF-κB signaling pathways (Figure 4a) [65, 66]. Alternatively, several Sop effector proteins provoke the malignant transformation of cells by activating the MAPK, STAT3 and Akt signaling pathways in a multistep process (Figure 4b) [58, 67, 68].

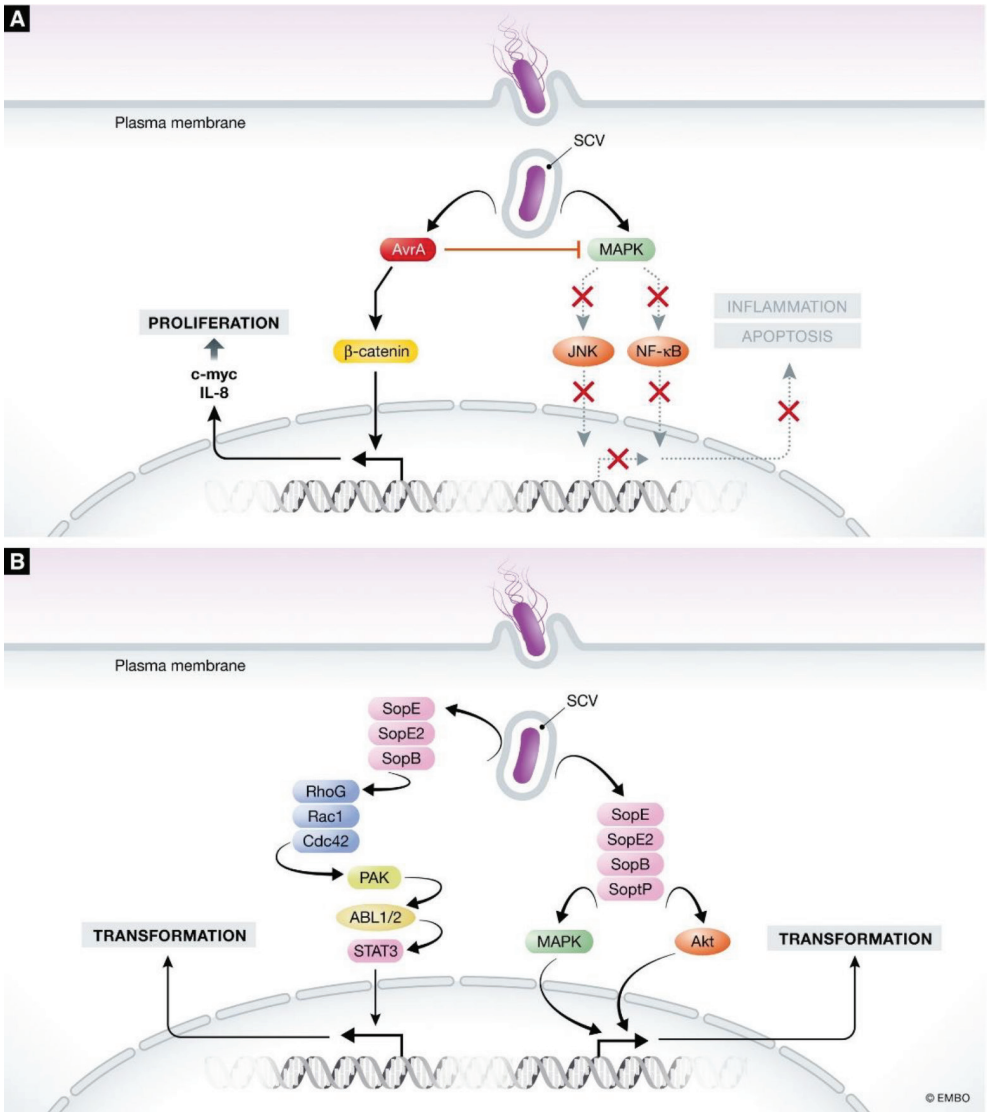


Figure 4. Examples of nontyphoidal *Salmonella* effector proteins involved in cellular transformation. Retrieved from Van Elsland, 2018 [58].

Pre-existing alterations in host cell processes, such as overexpression of the oncogene c-MYC, can act as predisposing factors for development of colon cancer after *Salmonella* infection. c-MYC plays an important role in promoting growth and proliferation of cells, inhibition of cell differentiation and apoptosis. Upregulation of c-MYC has been associated with many human cancers [69]. *In vivo* research demonstrated that mice heterozygous for APC (*Apc*$^{+/min}$, leading to c-MYC overexpression), developed colon tumors after oral infection with *S*. Typhimurium [69]. Unlike in humans, infection of mice with nontyphoidal *Salmonella* elicits a systemic infection resembling typhoid fever, rather than intestinal inflammation [71, 72]. However, pretreatment of mice with streptomycin depletes the normal gut flora and subsequent infection with *S*. Typhimurium allows for gut colonization and colitis, with symptoms similar human nontyphoidal *Salmonella* infections [72]. In contrast to mice infected with wild-type *S*. Typhimurium, mice infected with a mutated strain lacking the T3SS did not develop colon carcinomas [70]. Similarly, c-MYC overexpression was more frequently found in gallbladder tumors of cancer patients from India, a region with high *S*. Typhi incidence, as compared to Dutch gallbladder cancer patients generally lacking exposure to *S*. Typhi [70]. In the same study, the effect of inactivation of the tumor suppressor protein TP53, which is one of the most frequently observed mutations in gallbladder carcinoma patients, was assessed in a cell model [70, 73]. Single cells from murine gallbladder organoids (i.e. a simplified and miniature 3D *in vitro* version of the gallbladder organ) with inactive TP53 were cultured after being infected with either wild-type or mutant *S*. Typhimurium that cannot enter effectors for manipulation in host cells. In contrast to organoids infected with mutant *Salmonella*, organoids infected with wild-type *Salmonella* did not require growth supplements to form new organoids [70]. Also, the wild-type *Salmonella*-infected organoids had enlarged irregular nuclei and lost their cohesion and polarity, altogether demonstrating the transforming potential of *Salmonella* in predisposed cells even in absence of inflammation and immune responses [70]. The prerequisite for both TP53 inactivation and c-MYC overexpression was confirmed in an experiment where mouse embryonic fibroblasts (MEFs) were infected with *S*. Typhimurium or *S*. Typhi. MEF cells comprise a simple cell model extensively used in biomedical research for studying signaling pathways, inflammation and immunity. The MEFs harboring TP53 deficiency and overexpressed c-MYC formed significant more colonies in a soft agar assay, as compared to the MEFs with only one or none of these mutations, with similar results for *S*. Typhi and *S*. Typhimurium. Predisposed MEFs were able to maintain the transformed state when *Salmonella* was removed with antibiotics after several hours of infection [70]. The induction of cellular transformation and tumor formation in respectively MEFs and mice which harbor mutations resembling the mutations found in human gallbladder tumors, shows that *Salmonella* is able to induce cancer in these model systems under certain conditions.

Link between bacterial infection and cancer - the epidemiological perspective

The causal evidence of the contribution of bacteria to formation of malignancies is substantiated by epidemiological studies. The most studied association comprises the risk of gastric cancer (specifically mucosa associated lymphoid tissue [MALT] lymphoma) after *H. pylori* infection, with the earliest reports of a correlation dating back to the 1990s [74]. The association between *H. pylori* and MALT gastric lymphoma is restricted to strains expressing specific genes, including the *cagA* and *vacA* genes encoding for toxins which enhance the pathogenicity of *H. pylori* [75]. While estimates of the magnitude of gastric cancer risk after cagA-/vacA-positive *H. pylori* infection differ strongly between studies/countries, pooled risk estimates show over two-fold excess risk of gastric cancer [75]. Currently, *H. pylori* is the only bacterium which is officially declared as carcinogenic agent by the International Agency for Research on Cancer [76]. Yet, a plethora of research has been published postulating an association between other bacterial infections and cancer. Examples of these include *Chlamydia trachomatis* and human papillomavirus co-infection and cervical cancer [77], as well as some commensal bacteria such as *Fusobacterium nucleatum* and *Bacteroides fragilis* associated colorectal cancer [78]. With regard to *Salmonella*, multiple studies revealed an 4-5 times increased risk of gallbladder cancer associated with (chronic) infection with *S*. Typhi [79]. However, the epidemiological association between nontyphoidal *Salmonella* and cancer is limited to three studies [80-82]. Antibody titers against a *Salmonella*-specific antigen (flagellar subunit protein [FliC]) appeared significantly higher in colorectal cancer patients as compared to controls [80], whereas traces of *S*. Typhimurium were found in the tumor tissue and adjacent tissue of gallbladder patients [81]. Another study focused on the risk of colon cancer after *Salmonella* infection [82]. This nation-wide registry study found a significant association between a notified *Salmonella* infection and risk of colon cancer later in life among Dutch residents. The association particularly concerned the proximal part of the colon (1.5-fold increased risk) and was most pronounced for people infected between 20-60 years of age (2.1-fold increased risk). Also, the magnitude of the risk was serovar-dependent, with the highest observed incidences of proximal colon cancer after *S*. Enteritidis infection (3-fold increased risk) [82].

Knowledge gaps and thesis aims

Although the association between bacteria and cancers of the gastrointestinal tract is being gradually unraveled, several knowledge gaps appear from the above paragraphs. First, from an epidemiological perspective, there is only one large cohort study available in the

literature that assesses the association between nontyphoidal *Salmonella* infection and colon cancer. Therefore, substantiating this association with a similar study design applied to an independent cohort seems necessary. Moreover, this cohort study was based on notified *Salmonella* infections, which generally include patients that experienced a relatively severe course of the disease. However, people acquire multiple *Salmonella* infections throughout life, mostly a- or pauci-symptomatically. Whether these 'mild' infections also contribute to colon cancer development is yet to know. Similarly, the risk of gallbladder or biliary tract cancer after nontyphoidal *Salmonella* infection has never been studied before. From a causal perspective, limited knowledge is available about the drivers of the potential tumorigenic potential of *Salmonella* and whether this potential is restricted to the expression of specific virulence factors within certain serovars or even strains. Hence, unraveling the role of bacteria in the multistep process of developing cancer contributes to improve our understanding of the complex etiology of the disease. Ultimately, this might serve as input for improving early detection of cancer and reducing its burden. The general objective of this thesis is therefore to elucidate the role of nontyphoidal *Salmonella* infection in the development of cancers in the gastrointestinal tract, with particular focus on colon cancer. To this end, we used both experimental and epidemiological approaches to assess the association between *Salmonella* infection and gastrointestinal cancer.

Thesis outline

This thesis contains nine chapters. After this general introduction (**Chapter 1**), we present the results of a large epidemiological study assessing the occupational risk of (severe) salmonellosis and campylobacteriosis in the Dutch employed population, using data from several linked nationwide registries (**Chapter 2**). Here, the risk of salmonellosis and campylobacteriosis was assessed across the whole spectrum of registered occupations, including pre-defined groups of professions with a potentially higher degree of exposure to zoonotic gastrointestinal pathogens, such as *Salmonella*. We also verified whether the observed patterns in reported salmonellosis and campylobacteriosis among the occupational groups could be explained by differences in the magnitude of exposure to these pathogens, as defined by serological data. In **Chapter 3**, we present another nationwide registry study in which we assessed the association between occupation and colon cancer. Here we specifically emphasized those jobs with possible long-term (low-dose) exposure to zoonotic gastrointestinal pathogens by using the risk groups defined in the previous chapter. In **Chapter 4,** we studied the effect of repeated (low-dose) infections on colon cancer risk from both an epidemiological and experimental *in vivo* and *in vitro* perspective. Here we

used serological data from a large cross-sectional serosurvey in the Netherlands to assess whether colon cancer diagnosis is more likely to happen when it is preceded by (a period of) higher exposure to *Salmonella* earlier in life. Subsequently, we present the results of a study assessing the impact of *Salmonella* infection caused by multiple low-dose *versus* a single high dose of *Salmonella* bacteria in genetically predisposed mice on the formation and characteristics of colon tumors. Third, the effect of repeated exposure to *Salmonella* was studied at the cellular level using colony formation in an *in vitro* model. In **Chapter 5,** we investigated the risk of colon cancer after a (severe) nontyphoidal *Salmonella* infection in the Danish population in a nationwide population-based registry study. This study served as an corroboration of the previous Dutch cohort study in another independent Western population. In **Chapter 6**, we explored the association between severe nontyphoidal *Salmonella* or *Campylobacter* infection and biliary tract cancer in the Dutch population. For this registry-based epidemiological study, we linked the data from the nationwide salmonellosis and campylobacteriosis surveillance systems to biliary tract cancer diagnosis data in the Netherlands. **Chapter 7** summarizes the methodological characteristics and main outcomes of epidemiological studies assessing the associations between (commensal or pathogenic) bacteria and parasites and cancers in the gastrointestinal tract. In **Chapter 8**, we present the results of a case-control study including people with a history of severe nontyphoidal *Salmonella* infection who either developed proximal colon cancer later in life (i.e. the cases) or did not (i.e. the controls). The study aimed at exploring the genetic basis for the observed differences in the expression of virulence factors, nutrient metabolism and transformation capacities among the *Salmonella* strains isolated from the cases and controls. Lastly, we summarized and discussed the findings of this thesis in light of previous literature and the implications and future perspectives for this rapidly evolving field of science (**Chapter 9**).

References

1. Troeger C, Blacker BF, Khalil IA, Rao PC, Cao S, Zimsen SR, et al. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Infectious Diseases. 2018;18(11):1211-28.
2. Lagerweij GR, Pijnacker R, Friesema IHM, Mughini Gras L, Franz E. Disease burden of food-related pathogens in the Netherlands, 2019. Rijksinstituut voor Volksgezondheid en Milieu (RIVM). 2020;0117.
3. Klous G, Huss A, Heederik DJ, Coutinho RA. Human–livestock contacts and their relationship to transmission of zoonotic pathogens, a systematic review of literature. One Health. 2016;2:65-76.
4. De Wit MAS, Koopmans MPG, Kortbeek LM, Wannet WJB, Vinje J, van Leusden F, et al. Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. American Journal of Epidemiology. 2001;154(7):666-74.
5. Brenner FW, Villar RG, Angulo FJ, Tauxe R, Swaminathan B. *Salmonella* nomenclature. Journal of Clinical Microbiology. 2000;38(7):2465-7.
6. Eng SK, Pusparajah P, Ab Mutalib NS, Ser HL, Chan KG, Lee LH. *Salmonella*: a review on pathogenesis, epidemiology and antibiotic resistance. Frontiers in Life Science. 2015;8(3):284-93.
7. Gal-Mor O, Boyle EC, Grassl GA. Same species, different diseases: how and why typhoidal and non-typhoidal *Salmonella* enterica serovars differ. Frontiers in Microbiology. 2014;5:391.
8. Mughini-Gras L, Heck M, van Pelt W. Increase in reptile-associated human salmonellosis and shift toward adulthood in the age groups at risk, the Netherlands, 1985 to 2014. Eurosurveillance. 2016;21(34):30324.
9. Knodler LA. *Salmonella* enterica: living a double life in epithelial cells. Current Opinion in Microbiology. 2015;23:23-31.
10. Amani SA, Lang ML. Bacteria that cause enteric diseases stimulate distinct humoral immune responses. Frontiers in Immunology. 2020;11.
11. Viswanathan VK, Hodges K, Hecht G. Enteric infection meets intestinal function: how bacterial pathogens cause diarrhoea. Nature Reviews Microbiology. 2009;7(2):110-9.
12. Jajere SM. A review of *Salmonella* enterica with particular focus on the pathogenicity and virulence factors, host specificity and antimicrobial resistance including multidrug resistance. Veterinary World. 2019;12(4):504.
13. Srikanth CV, Mercado-Lubo R, Hallstrom K, McCormick BA. *Salmonella* effector proteins and host-cell responses. Cellular and Molecular Life Sciences. 2011;68(22):3687-97.
14. Arya G, Holtslander R, Robertson J, Yoshida C, Harris J, Parmley J, et al. Epidemiology, pathogenesis, genoserotyping, antimicrobial resistance, and prevention and control of non-typhoidal *Salmonella* serovars. Current Clinical Microbiology Reports. 2017;4(1):43-53.
15. Azimi T, Zamirnasta M, Sani MA, Dallal MMS, Nasser A. Molecular mechanisms of *Salmonella* effector proteins: a comprehensive review. Infection and Drug Resistance. 2020;13:11.
16. Haraga A, Ohlson MB, Miller SI. Salmonellae interplay with host cells. Nature Reviews Microbiology. 2008;6(1):53-66.
17. Santos RL, Raffatellu M, Bevins CL, Adams LG, Tükel Ç, Tsolis RM, et al. Life in the inflamed intestine, *Salmonella* style. Trends in Microbiology. 2009;17(11):498-506.

18. Mughini-Gras L, Pijnacker R, Duijster J, Heck M, Wit B, Veldman K, et al. Changing epidemiology of invasive non-typhoid *Salmonella* infection: a nationwide population-based registry study. Clinical Microbiology and Infection. 2020;26(7):941-e9.

19. Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) Results. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2020. Available from http://ghdx.healthdata.org/gbd-results-tool.

20. Vlaanderen F, Cuperus T, Keur I, De Rosa M, Rozendaal H, Maassen K. Staat of zoonoses 2019 [in Dutch]. Editor. 2020, RIVM: Bilthoven. Vlaanderen F, Cuperus T, Keur I, De Rosa M, Rozendaal H, Friesema I, et al. Staat van Zoönosen 2020. Rijksinstituut voor Volksgezondheid en Milieu (RIVM). 2021;0190.

21. Falkenhorst G, Simonsen J, Ceper TH, van Pelt W, de Valk H, Sadkowska-Todys M, et al. Serological cross-sectional studies on salmonella incidence in eight European countries: no correlation with incidence of reported cases. BMC Public Health. 2012;12(1):1-9.

22. Havelaar AH, Ivarsson S, Löfdahl M, Nauta MJ. Estimating the true incidence of campylobacteriosis and salmonellosis in the European Union, 2009. Epidemiology & Infection. 2013;141(2):293-302.

23. Hendriksen RS, Vieira AR, Karlsmose S, Lo Fo Wong DM, Jensen AB, Wegener HC, et al. Global monitoring of Salmonella serovar distribution from the World Health Organization Global Foodborne Infections Network Country Data Bank: results of quality assured laboratories from 2001 to 2007. Foodborne Pathogens and Disease. 2011;8(8):887-900.

24. European Food Safety Authority, & European Centre for Disease Prevention and Control. The European Union One Health 2019 Zoonoses Report. EFSA Journal, 2021;19(2):e06406.

25. Pires SM, Vigre H, Makela P, Hald T. Using outbreak data for source attribution of human salmonellosis and campylobacteriosis in Europe. Foodborne Pathogens and Disease. 2010;7(11):1351-61.

26. Danyluk MD, Goodrich-Schneider RM, Schneider KR, Harris LJ, Worobo RW. Outbreaks of foodborne disease associated with fruit and vegetable juices, 1922–2010. EDIS. 2012(3).

27. Mughini-Gras L, Enserink R, Friesema I, Heck, M, van Duynhoven Y, van Pelt W. Risk factors for human salmonellosis originating from pigs, cattle, broiler chickens and egg laying hens: a combined case-control and source attribution analysis. PloS one. 2014;9(2):e87933.

28. Mughini-Gras L, Franz E, van Pelt W. New paradigms for Salmonella source attribution based on microbial subtyping. Food Microbiology, 2018;71:60-67.

29. Cooper DM, Sunderland MA. (2000). The cell: A molecular approach. 2nd edition. Sinauer Associates.

30. ACS, American Cancer Society. (2021). Cancer staging. [Accessed at 2021 December 6]. Available at: https://www.cancer.org/treatment/understanding-your-diagnosis/staging.html

31. Wong HH, Chu P. Immunohistochemical features of the gastrointestinal tract tumors. Journal of Gastrointestinal Oncology. 2012;3(3):262.

32. Lee GH, Malietzis G, Askari A, Bernardo D, Al-Hassi HO, Clark SK. Is right-sided colon cancer different to left-sided colorectal cancer?–a systematic review. European Journal of Surgical Oncology (EJSO). 2015;41(3):300-8.

33. Di Pardo BJ, Bronson NW, Diggs BS, Thomas CR, Hunter JG, Dolan JP. The global burden of esophageal cancer: a disability-adjusted life-year approach. World Journal of Surgery, 2016;40(2):395-401.

34. Broder MS, Ailawadhi S, Beltran H, Blakely LJ, Budd GT, Carr L, et al. Estimates of stage-specific preclinical sojourn time across 21 cancer types. Journal of Clinical Oncology. 2021;39:no.15_suppl.

35. Luebeck EG, Curtius K, Jeon J, Hazelton WD. Impact of tumor progression on cancer incidence curves. Cancer Research. 2013;*73*(3):1086-96.

36. Simon K. Colorectal cancer development and advances in screening. Clinical Interventions in Aging. 2016;11:967.

37. Mik M, Berut M, Dziki L, Trzcinski R, Dziki A. Right-and left-sided colon cancer–clinical and pathological differences of the disease entity in one organ. Archives of Medical Science: AMS. 2017;13(1):157.

38. Labianca R, Nordlinger B, Beretta GD, Mosconi S, Mandalà M, Cervantes A, et al. Early colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals of Oncology. 2013;24:vi64-vi72.

39. Tamas K, Walenkamp AME, de Vries EGE, van Vugt MATM, Beets-Tan RG, van Etten B, et al. Rectal and colon cancer: Not just a different anatomic site. Cancer Treatment Reviews. 2015;41(8):671-9.

40. Colorectal Cancer Alliance. (2021). Understanding a diagnosis by stage. [Accessed at 2021 December 6]. Available at:

41. Colorectal Cancer Stages | Colorectal Cancer Alliance (ccalliance.org)

42. Bredenoord A, Tack J, Smout A. Het Colon. In: Functiestoornissen van het maag-darmkanaal. Bohn Stafleu van Loghum, Houten. 2010. Available at: https://doi.org/10.1007/978-90-313-7840-1_7

43. Dong M, Cioffi G, Wang J, Waite KA, Ostrom QT, Kruchko C, et al. Sex differences in cancer incidence and survival: A pan-cancer analysis. Cancer Epidemiology and Prevention Biomarkers. 2020;29(7):1389-97.

44. Hagen CE, Setia N, Lauwers GY. Familial adenomatous polyposis: a review of gastrointestinal manifestations. Diagnostic Histopathology. 2015;21(4):152-60.

45. Karamchandani DM, Zhang Q, Liao XY, Xu JH, Liu XL. Inflammatory bowel disease-and Barrett's esophagus-associated neoplasia: the old, the new, and the persistent struggles. Gastroenterology Report. 2019;7(6):379-95.

46. Rubin DC, Shaker A, Levin MS. Chronic intestinal inflammation: inflammatory bowel disease and colitis-associated colon cancer. Frontiers in Immunology. 2012;3:107.

47. Shivanna LM, Urooj A. A review on dietary and non-dietary risk factors associated with gastrointestinal cancer. Journal of Gastrointestinal Cancer. 2016;47(3):247-54.

48. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians. 2021;71(3):209-49.

49. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. The Lancet. 2018;391(10125):1023-75.

50. World Health Organization International Agency for Research on Cancer (IARC). (2021). *GLOBOCAN 2020: Estimated number of new cases in 2020, worldwide, both sexes, all ages.* [Accessed at 2021 September 15]. Available from: https://gco.iarc.fr/today/

51. Arnold M, Abnet CC, Neale RE, Vignat J, Giovannucci EL, McGlynn KA, et al. Global burden of 5 major types of gastrointestinal cancer. Gastroenterology. 2020;159(1):335-49.

52. Lin L, Yan L, Liu Y, Qu C, Ni J, Li H. The burden and trends of primary liver cancer caused by specific etiologies from 1990 to 2017 at the global, regional, national, age, and sex level results from the global burden of disease study 2017. Liver Cancer. 2020;9(5):563-82.

53. IKNL, Integraal Kankercentrum Nederland. (2021). Kankersoorten – cijfers. [Accessed at 2021 December 5]. Available from: Kankersoorten (iknl.nl)
54. Volksgezondheidenzorg.info. (2021). Dikkedarmkanker – Cijfers & context. [Accessed at 2021 December 5]. Available from: https://www.volksgezondheidenzorg.info/onderwerp/dikkedarmkanker/cijfers-context/
55. Vuik FE, Nieuwenburg SA, Bardou M, Lansdorp-Vogelaar I, Dinis-Ribeiro M, Bento MJ, et al. Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. Gut. 2019;68(10):1820-6.
56. Navarro M, Nicolas A, Ferrandez A, Lanas A. Colorectal cancer population screening programs worldwide in 2016: An update. World Journal of Gastroenterology. 2017;23(20):3632.
57. RIVM. (19th August 2021). Factsheet Bevolkingsonderzoek Darmkanker 2020. Factsheet bevolkingsonderzoek darmkanker | RIVM
58. Maida M, Macaluso FS, Ianiro G, Mangiola F, Sinagra E, Hold G, et al. Screening of colorectal cancer: present and future. Expert Review of Anticancer Therapy, 2017;17(12):1131-46.
59. van Elsland D, Neefjes J. Bacterial infections and cancer. EMBO Reports. 2018;19(11):e46632.
60. Song X, Xin N, Wang W, Zhao C. Wnt/β-catenin, an oncogenic pathway targeted by *H. pylori* in gastric carcinogenesis. Oncotarget. 2015;6(34):35579.
61. Padala RR, Karnawat R, Viswanathan SB, Thakkar AV, Das AB. Cancerous perturbations within the ERK, PI3K/Akt, and Wnt/β-catenin signaling network constitutively activate inter-pathway positive feedback loops. Molecular BioSystems, 2017;13(5):830-40.
62. Zhang K, Hornef M, Fulde M. The deadly bite of Salmonella Typhi. EMBO Reports. 2015;16(8):887-8.
63. Di Domenico EG, Cavallo I, Pontone M, Toma L, Ensoli F. Biofilm producing *Salmonella* typhi: chronic colonization and development of gallbladder cancer. International Journal of Molecular Sciences. 2017;18(9):1887.
64. Lu R, Wu S, Zhang YG, Xia Y, Zhou Z, Kato I, et al. *Salmonella* protein AvrA activates the STAT3 signaling pathway in colon cancer. Neoplasia. 2016;18(5):307-16.
65. Ye Z, Petrof EO, Boone D, Claud EC, Sun J. *Salmonella* effector AvrA regulation of colonic epithelial cell inflammation by deubiquitination. The American Journal of Pathology. 2007;171(3):882-92.
66. Jones RM, Wu H, Wentworth C, Luo L, Collier-Hyams L, Neish AS. *Salmonella* AvrA coordinates suppression of host immune and apoptotic defenses via JNK pathway blockade. Cell Host & Microbe. 2008;3(4):233-44.
67. Sun J. Pathogenic bacterial proteins and their anti-inflammatory effects in the eukaryotic host. Anti-Inflammatory & Anti-Allergy Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Inflammatory and Anti-Allergy Agents). 2009;8(3):214-27.
68. Kuijl C, Savage ND, Marsman M, Tuin AW, Janssen L, Egan DA, et al. Intracellular bacterial growth is controlled by a kinase network around PKB/AKT1. Nature. 2007;450(7170):725-30.
69. Hannemann S, Gao B, Galán JE. *Salmonella* modulation of host cell gene expression promotes its intracellular growth. PLoS pathogens. 2013;9(10):e1003668.
70. Pelengaris S, Khan M. The many faces of c-MYC. Archives of Biochemistry and Biophysics. 2003;416(2):129-36.
71. Scanu T, Spaapen RM, Bakker JM, Pratap CB, Wu LE, Hofland I, et al. *Salmonella* manipulation of host signaling pathways provokes cellular transformation associated with gallbladder carcinoma. Cell Host & Microbe. 2015;17(6):763-74.

72. Santos RL, Zhang S, Tsolis RM, Kingsley RA, Adams LG, Bäumler AJ. Animal models of *Salmonella* infections: enteritis versus typhoid fever. Microbes and Infection. 2001;3(14-15):1335-44.

73. Barthel M, Hapfelmeier S, Quintanilla-Martínez L, Kremer M, Rohde M, Hogardt M, et al. Pretreatment of mice with streptomycin provides a *Salmonella* enterica serovar Typhimurium colitis model that allows analysis of both pathogen and host. Infection and Immunity. 2003;71(5):2839-58.

74. Espinoza JA, García P, Bizama C, Roa JC. Molecular genetics of gallbladder cancer. eLS; 2015:1-7.

75. Nomura A, Stemmermann GN, Chyou PH, Kato I, Perez-Perez GI, Blaser MJ. Helicobacter pylori infection and gastric carcinoma among Japanese Americans in Hawaii. New England Journal of Medicine, 1991;325(16):1132-6.

76. Ghotaslou R, Leylabadlo HE, Nasiri MJ, Dabiri H, Hashemi A. Risk of gastric cancer in association with Helicobacter pylori different virulence factors: A systematic review and meta-analysis. Microbial Pathogenesis. 2018;118:214-9.

77. IARC. (2012). Biological Agents - IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Volume 100B, 385-435.

78. Karim S, Souho T, Benlemlih M, Bennani B. (2018). Cervical cancer induction enhancement potential of *Chlamydia trachomatis*: a systematic review. Current Microbiology. 2018;75(12):1667-74.

79. Viljoen KS, Dakshinamurthy A, Goldberg P, Blackburn JM. Quantitative profiling of colorectal cancer-associated bacteria reveals associations between fusobacterium spp., enterotoxigenic Bacteroides fragilis (ETBF) and clinicopathological features of colorectal cancer. PloS one. 2015;10(3):e0119462.

80. Koshiol J, Wozniak A, Cook P, Adaniel C, Acevedo J, Azócar L, et al. *Salmonella* enterica serovar Typhi and gallbladder cancer: a case–control study and meta-analysis. Cancer Medicine. 2016;5(11):3310-35.

81. Kato I, Boleij A, Kortman GA, Roelofs R, Djuric Z, Severson RK, et al. Partial associations of dietary iron, smoking and intestinal bacteria with colorectal cancer risk. Nutrition and Cancer. 2013;65(2):169-77.

82. Iyer P, Barreto SG, Sahoo B, Chandrani P, Ramadwar MR, Shrikhande SV, et al. Non-typhoidal *Salmonella* DNA traces in gallbladder cancer. Infectious Agents and Cancer. 2016;11(1):1-4.

83. Mughini-Gras L, Schaapveld M, Kramers J, Mooij S, Neefjes-Borst EA, Pelt WV, et al. Increased colon cancer risk after severe *Salmonella* infection. PloS one. 2018;13(1):e0189721.

Chapter 2

OCCUPATIONAL RISK OF SALMONELLOSIS AND CAMPYLOBACTERIOSIS: A NATION-WIDE POPULATION-BASED REGISTRY STUDY

J.W. Duijster
E. Franz
J.J.C. Neefjes
L. Mughini-Gras

Published in: *Occupational & Environmental Medicine*, 2019: 76(9)

Abstract

Objectives: Occupational exposure to animals and foods thereof is a poorly characterized risk factor for salmonellosis and campylobacteriosis, the main causes of bacterial gastroenteritis in the Western world. We performed a population-based registry study in the Netherlands to assess whether differences exist in the incidence of reported salmonellosis and campylobacteriosis cases among occupational groups, and whether they can be explained by differences in the magnitude of exposure to these pathogens, as defined by serology.

Methods: Person-level occupational data for all Dutch residents were linked to lab-confirmed salmonellosis and campylobacteriosis data, and to serological data from a previous national serosurvey. SIRs for salmonellosis and campylobacteriosis among occupational sectors and specific high-risk occupations were calculated based on the total employed population. Moreover, *Salmonella* and *Campylobacter* seroincidence rates were compared among sectors and high-risk occupations.

Results: Occupational exposure to live animals or manure and working in the sale of animal-derived food products were associated with significantly increased risks of salmonellosis (SIR 1.55–1.82) and campylobacteriosis (SIR 1.36–1.65). Moreover, incidences were significantly higher in specific industrial sectors, as well as healthcare and social work sectors. Mean seroincidence rates ranged from 1.28 to 2.30 infections/person-year for *Campylobacter*, and from 0.36 to 0.99 for *Salmonella*, with only slightly higher rates for people in high-risk occupations.

Conclusions: Significant differences in reported salmonellosis and campylobacteriosis incidence exist among occupational sectors, with the highest incidence in those persons occupationally exposed to live animals. These differences are only partially reflected in the serology.

Background

Salmonella and *Campylobacter* are the main causes of bacterial gastroenteritis in the Western world, including the European Union [1]. In the Netherlands, the annual number of salmonellosis cases is estimated at ~27 000, whereas for campylobacteriosis, this number is twofold to threefold higher [2]. In terms of disability-adjusted life years (DALY), both pathogens are estimated to cause altogether over 4000 DALYs in the Netherlands every year, with the associated cost amounting to ~€75 million/year. Such burden is mainly attributable to possible sequelae beyond gastroenteritis (i.e., Guillain-Barré syndrome, reactive arthritis, irritable bowel syndrome and inflammatory bowel disease) [3, 4]. Surveillance of salmonellosis and campylobacteriosis in the Netherlands is based on voluntary reporting of a network of diagnostic laboratories capturing mainly cases with more severe symptoms. Besides the main and extensively researched route of transmission via food, *Salmonella* and *Campylobacter* infections may be acquired through contact with animals or manure [5, 6]. The risk of *Salmonella* or *Campylobacter* transmission via contact with animals or manure has been shown to be significant in specific cohorts, including people occupationally exposed to live animals or animal-derived products (e.g., farmers and abattoir workers) [7, 8]. Studies assessing salmonellosis and campylobacteriosis incidence across different types of occupations on a national level are scarce [9]. Therefore, the aim of this study was to assess whether differences exist in the incidence of reported salmonellosis and campylobacteriosis cases among occupational groups, and whether they can be explained by differences in the magnitude of exposure to these pathogens, as defined by serology.

Methods

Data and study population

We linked two national registries and a national serosurvey in the Netherlands (~17 million inhabitants). One registry included deidentified person-level data on occupation as derived from Statistics Netherlands (CBS), which records the occupation of all Dutch residents at any moment in time based on tax returns. Occupations are coded based on the European Nomenclature of Economic Activities (NACE) (second revision) classification of productive economic activities, which is part of the integrated International Standard Industrial Classification of All Economic Activities (ISIC) system [10]. NACE codes consist of five digits, allowing for four hierarchical levels (i.e., sections, divisions, groups, classes) [10]. At the time of analysis, the data set included data for 12 566 846 individuals of legal working age with recorded type of occupation (5-digit level) and the dates of start and end of employment, between January 1999 and December 2016.

The second data set contained data on reported human salmonellosis and campylobacteriosis cases in the Netherlands, derived from the national laboratory surveillance network coordinated by the Dutch National Institute for Public Health and the Environment. Estimated population coverage is 64% and 52% for salmonellosis and campylobacteriosis, respectively [11]. At the time of analysis, the data set included 27 425 records of culture-confirmed non-typhoid *Salmonella* infection and 31 855 records of culture-confirmed *Campylobacter* infections (both among outpatients and hospitalized patients), with relevant metadata (i.e., gender, birth date, residence location). The salmonellosis data set contained data between January 1999 and December 2016, whereas campylobacteriosis data between January 2004 and December 2016.

The third data set contained serological data for *Salmonella* and *Campylobacter* from participants of a population-based cross-sectional serosurvey in the Netherlands in 2006–2007. This serosurvey has already been presented in detail before [12]. Briefly, participants provided a blood sample and completed an epidemiological questionnaire. In total, serum samples from 7904 individuals were available, 1304 of which were tested for anti-*Salmonella* and anti-*Campylobacter* IgA, IgM and IgG concentrations (optical density values) using a mixed ELISA based on lipopolysaccharides of *S*. Enteritidis and *S*. Typhimurium [13, 14] and an acid glycine extract of *C*. *jejuni* strain SSDZ-01 [15] as capture antigens. This data set has been used in several previous studies on immunodynamic modelling of *Salmonella* [16, 17] and *Campylobacter* infections [18–20].

Data linkage and exposure/outcome definition

All three data sets were transferred to CBS, which acted as trusted third party for data anonymization by adding a unique Record Identification Number (RIN), based on persons' gender, birth date and residence location. On generation of the RINs, all the personal identifiers were removed, and the RINs were then used for linkage to the study data [21]. We limited the analyses to people aged 16–69 years, as compulsory education applies until 16 years and almost all people retire by the age of 70 years. The data set was cleared from duplicate isolations of the same *S. enterica* subsp. *enterica* serovar or *Campylobacter* sp. within 3 months after the initial infection. Participants of the serosurvey with start of employment after the sampling date were excluded. We excluded also participants who ended employment >1 year before the sampling date, to account for waning immunity, leaving 733 participants with *Salmonella* and/or *Campylobacter* serology data for analysis.

The structure of the NACE framework allows for analysis at different classification levels. We performed the first analysis at the division level where all occupations are mutually

classified into 86 divisions, hereafter also referred to as 'sectors'. Due to revision of the NACE classification in 2008, some occupations could not be classified into a single sector in the period before the revision; hence, these were excluded from the analysis. The serology analysis was performed at section level due to sample size constraints. In total, 21 sections exist, each containing one or multiple sectors (mutually classified). To test for differences between *Salmonella* serovars, these were classified as *S*. Typhimurium and its monophasic variant (35.3%), *S*. Enteritidis (35.8%), and other serovars (28.9%). Based on the type of sample the *Salmonella* isolate originated from (i.e., faeces, blood, urine, and so on), *Salmonella* infections were classified as enteric (faeces, 91.1%), septicaemic (blood, 3.7%) or others (mostly urinary tract and wound infections, 5.2%). For *Campylobacter*, the analysis was limited to the most frequently reported species in the Netherlands: *C. jejuni* (92.9%) and *C. coli* (7.1%); further information on the *Campylobacter* isolates was not available.

Statistical analysis

Time at risk (age ≥16) started at the date of employment and ended at the date of first reported *Salmonella* or *Campylobacter* infection, end date of that employment (when this corresponded to the start of an unemployment period) or the end of the study period (1 January 2017), whichever occurred first. As long as no *Salmonella* or *Campylobacter* infection occurred, individuals were allowed to re-enter the study cohort at any point in time in case of intermittent employment periods and shifts between occupations, and they could be included in multiple sectors (either subsequently or simultaneously). Accounting for an average reporting delay of 3 weeks [11], the reporting date of *Salmonella* or *Campylobacter* infection minus 3 weeks was used for analysis. An event was therefore defined as a reported salmonellosis or campylobacteriosis case with estimated date of infection during an employment period. Separate analyses were performed for *Salmonella* and *Campylobacter*, allowing for occurrence of both infections in one individual.

Incidence rates (IR) per 100 000 person-years at risk of salmonellosis and campylobacteriosis in the employed population were calculated by *Salmonella* serovar and type of infection, *Campylobacter* sp., gender, age (5-year bands) and calendar year. SIRs for salmonellosis and campylobacteriosis were calculated for each occupational sector by dividing the observed number of reported *Salmonella* and *Campylobacter* infections by the expected number of infections based on the IRs in the employed population (matched by gender, age and calendar year); 95% CIs were estimated based on Poisson distribution of person-time data. For sectors with significantly increased or decreased SIRs and ≥10 cases, analyses were stratified by *Salmonella* serovar, *Campylobacter* sp., gender and age group (16–19, 20–29, 30–39, 40–49,

≥50 years). Next, based on the 5-digit NACE codes, we classified 42 occupations with potential risk of occupational exposure to *Salmonella* and/or *Campylobacter* into three specific groups (Supplementary table S1). Those risk groups entailed occupations with possible contact with live animals or manure (e.g., farmers and abattoir workers), occupations in food production/preparation (e.g., bakers, cooks/chefs) and occupations in sale of animal-derived products (e.g., butchers). Overall and stratified SIRs of salmonellosis and campylobacteriosis were calculated for each of these high-risk groups using the rates in the employed population as reference. Cumulative incidence plots with years of employment as timescale were made to graphically display the risk of infection in the risk groups versus the employed population.

Seroincidence rates for *Salmonella* and *Campylobacter* infections, defined as the average number of infections per person-year, were determined based on the optical density values of serum IgA, IgG and IgM as described in detail elsewhere [20]. Briefly, we used the European Centre for Disease Prevention and Control's (ECDC) seroincidence calculator tool (https://ecdc.europa.eu/en/publications-data/seroincidence-calculator-tool), which uses the combination of IgG, IgM and IgA values at a given point in time to estimate the time since seroconversion, thereby providing an estimate of the annual 'force of infection' for each individual using a Bayesian back-calculation model. This model is based on the kinetics of IgG, IgM and IgA observed during previous longitudinal studies of adult patients with stool culture-confirmed *Salmonella* or *Campylobacter* infections, which provided reference values for peak levels and decay rates of Ig concentrations and their relationship over time. Following the analytical approach of Monge *et al* [18], we tested for differences in log-transformed seroincidence rates between sections using a multivariate linear regression model including also gender and years of employment as covariates. We then compared seroincidences in high-risk occupations with those of other occupations. All statistical analyses were performed using STATA V.14.2. P values <0.05 were considered statistically significant.

Results

Cohort description

The cohort consisted of 12 566 831 individuals aged 16–69 years employed between 1999 and 2016 (Supplementary figure S1). People entering the cohort after the start of the study (1 January 1999) had a median age of 17 years (IQR: 16–29). Overall, 8220 individuals with a reported *Salmonella* infection during employment were observed, corresponding to an IR of 6.51 infections per 100 000 person-years at risk (95% CI 6.36 to 6.65). Supplementary

table S2 shows the IRs of salmonellosis by serovar, type of infection, gender and age group. Highest IRs were observed for age groups 16–19 years (IR: 12.72, 95% CI 11.94 to 13.50) and 20–29 years (IR: 10.85, 95% CI 10.45 to 11.24). Infection occurred after a median of 5 years of registered employment (IQR: 2–9).

For *Campylobacter*, the study period was limited to 2004–2016, with a total of 11 615 429 people in the cohort, of which 14 352 with a reported *Campylobacter* infection. The overall IR was 15.54 infections per 100 000 person-years at risk (95% CI 15.29 to 15.79). As for salmonellosis, the IRs for campylobacteriosis were higher in the younger age groups (Supplementary table S2). The median time of registered employment at infection was 5 years (IQR: 2–8).

Increased occupational risks

Among the 86 sectors, a median of 31 *Salmonella* infections (IQR: 11–87) and 53 Campylobacter infections (IQR: 19–149) were reported. Supplementary table S3 shows the SIRs of salmonellosis and campylobacteriosis per sector. Among sectors with ≥10 reported cases, 12 sectors showed a significantly increased SIR for salmonellosis, campylobacteriosis or both (Table 1, Supplementary tables S4–S15). The highest SIRs were observed for the sector 'veterinary activities', with a twofold increased risk for salmonellosis (SIR 2.03, 95% CI 1.22 to 3.37) and campylobacteriosis (SIR 1.96, 95% CI 1.33 to 2.87). Most reported cases within this sector were female (salmonellosis: 80.0%; campylobacteriosis: 84.6%) and aged 20–29 years (Supplementary table S4). Increased SIRs, mainly for campylobacteriosis, were found in five industrial sectors, including the manufacturing of chemicals, paper and machinery, and the extraction and supply of petroleum, gas and electricity. Among these sectors, SIRs were only significant for males (salmonellosis: 1.40–1.42; campylobacteriosis: 1.44–2.59) and people aged ≥30 years. Within the 'other manufacturing' sector, reported cases were mostly attributable to the occupation 'social employment' (*Salmonella* n=165; *Campylobacter* n=301), whereas the other occupations within this sector had <15 reported cases each. Social employment includes customised and supervised occupations for people with physical or mental disabilities. Most cases were reported among people being ≥40 years (salmonellosis: 69.4%; campylobacteriosis: 77.7%). Marginally increased risks were also observed among healthcare and social workers (mean SIR 1.13 salmonellosis; mean SIR 1.17 campylobacteriosis), with most cases being females (77.6%–92.5% salmonellosis; 76.7%–90.2% campylobacteriosis). SIRs for salmonellosis were highest in the youngest age group (16–19 years), whereas this was not the case for campylobacteriosis (Supplementary tables S11–S13). In the healthcare sector, most cases were reported among people working

within hospitals (salmonellosis: n=413; campylobacteriosis: n=833). Within 'residential care activities', which includes occupations in nursing homes, psychiatric hospitals, home care for elderly and disabled people, reported infections were evenly distributed across occupations, with SIRs higher for males (Supplementary table S12). Furthermore, an increased risk for campylobacteriosis was found for the 'accommodation' sector (e.g., hotels and campsites), with highest risks in the younger age groups (Supplementary table S14).

Table 1. Sectors with significantly increased or decreased standardized incidence ratios (SIRs) for overall salmonellosis and/or campylobacteriosis.

	Salmonellosis			**Campylobacteriosis**		
Sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Higher SIRs						
Veterinary activities	15	7.4	2.03 (1.22-3.37)**	26	13.3	1.96 (1.33-2.87)**
Manufacture of chemicals	59	42.6	1.38 (1.07-1.79)*	112	81.0	1.38 (1.15-1.66)**
Manufacture of paper (products)	15	19.5	1.03 (0.64-1.66)	51	34.1	1.50 (1.14-1.97)**
Extraction of crude petroleum and natural gas	6	3.4	1.75 (0.79-3.89)	16	6.3	2.54 (1.56-4.15)***
Electricity, gas, steam and air conditioning supply	17	23.2	0.73 (0.46-1.18)	67	43.9	1.53 (1.20-1.94)**
Manufacture of machinery and equipment	59	45.5	1.30 (1.01-1.67)*	114	124.7	0.91 (0.76-1.10)
Other manufacturing	183	115.8	1.58 (1.37-1.83)***	323	230.7	1.40 (1.26-1.56)***
Human health activities	550	492	1.12 (1.03-1.21)*	1,123	918.6	1.22 (1.15-1.30)***
Residential care activities	559	478.3	1.17 (1.08-1.27)***	981	835.4	1.17 (1.10-1.25)***
Social work activities without accommodation	362	331.6	1.09 (0.98-1.21)	686	620.5	1.11 (1.03-1.19)**
Accommodation	102	95.5	1.06 (0.87-1.28)	213	155.0	1.37 (1.20-1.57)***
Activities of households as employers of domestic personnel	26	23.2	1.12 (0.76-1.65)	82	57.8	1.42 (1.14-1.76)**
Lower SIRs						
Architectural and engineering activities; technical testing and analysis[§]	79	102.8	0.77 (0.62-0.96)*	162	191.3	0.85 (0.73-0.99)*
Computer programming and consultancy[§]	80	91.9	0.87 (0.70-1.08)	149	240.5	0.62 (0.53-0.73)***
Financial service activities[§]	125	145.8	0.86 (0.72-1.02)	200	256.3	0.78 (0.68-0.90)***
Activities auxiliary to financial services and insurance activities[§]	59	58.8	1.01 (0.78-1.30)	72	100.1	0.72 (0.57-0.91)**
Activities of head offices; management consultancy activities[§]	134	141.5	0.95 (0.80-1.12)	191	279.0	0.68 (0.59-0.79)***

	Salmonellosis			Campylobacteriosis		
Sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Activities of membership organizations§	74	81.1	0.91 (0.73-1.15)	113	140.0	0.81 (0.67-0.97)*
Education	428	502.3	0.85 (0.78-0.94)**	857	934.6	0.92 (0.86-0.98)*
Crop and animal production and hunting	105	119.2	0.88 (0.73-1.07)	141	181.2	0.78 (0.66-0.92)**
Construction of buildings	64	64.1	1.00 (0.78-1.27)	141	169.7	0.83 (0.70-0.98)*
Wholesale trade	403	431.5	0.93 (0.85-1.03)	716	845.7	0.85 (0.79-0.91)***
Land transport	157	185.1	0.85 (0.73-0.99)*	365	345	1.06 (0.95-1.17)
Services to buildings and landscape activities	184	176.2	1.04 (0.90-1.21)	247	287.5	0.86 (0.76-0.97)*

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio; *p<0.05; **p<0.01; ***p<0.001; § white collar sector

Decreased occupational risks

Twelve sectors (with ≥10 cases) showed a significantly lower SIR for salmonellosis and/or campylobacteriosis (Supplementary tables S3, S16–27). SIRs were 0.77–0.85 for salmonellosis and 0.62–0.92 for campylobacteriosis (Table 1). Some of these sectors are 'white collar' sectors, which includes jobs at professional, administrative or managerial level, generally associated with a higher socioeconomic status (SES). No consistent differences were observed among age groups or gender in the white collar sectors. Within the educational sector, risk was significantly reduced only for females (SIR 0.83, 95% CI 0.74 to 0.94 salmonellosis; SIR 0.85, 95% CI 0.78 to 0.93 campylobacteriosis).

High-risk occupations

Supplementary table S28 shows the characteristics of the three high-risk groups. The group occupationally exposed to live animals or manure ('live animals') consisted of 240 993 and 172 978 people for the salmonellosis and campylobacteriosis analysis, respectively, with the majority being male (63.5%). Within this group, 93 *Salmonella* and 147 *Campylobacter* infections were reported. The second group included 2 037 210 people with occupational exposure through food production/preparation ('food production') for the salmonellosis analysis and 1 666 621 people for the campylobacteriosis analysis, with 423 and 762 salmonellosis and campylobacteriosis cases, respectively. The third group included 244 051 people involved in the sale of animal-derived food products ('food sale') for the salmonellosis analysis and 178 427 for the campylobacteriosis analysis, in which 78 salmonellosis and 109 campylobacteriosis cases were reported. Analysis of

the three risk groups showed a significantly increased risk for both salmonellosis and campylobacteriosis in the live animals group and in the food sale group (Table 2, Figure 1). For salmonellosis, the SIR was 1.82 (95% CI 1.49 to 2.23) for the live animals group and 1.55 (95% CI 1.24 to 1.93) for the food sale group. In both groups, risk was most pronounced for *S*. Typhimurium, whereas the risk for serovars other than Enteritidis and Typhimurium was not significantly elevated. SIRs were generally higher in the younger age groups (Table 2).

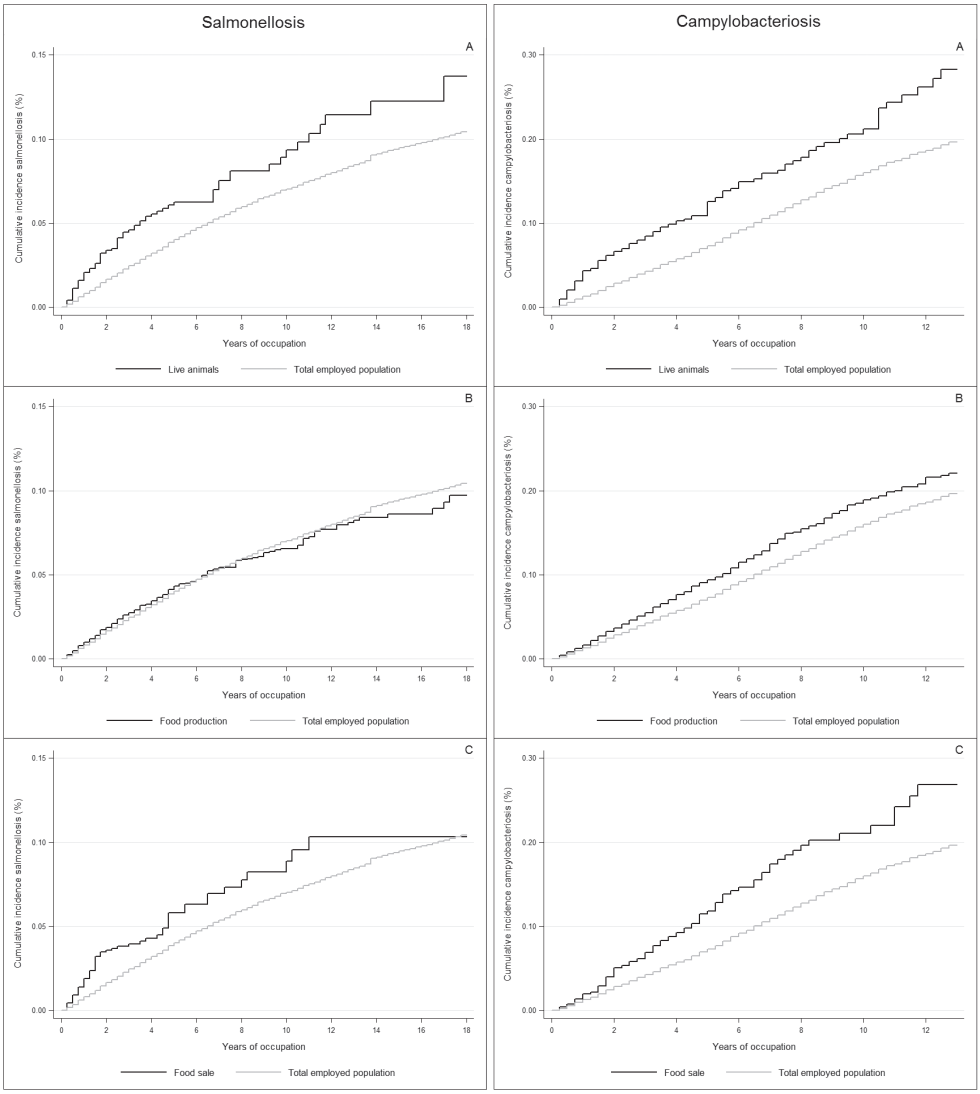


Figure 1. Cumulative incidence (CI) plots of salmonellosis and campylobacteriosis by risk group. (A) Live animals. (B) Food production. (C) Food sale.

Table 2. Standardized incidence ratios (SIRs) for salmonellosis and campylobacteriosis by risk group.

	Live animals			**Food production**			**Food sale**		
SALMONELLOSIS	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender									
Overall	93	51.5	1.82 (1.49-2.23)***	423	445.1	0.95 (0.86-1.05)	78	50.5	1.55 (1.24-1.93)***
Male	60	32.0	1.88 (1.46-2.42)***	195	202.5	0.96 (0.84-1.11)	32	19.6	1.63 (1.15-2.30)**
Female	33	19.1	1.73 (1.23-2.43)**	228	242.5	0.94 (0.83-1.07)	46	30.8	1.49 (1.12-1.99)**
***Salmonella* serovar**									
Typhimurium	38	13.6	2.79 (2.03-3.84)***	130	137.4	0.95 (0.80-1.12)	32	14.4	2.23 (1.58-3.15)***
Enteritidis	34	20.0	1.70 (1.21-2.38)**	177	167.9	1.05 (0.91-1.22)	31	20.1	1.54 (1.08-2.19)*
Other	21	17.4	1.20 (0.78-1.85)	116	139.7	0.83 (0.69-1.00)*	15	16.0	0.94 (0.57-1.56)
Type of infection									
Enteric	89	47.7	1.87 (1.52-2.30)***	399	418.5	0.95 (0.86-1.05)	74	47.4	1.56 (1.24-1.96)***
Septicemic	0	1.5	-	9	10.3	0.87 (0.45-1.68)	2	1.2	1.71 (0.43-6.83)
Other[$]	4	1.9	2.06 (0.77-5.48)	15	16.3	0.92 (0.56-1.53)	2	1.9	1.05 (0.26-4.21)
Age									
16-19 years	22	7.18	3.07 (2.02-4.66)***	155	147.7	1.05 (0.90-1.23)	30	14.3	2.09 (1.46-2.99)***
20-29 years	33	18.3	1.80 (1.28-2.53)**	164	185.0	0.89 (0.76-1.03)	26	17.9	1.45 (0.99-2.13)
30-39 years	18	9.1	1.99 (1.25-3.15)**	46	41.8	1.10 (0.82-1.47)	8	6.5	1.23 (0.61-2.45)
40-49 years	12	8.1	1.49 (0.85-2.62)	31	35.2	0.88 (0.62-1.25)	10	5.9	1.69 (0.91-3.15)
≥50 years	8	8.4	0.95 (0.47-1.90)	27	35.3	0.77 (0.53-1.12)	4	5.8	0.69 (0.26-1.84)
P-trend	<0.001			0.123			0.026		
CAMPYLOBACTERIOSIS									
Gender									
Overall	147	88.9	1.65 (1.41-1.94)***	762	744.3	1.02 (0.95-1.10)	109	80.1	1.36 (1.13-1.64)**
Male	94	57.2	1.64 (1.34-2.01)***	366	345.9	1.06 (0.96-1.17)	36	33.7	1.07 (0.77-1.48)
Female	53	31.7	1.67 (1.28-2.19)***	396	398.4	0.99 (0.90-1.10)	73	46.4	1.57 (1.25-1.98)***
***Campylobacter* species**									
Jejuni	140	82.9	1.69 (1.43-2.00)***	718	696.0	1.03 (0.96-1.11)	99	74.8	1.32 (1.09-1.61)**
Coli	7	6.0	1.17 (0.56-2.45)	44	48.3	0.91 (0.68-1.22)	10	5.3	1.89 (1.02-3.51)*

	Live animals			Food production			Food sale		
SALMONELLOSIS	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Age									
16-19 years	29	9.5	3.05 (2.12-4.39)***	216	213.2	1.01 (0.8-1.16)	28	18.4	1.52 (1.05-2.20)*
20-29 years	43	26.5	1.62 (1.20-2.19)**	333	296.7	1.12 (1.01-1.25)*	38	25.2	1.51 (1.10-20.7)*
30-39 years	15	13.8	1.08 (0.65-1.80)	65	73.6	0.88 (0.69-1.13)	17	10.6	1.60 (1.00-2.58)
40-49 years	26	17.0	1.52 (1.04-2.24)*	70	75.8	0.92 (0.73-1.17)	18	12.4	1.45 (0.92-2.30)
≥50 years	34	22.0	1.55 (1.11-2.17)*	78	85.0	0.92 (0.74-1.15)	8	13.4	0.60 (0.30-1.19)
P-trend	<0.001			0.909			0.070		

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio; *p<0.05; **p<0.01; ***p<0.001; $ *Salmonella* isolated from urine or wounds.

Serology

Data from 732 serosurvey participants remained for analysis (294 males; 438 females). Mean age at sampling was 37 years (SD: 12). Duration of registered employment (since ≥1999) at sampling increased with age, from a median of 1.1 year (IQR: 0.5–2.2) for those aged 16–19 years to 7.8 years (IQR: 7.4–8.2) for people ≥50 years.

Mean seroincidence adjusted for gender and years of employment was 0.74 infections/person-year (95% CI 0.73 to 0.75, n=721) for *Salmonella* and 1.81 infections/person-year (95% CI 1.80 to 1.83, n=725) for *Campylobacter* (Table 3). Seroincidence in females versus males was slightly lower for *Salmonella* (exp(b): 0.915 (95% CI 0.770 to 1.086, p=0.308)) and slightly higher for *Campylobacter* (exp(b): 1.086 (95% CI 0.996 to 1.185, p=0.062)), though both non-significant. Age at sampling and years of employment showed a significant positive association with *Salmonella* and *Campylobacter* seroincidences; however, due to high collinearity between these two variables, we only considered years of employment (exp(b): 1.074 (95% CI 1.036 to 1.113) per year of employment for *Salmonella*; exp(b): 1.038 (95% CI 1.019 to 1.057 for *Campylobacter*)). Serology data were available for 19 occupational sections (Table 4). Mean seroincidence of *Salmonella* per section ranged from 0.36 to 0.99 infections/person-year. No significant differences were observed in the seroincidence for *Salmonella* between sections (i.e., comparisons of each section *vs.* all others). Among sections with ≥10 participants, seroincidences were highest in the sections 'transportation and storage', 'financial and insurance activities' and 'real estate activities'. For *Campylobacter*, the mean seroincidence ranged from 1.28 to 2.30 infections/person-year, with a significantly higher seroincidence rate in the 'other service activities' section (2.30; 95% CI 2.18 to 2.43) compared with other sections. Table 4 shows the seroincidence for *Salmonella* and *Campylobacter* in people exposed to the high-risk occupations: for both pathogens, seroincidence was slightly increased in these high-risk occupations (exp(b): 1.08, 95% CI 0.75 to 1.56, p=0.677, for *Salmonella*; exp(b): 1.03, 95% CI 0.86 to 1.24, p=0.732, for *Campylobacter*), although non-significant.

Table 3. Mean and 95% confidence interval of seroincidence rates (i.e. number of infections/person-year) of Salmonella and Campylobacter, by age at sampling

	Salmonella			*Campylobacter*		
Age at sampling	**N**	**Mean§**	**95%CI**	**N**	**Mean§**	**95%CI**
16-19 years	49	0.52	0.51-0.53	49	1.53	1.27-1.83
20-29 years	176	0.69	0.68-0.71	178	1.76	1.73-1.78
30-39 years	193	0.78	0.77-0.80	194	1.87	1.86-1.89
40-49 years	160	0.78	0.77-0.80	159	1.86	1.84-1.88
≥50 years	143	0.79	0.77-0.81	145	1.87	1.84-1.89
Total	721	0.74	0.73-0.75	725	1.81	1.80-1.83

§ Adjusted for gender and years of employment at time of sampling.

Table 4. Mean and confidence interval of seroincidence rates (number of infections/person-year) of *Salmonella* and *Campylobacter,* by section

	Salmonella		***Campylobacter***	
Section	N	Mean (95%CI)[§]	N	Mean (95%CI)[§]
Agriculture, forestry and fishing	6	0.77 (0.57-1.04)	8	1.62 (1.46-1.79)
Manufacturing	76	0.70 (0.67-0.73)	76	1.78 (1.74-1.82)
Electricity, gas, steam and air conditioning supply	4	0.44 (0.31-0.63)	4	1.55 (1.21-1.98)
Water supply; sewerage, waste management and remediation activities	4	0.36 (0.25-0.53)	4	1.28 (1.06-1.56)
Construction	36	0.61 (0.57-0.68)	35	1.72 (1.66-1.78)
Wholesale and retail trade	133	0.63 (0.61-0.65)	134	1.81 (1.77-1.84)
Transportation and storage	26	0.99 (0.91-1.09)	26	1.76 (1.69-1.84)
Accommodation and food service activities	29	0.48 (0.44-0.51)	29	1.78 (1.71-1.85)
Information and communication	19	0.76 (0.69-0.84)	19	1.57 (1.51-1.64)
Financial and insurance activities	35	0.97 (0.91-1.04)	36	2.07 (1.98-2.15)
Real estate activities	10	0.95 (0.80-1.13)	10	1.64 (1.51-1.78)
Professional, scientific and technical activities	61	0.68 (0.64-0.71)	60	1.81 (1.77-1.86)
Administrative and support service activities	109	0.67 (0.65-0.68)	108	1.72 (1.69-1.75)
Public administration and defence; compulsory social security	47	0.77 (0.73-0.81)	48	1.89 (1.84-1.94)
Education	72	0.78 (0.75-0.82)	73	1.82 (1.78-1.87)
Human health and social work activities	167	0.70 (0.68-0.72)	169	1.78 (1.76-1.80)
Arts, entertainment and recreation	20	0.45 (0.43-0.49)	20	1.54 (1.48-1.59)
Other service activities	21	0.91 (0.84-0.99)	21	2.30 (2.18-2.43)**
Activities of households as employers	2	0.95 (0.72-1.26)	2	2.26 (1.96-2.62)
High risk occupations[†]	43	0.65 (0.60-0.70)	44	1.69 (1.64-1.75)
All occupations	721	0.74 (0.73-0.75)	725	1.81 (1.80-1.83)

§ Adjusted for gender and years of employment at time of sampling; *p<0.05; **p<0.01; ***p<0.001
† Risk groups 'live animals', 'food production' and 'food sale' combined.

Discussion

We assessed the distribution of reported salmonellosis and campylobacteriosis cases, as well as the magnitude of exposure to these pathogens, among different occupational groups in the Dutch-employed population. We identified significantly increased SIRs for both salmonellosis and campylobacteriosis in several occupations. These observations can be explained by a combination of multiple coexisting factors entailing exposure levels to the pathogens, susceptibility to infection and medical awareness/knowledge associated with the occupations in question.

The risk of reported salmonellosis and campylobacteriosis was almost twofold higher in people in the 'live animals' group, presumably caused by increased exposure to both pathogens. Similar, although stronger, associations were found in a registry study in the USA where the relative risk of salmonellosis and campylobacteriosis among people working in occupations including farming was respectively 10-fold and threefold higher compared with other occupations [9], whereas in another study, 17% of the campylobacteriosis cases reported occupational exposure to animals [7]. Among people with occupational exposure to animal-derived food products, we observed a significantly increased risk of infection in the 'food sale' group and in the 'accommodation' sector. However, we did not observe it in the overall risk group involved in food production/preparation. Acquired immunity against *Salmonella* and *Campylobacter* might be an explanation for the latter observation. Furthermore, SIRs (for salmonellosis and/or campylobacteriosis) were significantly higher in five industrial sectors, mainly those associated with the use of chemicals. Long-term exposure to chemical substances is associated with altered composition of gut microbiota, resulting in dysregulation of the gut mucosal immune function, which in turn might lead to adverse health effects and possibly increased susceptibility to enteric infections [22]. Generally, frailty and low SES are risk factors for increased morbidity and mortality of disease [23]. This could explain the increased SIRs among people working in the 'other manufacturing' sector, whereas the decreased SIRs among people working in 'white collar' sectors could be explained by a higher educational level, SES and general health. Moreover, marginally higher SIRs were observed among people working in healthcare and social work. An underlying factor might be increased infection pressure in such facilities, as nosocomial outbreaks and outbreaks in long-term care facilities are documented for *Salmonella* [24, 25]. On the other hand, the increased SIRs might also be partially attributable to increased healthcare-seeking behaviour caused by medical awareness/knowledge among people working in these sectors. Shift work, being common in healthcare and industrial sectors, has been proposed to increase the risk of infection as an indirect result of sleep rhythm and health behaviour on the immune function [26, 27]. In addition to the infection risk from occupational exposure, the observed distribution of salmonellosis and campylobacteriosis cases in our study is, to some extent, influenced by surveillance/detection bias, potential healthy worker effect, as well as the confounding effect of lifestyle, which we could not fully control for [28]. Indeed, here we could not account for other potential risk factors related to, for example, eating habits, pet ownership, travel and ethnicity, which might have played a role as well. Moreover, the study period covered 17 years in which diagnostics of gastrointestinal infections and hygiene standards in, for example, abattoirs, might have changed. However, we consistently used the same type of data (i.e., culture-confirmed *Salmonella* and *Campylobacter* infections, as culture is still performed for antimicrobial resistance determination after positive PCR

screening) and temporal trends in reported salmonellosis/campylobacteriosis were not assessed, that is, data were analysed retrospectively by including (cumulative) employment time of each individual in each sector, and not chronological time per se. Thus, while the strength of some associations might differ between periods, our study was meant to provide overall estimates for the average effects of occupation during the whole study period.

Salmonellosis and campylobacteriosis IRs are based on laboratory-confirmed cases reported to public health surveillance. These cases constitute only a small fraction of all cases occurring in the community and are usually patients with severe or prolonged symptoms. The extent of under-reporting is influenced by healthcare-seeking behaviours and patient-related sensitivity of the healthcare and surveillance systems (e.g., patients with travel history or those with underlying chronic diseases are more likely to undergo increased medical scrutiny on presentation of symptoms). Serology allows us to assess infection risks independently of these factors, as it also includes asymptomatic infections, hence it sheds light on the epidemiology of *Salmonella* and *Campylobacter* from a different perspective [18]. We found the seroincidence to be only slightly associated with occupational groups with a higher incidence of reported salmonellosis and campylobacteriosis cases. Previous studies comparing seroincidence rates among countries have found no significant correlations with incidence of reported cases, with seroincidence rates being up to 130-fold higher than reported incidence [17]. Besides surveillance artefacts, possible explanations could be the intrinsic limitations of seroincidence data, such as differential antibody decay over time in different groups of the population. It is difficult to predict how this may have affected the seroincidence estimates of our high-risk groups. If the antibody response is stronger, seroincidence would be overestimated. However, if frequent infections induce a weaker immune response, especially lower IgM production, seroincidence would be underestimated, as pointed out before [17].

A limitation of the NACE classification is that a person's NACE code is based on the economic activity of the company/organisation employing the linked person, rather than the actual job tasks. The proportion of people employed via an employment agency differs among sectors, with most people in the 'employment activities' sector working in industry (24% males; 14% females) [29]. This might affect the observed risk of infection among occupational groups. Furthermore, serological data were limited by the sample size of serosurvey participants in some sectors (probably due to participation bias), which hampered comparisons between groups. In conclusion, we found significantly increased occupational risks for salmonellosis and campylobacteriosis among people with occupational exposure to animals or animal-derived products, healthcare and social workers, as well as people working in specific industrial sectors. Seroincidence in these high-risk groups was only slightly increased, suggesting possible

differential antibody response and decay over time (on increased exposure to *Salmonella* and *Campylobacter*) in different groups. Campylobacteriosis and salmonellosis should be considered when workers in occupations at increased risk for infection have symptoms compatible with these diseases. Although the exact transmission routes in these occupational groups are yet to be fully understood, targeting education and prevention strategies may help reduce disease and provide a better understanding of these occupationally acquired infections.

Contributors

All authors conceived and designed the study. JWD performed the data analysis together with LMG. JWD drafted the paper. All authors have substantially contributed to critical interpretation of the results and drafting/revising of the paper.

Funding

This study was supported by the Dutch Ministry for Public Health, Welfare and Sport via the 'ZonMw' (grant number 522004001).

Competing interests

None declared.

Patient consent for publication

Not required.

Ethics approval

The serosurvey (PIENTER-2) was approved by the Medical Ethics Testing Committee of the Foundation of Therapeutic Evaluation of Medicines (METC-STEG) in Almere (ISRCTN 20164309). This study was performed on deidentified data and no person identifying information was generated.

References

1. European Food Safety Authority and European Centre for Disease Prevention and Control (EFSA and ECDC). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. EFSa Journal 2018;16:e05500.
2. National Institute for Public Health and the Environment (RIVM). Staatvan Zoönosen 2017 [State of Zoonotic Diseases 2017]. Bilthoven: RIVM, 2018.
3. Mangen MJ , Friesema IHM , Pijnacker R , et al . Disease burden of food-related pathogens in the Netherlands, 2017. 2018. Bilthoven: RIVM, 2018.
4. Mangen M-JJ , Bouwknegt M , Friesema IHM , et al . Cost-Of-Illness and disease burden of food-related pathogens in the Netherlands, 2011. Int J Food Microbiol 2015;196:84–93.
5. Mughini-Gras L , Enserink R , Friesema I , et al . Risk factors for human salmonellosis originating from pigs, cattle, broiler chickens and egg laying hens: a combined case-control and source attribution analysis. PLoS One 2014;9:e87933.
6. Mughini Gras L , Smid JH , Wagenaar JA , et al . Risk factors for campylobacteriosis of chicken, ruminant, and environmental origin: a combined case-control and source attribution analysis. PLoS One 2012;7:e42599.
7. Su C-P , Stover DT , Buss BF , et al . Occupational Animal Exposure Among Persons with Campylobacteriosis and Cryptosporidiosis - Nebraska, 2005-2015. MMWR Morb Mortal Wkly Rep 2017;66:955–8.
8. Ellström P , Hansson I , Söderström C , et al . A prospective follow-up study on transmission of Campylobacter from poultry to abattoir workers. Foodborne Pathog Dis 2014;11:684–8.
9. Su C-P , de Perio MA , Fagan K , et al . Occupational Distribution of Campylobacteriosis and Salmonellosis Cases - Maryland, Ohio, and Virginia, 2014. MMWR Morb Mortal Wkly Rep 2017;66:850–3.
10. Eurostat. NACE Rev. 2—Statistical classification of economic activities, 2008. Available: http://ec.europa.eu/eurostat/web/nace-rev2
11. van Pelt W , de Wit MAS , Wannet WJB , et al . Laboratory surveillance of bacterial gastroenteric pathogens in the Netherlands, 1991-2001. Epidemiol Infect 2003;130:431–41.
12. van der Klis FRM , Mollema L , Berbers GAM , et al . Second National serum bank for population-based seroprevalence studies in the Netherlands. Neth J Med 2009;67:301–8.
13. Mollema L , De Melker HE , Hahné SJM , et al . PIENTER 2-project: second research project on the protection against infectious diseases offered by the National immunization programme in the Netherlands 2010.
14. Strid MA , Dalby T , Mølbak K , et al . Kinetics of the human antibody response against Salmonella enterica serovars enteritidis and typhimurium determined by lipopolysaccharide enzyme-linked immunosorbent assay. Clin Vaccine Immunol 2007;14:741–7.
15. Ang CW , Krogfelt K , Herbrink P , et al . Validation of an ELISA for the diagnosis of recent Campylobacter infections in Guillain-Barré and reactive arthritis patients. Clin Microbiol Infect 2007;13:915–22.
16. Simonsen J , Teunis P , van Pelt W , et al . Usefulness of seroconversion rates for comparing infection pressures between countries. Epidemiol Infect 2011;139:636–43.
17. Falkenhorst G , Simonsen J , Ceper TH , et al . Serological cross-sectional studies on Salmonella incidence in eight European countries: no correlation with incidence of reported cases. BMC Public Health 2012;12:523.

18. Monge S , Teunis P , Friesema I , et al . Immune response-eliciting exposure to Campylobacter vastly exceeds the incidence of clinically overt campylobacteriosis but is associated with similar risk factors: a nationwide serosurvey in the Netherlands. J Infect 2018;77:171–7.
19. Ang CW , Teunis PFM , Herbrink P , et al . Seroepidemiological studies indicate frequent and repeated exposure to Campylobacter spp. during childhood. Epidemiol Infect 2011;139:1361–8.
20. Teunis PFM , van Eijkeren JCH , Ang CW , et al . Biomarker dynamics: estimating infection rates from serological data. Stat Med 2012;31:2240–8.
21. Mughini-Gras L , Schaapveld M , Kramers J , et al . Increased colon cancer risk after severe Salmonella infection. PLoS One 2018;13:e0189721.
22. Jin Y , Wu S , Zeng Z , et al . Effects of environmental pollutants on gut microbiota. Environ Pollut 2017;222:1–9.
23. Mughini-Gras L , Pijnacker R , Heusinkveld M , et al . Societal burden and correlates of acute gastroenteritis in families with preschool children. Sci Rep 2016;6:22144.
24. Lee MB , Greig JD . A review of nosocomial Salmonella outbreaks: infection control interventions found effective. Public Health 2013;127:199–206.
25. Greig JD , Lee MB . Enteric outbreaks in long-term care facilities and recommendations for prevention: a review. Epidemiol Infect 2009;137:145–55.
26. Mohren DCL , Jansen NWH , Kant IJ , et al . Prevalence of common infections among employees in different work schedules. J Occup Environ Med 2002;44:1003–11.
27. Almeida CMOde , Malheiro A . Sleep, immunity and shift workers: a review. Sleep Sci 2016;9:164–8.
28. Spierenburg EAJ , Smit LAM , Heederik D , et al . Healthy worker survivor analysis in an occupational cohort study of Dutch agricultural workers. Int Arch Occup Environ Health 2015;88:1165–73.
29. CBS. De uitzendbranche in Nederland sinds 2005. retrieved at 4 February 2019 from, 2017. Available: https://www.cbs.nl/nl-nl/achtergrond/2017/43/de-uitzendbranche-in-nederland-sinds-2005

Supplementary material

Table S1. Classification of risk groups.

NACE code*	Description
Risk group 1 – Contact with live animals and animal manure	
01410	Raising of dairy cattle
01420	Raising of other cattle and buffaloes
01430	Raising of horses and other equines
01450	Raising of sheep and goats
01460	Raising of swine and pigs
01470	Raising of poultry
01490	Raising of other animals
01500	Mixed farming
01620	Support activities for animal production
10110	Processing and preserving of meat
10120	Processing and preserving of poultry meat
20150	Manufacture of fertilizers and nitrogen compounds
37000	Sewerage
38210	Treatment and disposal of non-hazardous waste**
46231	Wholesale of live farm animals
46232	Wholesale of live pets
75000	Veterinary activities
91041	(Petting) zoos
Risk group 2 – Production and handling of food products	
10130	Production of meat and poultry meat products
10510	Operation of dairies and cheese making
10520	Manufacture of ice cream
10710	Manufacture of bread; manufacture of fresh pastry goods and cakes
10730	Manufacture of macaroni, noodles, couscous and similar farinaceous products
10840	Manufacture of condiments and seasonings
10850	Manufacture of prepared meals and dishes
10890	Manufacture of other food products n.e.c.
10910	Manufacture of prepared feeds for farm animals
10920	Manufacture of prepared pet foods
55101	Hotel-restaurants
56101	Restaurants
56102	Fast food restaurants and mobile food service activities
56210	Event catering activities
56290	Other food service activities (e.g. canteens)
Risk group 3 – Sale of animal products	
46320	Wholesale of meat and meat products
46331	Wholesale of dairy products

NACE code*	Description
46332	Wholesale of eggs
47221	Retail sale of meat and charcuterie in specialized stores
47222	Retail sale of game meat and poultry meat in specialized stores
47230	Retail sale of fish, crustacean and mollusks in specialized stores
47241	Retail sale of bread, cakes, flour confectionary in specialized stores
47291	Retail sale of cheese in specialized stores

*NACE code (first 4 digits), Netherlands' specific digit (fifth). ** includes processing/disposal/ destruction of slurry and animal carcasses.

Table S2. IRs of salmonellosis and campylobacteriosis in the Dutch employed population.

	Salmonellosis		**Campylobacteriosis**	
	Obs	IR (95%CI)[§]	Obs	IR (95%CI)[§]
Total	8,220	6.51 (6.36-6.65)	14,352	15.54 (15.29-15.79)
Gender				
Male	4,174	6.06 (5.88-6.25)	7,584	15.31 (14.97-15.65)
Female	4,046	7.04 (6.82-7.25)	6,768	15.81 (15.43-16.18)
***Salmonella* serovar or *Campylobacter* species**				
S. Typhimurium	2,113	1.67 (1.60-1.74)	-	-
S. Enteritidis	3,263	2.58 (2.49-2.67)	-	-
Other *Salmonella* serovars	2,844	2.25 (2.17-2.33)	-	-
C. jejuni	-	-	13,377	14.48 (14.24-14.73)
C. coli	-	-	975	1.06 (0.99-1.12)
Type of infection				
Enteric	7,672	6.07 (5.94-6.21)	-	-
Septicemic	226	0.18 (0.16-0.20)	-	-
Other[ǂ]	322	0.25 (0.23-0.28)	-	-
Age group				
16-19 years	1,024	12.72 (11.94-13.50)	1332	22.89 (21.67-24.12)
20-29 years	2,930	10.85 (10.45-11.24)	4285	22.46 (21.79-23.13)
30-39 years	1,428	4.69 (4.45-4.93)	2419	11.52 (11.06-11.98)
40-49 years	1,352	4.33 (4.10-4.56)	2763	11.98 (11.53-12.43)
≥50 years	1,486	5.02 (4.77-5.28)	3553	15.19 (14.69-15.69)
Total number of individuals exposed	12,566,831		11,615,429	
Years of data	1999-2016		2004-2016	

Obs: observed numbers. IR: incidence rate. § per 100 000 person-years. ǂ *Salmonella* isolated from urinary or wounds.

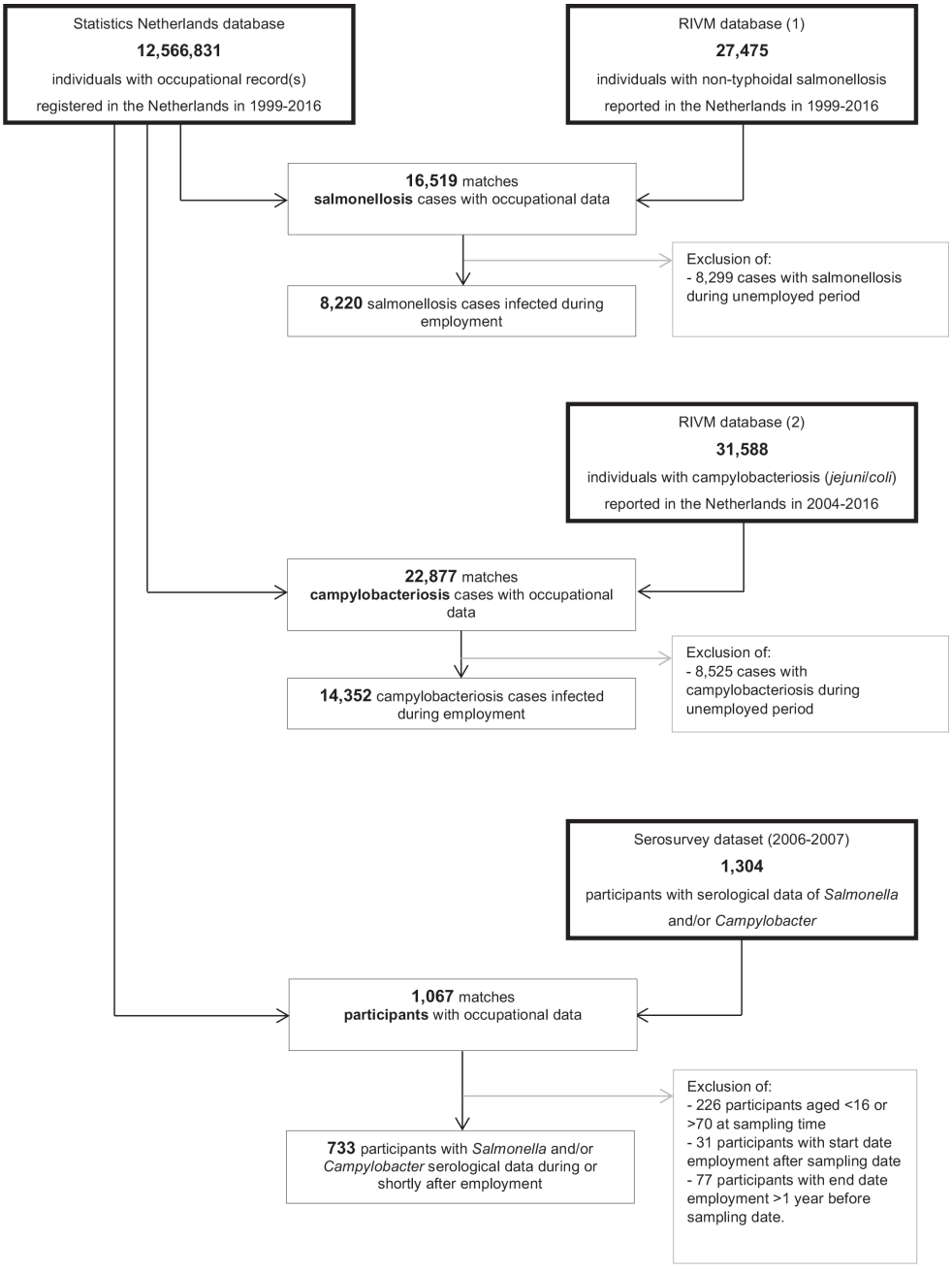


Figure S1. Schematic representation of the data management process.

Table S3. SIRs of salmonellosis and campylobacteriosis by division

	Salmonellosis			Campylobacteriosis		
Div	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Crop and animal production, hunting	105	119.2	0.88 (0.73-1.07)	141	181.2	0.78 (0.66-0.92)**
Forestry and logging	1	1.4	0.69 (0.09-4.93)	1	2.5	0.39 (0.06-2.78)
Fishing and aquaculture	2	1.5	1.36 (0.34-5.45)	5	2.4	2.06 (0.86-4.95)
Extraction of crude petroleum and natural gas	6	3.4	1.75 (0.79-3.89)	16	6.3	2.54 (1.56-4.15)***
Other mining and quarrying	5	1.9	2.61 (1.08-6.26)	5	3.5	1.42 (0.59-3.41)
Mining support service activities	5	2.8	1.78 (0.74-4.27)	9	5.8	1.55 (0.80-2.97)
Manufacture of food products	126	132.8	0.95 (0.80-1.13)	230	225.2	1.02 (0.90-1.16)
Manufacture of beverages	5	7.8	0.64 (0.27-1.54)	8	13.1	0.61 (0.31-1.22)
Manufacture of tobacco products	4	3.6	1.11 (0.42-2.97)	7	5.6	1.24 (0.59-2.60)
Manufacture of textiles	16	13.4	1.19 (0.73-1.94)	20	21.0	0.95 (0.61-1.47)
Manufacture of wearing apparel	7	4.0	1.75 (0.84-3.68)	4	4.6	0.87 (0.33-2.31)
Manufacture of leather (products)	1	2.1	0.47 (0.07-3.31)	3	3.2	0.95 (0.31-2.95)
Manufacture of wood/cork etc	17	16.5	1.03 (0.64-1.66)	22	27.1	0.82 (0.53-1.23)
Manufacture of paper (products)	15	19.5	0.77 (0.46-1.28)	51	34.1	1.50 (1.14-1.97)**
Printing and reproduction of recorded media	28	35.0	0.80 (0.55-1.16)	44	53.1	0.83 (0.62-1.11)
Manufacture of coke and refined petroleum products	7	5.2	1.34 (0.64-2.82)	13	10.6	1.22 (0.71-2.10)
Manufacture of chemicals	59	42.6	1.38 (1.07-1.79)*	112	81.0	1.38 (1.15-1.66)**
Manufacture of basic pharmaceutical products and preparations	8	8.0	1.00 (0.50-1.99)	15	22.6	0.66 (0.40-1.10)
Manufacture of rubber and plastic	38	31.1	1.22 (0.89-1.68)	65	56.0	1.16 (0.91-1.48)
Manufacture of other non-metallic mineral products	23	25.1	0.92 (0.61-1.38)	45	44.6	1.01 (0.75-1.35)
Manufacture of basic metals	13	20.3	0.64 (0.37-1.10)	38	38.5	0.99 (0.72-1.36)
Manufacture of fabricated metal products	99	89.2	1.11 (0.92-1.35)	153	158.6	0.96 (0.82-1.13)
Manufacture of computer, electronic and optical products	33	28.6	1.15 (0.82-1.62)	45	48.3	0.93 (0.70-1.25)
Manufacture of electrical equipment	9	13.2	0.68 (0.36-1.31)	43	33.2	1.29 (0.96-1.75)
Manufacture of machinery and equipment	59	45.5	1.30 (1.01-1.67)*	114	124.7	0.91 (0.76-1.10)
Manufacture of motor vehicles	22	16.8	1.31 (0.86-1.99)	28	34.2	0.82 (0.56-1.18)

	Salmonellosis			Campylobacteriosis		
Div	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Manufacture of other transport equipment	12	10.9	1.11 (0.63-1.95)	37	28.2	1.31 (0.95-1.81)
Manufacture of furniture	9	15.8	0.57 (0.30-1.09)	33	33.2	0.99 (0.71-1.40)
Other manufacturing	183	115.8	1.58 (1.37-1.83)***	323	230.7	1.40 (1.26-1.56)***
Repair/installation of machinery	31	22.3	1.39 (0.98-1.98)	61	62.2	0.98 (0.76-1.26)
Electricity, gas, steam and air conditioning supply	17	23.2	0.73 (0.46-1.18)	67	43.9	1.53 (1.20-1.94)**
Water collection, treatment and supply	4	5.3	0.75 (0.28-2.01)	13	10.0	1.30 (0.76-2.24)
Sewerage	3	4.0	0.75 (0.24-2.33)	9	7.7	1.18 (0.61-2.26)
Waste collection, treatment and disposal activities	31	23.3	1.33 (0.94-1.89)	54	44.8	1.21 (0.92-1.57)
Remediation activities and other waste management services	4	1.19	3.35 (1.26-8.94)	6	2.9	2.06 (0.92-4.58)
Construction of buildings	64	64.1	1.00 (0.78-1.27)	141	169.7	0.83 (0.70-0.98)*
Civil engineering	39	53.3	0.73 (0.53-1.00)	97	100.1	0.97 (0.79-1.18)
Specialized construction activities	24	218.1	0.98 (0.86-1.12)	370	373.9	0.99 (0.89-1.10)
Wholesale, retail trade and repair of motor vehicles	117	136.6	0.86 (0.71-1.03)	224	231.0	0.97 (0.85-1.11)
Wholesale trade	403	431.5	0.93 (0.85-1.03)	716	845.7	0.85 (0.79-0.91)***
Retail trade	1000	1041.5	0.96 (0.90-1.02)	1621	1621.3	1.00 (0.95-1.05)
Land transport	157	185.1	0.85 (0.73-0.99)*	365	345.0	1.06 (0.95-1.17)
Water transport	23	15.5	1.48 (0.99-2.23)	32	27.8	1.15 (0.81-1.63)
Air transport	83	69.4	1.20 (0.96-1.48)	43	49.6	0.87 (0.64-1.17)
Warehousing and support activities for transportation	87	74.7	1.17 (0.94-1.44)	153	144.0	1.06 (0.91-1.25)
Postal and courier activities	84	85.5	0.98 (0.79-1.22)	140	144.2	0.97 (0.82-1.15)
Accommodation	102	95.5	1.06 (0.87-1.28)	213	155.0	1.37 (1.20-1.57)***
Food and beverage service activities	396	393.2	1.01 (0.91-1.11)	641	632.0	1.01 (0.94-1.10)
Publishing activities	31	39.7	0.78 (0.55-1.11)	69	60.0	1.15 (0.91-1.46)
Motion picture, video, television program production etc.	19	17.7	1.07 (0.68-1.68)	20	28.2	0.71 (0.46-1.10)
Programming and broadcasting activities	9	7.5	1.20 (0.62-2.31)	6	13.8	0.43 (0.20-0.97)
Telecommunications	36	42.5	0.85 (0.61-1.17)	59	64.5	0.92 (0.71-1.18)
Computer programming, consultancy and related activities	80	91.9	0.87 (0.70-1.08)	149	240.5	0.62 (0.53-0.73)***
Information service activities	11	7.9	1.39 (0.77-2.50)	17	19.6	0.87 (0.54-1.40)
Financial service activities, except insurance and pension funding	125	145.8	0.86 (0.72-1.02)	200	256.3	0.78 (0.68-0.90)***

	Salmonellosis			Campylobacteriosis		
Div	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Insurance, reinsurance and pension funding	56	63.4	0.88 (0.68-1.15)	126	121.7	1.04 (0.87-1.23)
Activities auxiliary to financial services and insurance activities	59	58.5	1.01 (0.78-1.30)	72	100.1	0.72 (0.57-0.91)**
Real estate activities	83	68.8	1.21 (0.97-1.50)	114	128.6	0.89 (0.74-1.07)
Legal and accounting activities	133	139.7	0.95 (0.80-1.13)	220	241.0	0.91 (0.80-1.04)
Activities of head offices; management consultancy activities	134	141.5	0.95 (0.80-1.12)	191	279.0	0.68 (0.59-0.79)***
Architectural and engineering activities; technical testing and analysis	79	102.8	0.77 (0.62-0.96)*	162	191.3	0.85 (0.73-0.99)*
Scientific research and development	34	34.8	0.98 (0.70-1.37)	49	63.0	0.78 (0.59-1.03)
Advertising and market research	61	62.7	0.97 (0.76-1.25)	76	93.4	0.81 (0.65-1.02)
Other professional, scientific and technical activities	12	20.0	0.60 (0.34-1.06)	36	44.1	0.82 (0.59-1.13)
Veterinary activities	15	7.4	2.03 (1.22-3.37)**	26	13.3	1.96 (1.33-2.87)**
Rental and leasing activities	37	30.3	1.22 (0.88-1.68)	48	53.4	0.90 (0.68-1.19)
Employment activities	640	672.8	0.95 (0.88-1.03)	1,170	1106.3	1.05 (1.00-1.12)
Travel agency, tour operator reservation service etc.	28	28.5	0.98 (0.68-1.42)	46	44.1	1.04 (0.78-1.39)
Security and investigation activities	34	33.1	1.03 (0.73-1.44)	62	64.8	0.96 (0.75-1.23)
Services to buildings and landscape activities	184	176.2	1.04 (0.90-1.21)	247	287.5	0.86 (0.76-0.97)*
Office administrative, office support etc.	30	32.4	0.93 (0.65-1.33)	60	67.7	0.89 (0.69-1.14)
Public administration and defence; compulsory social security	513	506.6	1.01 (0.93-1.10)	987	944.0	1.05 (0.98-1.11)
Education	428	502.3	0.85 (0.78-0.94)**	857	934.6	0.92 (0.86-0.98)*
Human health activities	550	492	1.12 (1.03-1.21)*	1,123	918.6	1.22 (1.15-1.30)***
Residential care activities	559	478.3	1.17 (1.08-1.27)***	981	835.4	1.17 (1.10-1.25)***
Social work activities without accommodation	362	331.6	1.09 (0.98-1.21)	686	620.5	1.11 (1.03-1.19)**
Creative, arts and entertainment activities	20	24.5	0.82 (0.53-1.26)	47	47.5	0.99 (0.74-1.32)
Libraries, archives, museums and other cultural activities	26	22.5	1.15 (0.79-1.70)	50	42.1	1.19 (0.90-1.57)
Gambling and betting activities	8	9.0	0.89 (0.44-1.78)	19	15.8	1.20 (0.77-1.88)
Sports activities and amusement and recreation activities	95	84.7	1.12 (0.92-1.37)	157	152.9	1.03 (0.88-1.20)

	Salmonellosis			Campylobacteriosis		
Div	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Activities of membership organizations	74	81.1	0.91 (0.73-1.15)	113	140.0	0.81 (0.67-0.97)*
Repair of computers and personal/household goods	3	5.6	0.54 (0.17-1.67)	9	10.7	0.84 (0.44-1.62)
Other personal service activities	75	79.7	0.94 (0.75-1.18)	144	131.1	1.10 (0.93-1.29)
Activities of households as employers of domestic personnel	26	23.2	1.12 (0.76-1.65)	82	57.8	1.42 (1.14-1.76)**
Undifferentiated goods- and services-producing activities of private households for own use	0	0.003	-	0	0.01	-
Activities of extraterritorial organizations and bodies	0	1.6	-	4	3.6	1.12 (0.42-2.99)

Obs: observed numbers; Exp: expected numbers; *p<0.05; **p<0.01; ***p<0.001.

Table S4. SIRs of salmonellosis and campylobacteriosis in the 'veterinary activities' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	3	0.9	3.26 (1.05-10.09)*	4	1.8	2.22 (0.84-5.93)
Female	12	6.5	1.86 (1.05-3.27)*	22	11.5	1.91 (1.26-2.91)**
Serovar/species						
S. Typhimurium	9	1.82	4.94 (2.57-9.50)***	-	-	-
S. Enteritidis	2	2.8	0.71 (0.18-2.85)	-	-	-
Salmonella other	4	2.8	1.45 (0.54-3.86)	-	-	-
C. jejuni	-	-	-	24	12.4	1.93 (1.29-2.88)**
C. coli	-	-	-	2	0.9	2.30 (0.58-9.20)
Age group						
16-19 years	1	0.2	4.34 (0.62-30.82)	0	0.3	-
20-29 years	10	4.1	2.44 (1.31-4.54)**	14	6.7	2.10 (1.24-3.54)**
30-39 years	0	1.6	-	5	2.9	1.70 (0.71-4.08)
40-49 years	2	0.8	2.36 (0.59-9.43)	3	1.8	1.62 (0.52-5.03)
≥50 years	2	0.6	3.12 (0.78-12.46)	4	1.6	2.57 (0.96-6.83)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S5. SIRs of salmonellosis and campylobacteriosis in the 'manufacture of chemicals' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	48	33.9	1.42 (1.07-1.88)*	96	66.5	1.44 (1.18-1.76)***
Female	11	8.7	1.26 (0.70-2.27)	16	14.5	1.10 (0.68-1.80)
Serovar/species						
S. Typhimurium	9	9.6	0.94 (0.49-1.80)	-	-	-
S. Enteritidis	26	17.6	1.48 (1.01-2.17)*	-	-	-
Salmonella other	24	15.4	1.55 (1.04-2.32)*	-	-	-
C. jejuni	-	-	-	101	75.5	1.34 (1.10-1.63)**
C. coli	-	-	-	11	5.6	1.97 (1.09-3.56)*
Age group						
16-19 years	0	0.7	-	0	0.9	-
20-29 years	9	8.7	1.03 (0.54-1.98)	15	11.5	1.30 (0.78-2.16)
30-39 years	13	10.2	1.28 (0.74-2.21)	24	14.9	1.61 (1.08-2.40)*
40-49 years	18	11.3	1.59 (1.00-2.52)*	32	23.8	1.35 (0.95-1.90)
≥50 years	19	11.7	1.63 (1.04-2.55)*	41	29.9	1.37 (1.01-1.86)*

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S6. SIRs of salmonellosis and campylobacteriosis in the 'manufacture of paper (products)' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	15	16.0	0.94 (0.57-1.56)	49	28.9	1.70 (1.28-2.24)***
Female	0	3.5	-	2	5.2	0.38 (0.10-1.53)
Serovar/species						
S. Typhimurium	7	4.3	1.61 (0.77-3.38)	-	-	-
S. Enteritidis	3	8.3	0.36 (0.12-1.12)	-	-	-
Salmonella other	5	6.9	0.73 (0.30-1.75)	-	-	-
C. jejuni	-	-	-	50	31.7	1.58 (1.19-2.08)**
C. coli	-	-	-	1	2.4	0.42 (0.06-2.99)
Age group						
16-19 years	0	0.5	-	0	0.5	-
20-29 years	2	3.9	0.51 (0.13-2.03)	4	4.4	0.90 (0.34-2.40)
30-39 years	2	4.8	0.42 (0.10-1.66)	9	6.1	1.46 (0.76-2.81)
40-49 years	4	5.3	0.75 (0.28-2.01)	19	10.9	1.75 (1.11-2.74)*
≥50 years	7	4.9	1.42 (0.68-2.99)	19	12.1	1.57 (1.00-2.47)*

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S7. SIRs of salmonellosis and campylobacteriosis in the 'extraction of crude petroleum and natural gas' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	4	2.9	1.38 (0.51-3.67)	14	5.4	2.59 (1.54-4.38)***
Female	2	0.5	3.78 (0.95-15.12)	2	0.9	2.22 (0.56-8.89)
Serovar/species						
S. Typhimurium	0	0.7	-	-	-	-
S. Enteritidis	4	1.4	2.76 (1.03-7.37)*	-	-	-
Salmonella other	2	1.2	1.60 (0.40-6.41)	-	-	-
C. jejuni	-	-	-	16	5.9	2.72 (1.67-4.45)***
C. coli	-	-	-	0	0.4	-
Age group						
16-19 years	0	0.01	-	0	0.01	-
20-29 years	0	0.5	-	2	0.8	2.66 (0.66-10.62)
30-39 years	1	0.7	1.42 (0.20-10.08)	3	1.1	2.65 (0.86-8.23)
40-49 years	3	1.1	2.72 (0.88-8.43)	4	1.7	2.42 (0.91-6.45)
≥50 years	2	1.1	1.82 (0.45-7.27)	7	2.8	2.54 (1.21-5.33)*

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S8. SIRs of salmonellosis and campylobacteriosis in the 'electricity, gas, steam and air conditioning supply' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	10	17.9	0.56 (0.30-1.04)	52	34.1	1.53 (1.16-2.00)**
Female	7	5.3	1.32 (0.63-2.77)	15	9.8	1.53 (0.92-2.54)
Serovar/species						
S. Typhimurium	3	5.2	0.58 (0.19-1.80)	-	-	-
S. Enteritidis	7	9.3	0.75 (0.36-1.58)	-	-	-
Salmonella other	7	8.7	0.81 (0.39-1.70)	-	-	-
C. jejuni	-	-	-	60	41.0	1.46 (1.14-1.89)**
C. coli	-	-	-	7	2.9	2.41 (1.15-5.05)*
Age group						
16-19 years	0	0.2	-	0	0.3	-
20-29 years	5	4.5	1.10 (0.46-2.65)	9	7.7	1.18 (0.61-2.26)
30-39 years	6	4.7	1.29 (0.58-2.87)	15	8.1	1.84 (1.11-3.06)*
40-49 years	3	5.8	0.52 (0.17-1.60)	10	9.9	1.01 (0.54-1.87)
≥50 years	3	7.9	0.38 (0.12-1.17)	33	17.9	1.85 (1.31-2.60)***

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S9. SIRs of salmonellosis and campylobacteriosis in the 'manufacturing of machinery and equipment' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	55	39.2	1.40 (1.08-1.83)*	101	110.1	0.92 (0.75-1.11)
Female	4	6.2	0.64 (0.24-1.71)	13	14.6	0.89 (0.52-1.53)
Serovar/species						
S. Typhimurium	18	12.5	1.44 (0.91-2.29)	-	-	-
S. Enteritidis	19	15.3	1.24 (0.79-1.95)	-	-	-
Salmonella other	22	17.7	1.24 (0.82-1.89)	-	-	-
C. jejuni	-	-	-	103	116.2	0.89 (0.73-1.08)
C. coli	-	-	-	11	8.5	1.29 (0.71-2.33)
Age group						
16-19 years	0	2.3	-	4	4.0	1.00 (0.38-2.67)
20-29 years	18	12.8	1.41 (0.89-2.24)	20	24.3	0.82 (0.53-1.28)
30-39 years	15	9.5	1.58 (0.95-2.62)	24	23.8	1.01 (0.67-1.50)
40-49 years	20	9.8	2.04 (1.31-3.16)**	26	32.4	0.80 (0.55-1.18)
≥50 years	6	11.1	0.54 (0.24-1.21)	40	40.2	1.00 (0.73-1.36)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S10. SIRs of salmonellosis and campylobacteriosis in the 'other manufacturing' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	127	79.1	1.61 (1.35-1.91)***	233	164.2	1.42 (1.25-1.61)***
Female	56	36.8	1.52 (1.17-1.98)**	90	66.5	1.35 (1.10-1.66)**
Serovar/species						
S. Typhimurium	56	26.3	2.13 (1.64-2.76)***	-	-	-
S. Enteritidis	58	46.8	1.24 (0.96-1.60)	-	-	-
Salmonella other	69	42.7	1.62 (1.28-2.05)***	-	-	-
C. jejuni	-	-	-	300	214.7	1.40 (1.25-1.56)***
C. coli	-	-	-	23	16.0	1.44 (0.95-2.16)
Age group						
16-19 years	3	2.4	1.23 (0.40-3.81)	1	2.6	0.39 (0.05-2.75)
20-29 years	26	24.0	1.08 (0.74-1.59)	20	34.5	0.58 (0.37-0.90)*
30-39 years	27	19.1	1.41 (0.97-2.06)	51	31.3	1.63 (1.24-2.14)**
40-49 years	46	28.0	1.64 (1.23-2.20)**	101	55.3	1.83 (1.50-2.22)***
≥50 years	81	42.3	1.92 (1.54-2.38)***	150	107.0	1.40 (1.19-1.64)***

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S11. SIRs of salmonellosis and campylobacteriosis in the 'human health activities' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	123	102.8	1.20 (1.00-1.43)*	262	213.4	1.23 (1.09-1.39)**
Female	427	389.9	1.10 (1.00-1.20)	861	705.2	1.22 (1.14-1.31)***
Serovar/species						
S. Typhimurium	121	113.5	1.07 (0.89-1.27)	-	-	-
S. Enteritidis	212	193.4	1.10 (0.96-1.25)	-	-	-
Salmonella other	217	185.8	1.17 (1.02-1.33)*	-	-	-
C. jejuni	-	-	-	1,055	855.4	1.23 (1.16-1.31)***
C. coli	-	-	-	68	63.3	1.08 (0.85-1.36)
Age group						
16-19 years	27	19.2	1.41 (0.96-2.05)	32	26.0	1.23 (0.87-1.74)
20-29 years	205	176.4	1.16 (1.01-1.33)*	354	281.0	1.26 (1.14-1.40)***
30-39 years	92	93.7	0.98 (0.80-1.20)	182	166.4	1.09 (0.95-1.26)
40-49 years	107	98.9	1.08 (0.89-1.31)	244	200.2	1.22 (1.07-1.38)**
≥50 years	119	104.4	1.14 (0.95-1.36)	311	245.0	1.27 (1.14-1.42)***

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S12. SIRs of salmonellosis and campylobacteriosis in the 'residential care activities' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	81	60.2	1.34 (1.08-1.67)**	153	116.6	1.31 (1.12-1.54)**
Female	478	418.1	1.14 (1.05-1.25)**	828	718.8	1.15 (1.08-1.23)***
Serovar/species						
S. Typhimurium	144	113.9	1.26 (1.07-1.49)**	-	-	-
S. Enteritidis	220	188.8	1.17 (1.02-1.33)*	-	-	-
Salmonella other	195	175.6	1.11 (0.96-1.28)	-	-	-
C. jejuni	-	-	-	912	778.2	1.17 (1.10-1.25)***
C. coli	-	-	-	69	57.3	1.20 (0.95-1.53)
Age group						
16-19 years	57	45.7	1.25 (0.96-1.62)	67	64.4	1.04 (0.82-1.32)
20-29 years	214	176.5	1.21 (1.06-1.39)	329	271.2	1.21 (1.09-1.35)***
30-39 years	82	72.4	1.13 (0.91-1.41)	155	120.7	1.28 (1.10-1.50)**
40-49 years	98	88.6	1.11 (0.91-1.35)	196	169.0	1.16 (1.01-1.33)*
≥50 years	108	95.0	1.14 (0.94-1.37)	234	210.1	1.11 (0.98-1.27)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S13. SIRs of salmonellosis and campylobacteriosis in the 'social work activities without accommodation' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	27	31.3	0.86 (0.59-1.26)	67	61.9	1.08 (0.85-1.38)
Female	335	300.3	1.12 (1.00-1.24)*	619	558.6	1.10 (1.02-1.20)*
Serovar/species						
S. Typhimurium	95	77.8	1.22 (1.00-1.49)	-	-	-
S. Enteritidis	135	127.7	1.06 (0.89-1.25)	-	-	-
Salmonella other	132	126.1	1.05 (0.88-1.24)	-	-	-
C. jejuni	-	-	-	636	577.7	1.10 (1.02-1.19)*
C. coli	-	-	-	50	42.8	1.17 (0.89-1.54)
Age group						
16-19 years	37	18.7	1.98 (1.43-2.73)***	24	29.0	0.83 (0.56-1.24)
20-29 years	128	127.2	1.01 (0.85-1.20)	229	216.6	1.06 (0.93-1.20)
30-39 years	67	52.2	1.28 (1.01-1.63)*	102	97.9	1.04 (0.86-1.26)
40-49 years	59	61.8	0.95 (0.74-1.23)	155	124.5	1.25 (1.06-1.46)**
≥50 years	71	71.7	0.99 (0.78-1.25)	176	152.5	1.15 (1.00-1.34)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S14. SIRs of salmonellosis and campylobacteriosis in the 'accommodation' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	53	39.3	1.35 (1.03-1.76)*	103	63.7	1.62 (1.33-1.96)***
Female	49	57.2	0.86 (0.65-1.13)	110	91.3	1.20 (1.00-1.45)
Serovar/species						
S. Typhimurium	32	28.6	1.12 (0.79-1.58)	-	-	-
S. Enteritidis	45	37.2	1.21 (0.90-1.62)	-	-	-
Salmonella other	25	30.8	0.81 (0.55-1.20)	-	-	-
C. jejuni	-	-	-	202	144.9	1.39 (1.21-1.60)***
C. coli	-	-	-	11	10.2	1.08 (0.60-1.95)
Age group						
16-19 years	26	24.6	1.06 (0.72-1.55)	51	33.2	1.54 (1.17-2.02)**
20-29 years	54	46.1	1.17 (0.90-1.53)	100	70.5	1.42 (1.17-1.73)***
30-39 years	9	10.2	0.88 (0.46-1.70)	17	17.5	0.97 (0.61-1.57)
40-49 years	7	7.8	0.90 (0.43-1.88)	21	16.1	1.30 (0.85-2.00)
≥50 years	6	7.8	0.76 (0.34-1.70)	24	17.8	1.35 (0.90-2.01)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S15. SIRs of salmonellosis and campylobacteriosis in the 'activities of households as employers of domestic personnel' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	1	4.4	0.23 (0.03-1.61)	15	12.7	1.18 (0.71-1.95)
Female	25	18.8	1.33 (0.90-1.97)	67	45.1	1.49 (1.17-1.89)**
Serovar/species						
S. Typhimurium	3	6.0	0.50 (0.17-1.56)	-	-	-
S. Enteritidis	9	7.0	1.29 (0.67-2.47)	-	-	-
Salmonella other	14	10.2	1.37 (0.81-2.31)	-	-	-
C. jejuni	-	-	-	76	53.9	1.41 (1.13-1.77)**
C. coli	-	-	-	6	4.0	1.51 (0.68-3.37)
Age group						
16-19 years	4	1.7	2.36 (0.89-6.29)	4	3.2	1.24 (0.46-3.29)
20-29 years	4	8.4	0.48 (0.18-1.27)	18	16.9	1.06 (0.67-1.69)
30-39 years	4	2.9	1.38 (0.52-3.69)	12	7.4	1.62 (0.92-2.86)
40-49 years	8	4.0	2.02 (1.01-4.04)*	23	12.7	1.81 (1.20-2.72)**
≥50 years	6	6.3	0.95 (0.43-2.12)	25	17.6	1.42 (0.96-2.10)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S16. SIRs of salmonellosis and campylobacteriosis in the 'architectural and engineering activities; technical testing and analysis' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	59	76.8	0.77 (0.59-0.99)*	122	147.1	0.83 (0.69-0.99)*
Female	20	26.0	0.77 (0.50-1.19)	40	44.2	0.90 (0.66-1.23)
Serovar/species						
S. Typhimurium	17	25.2	0.67 (0.42-1.08)	-	-	-
S. Enteritidis	33	40.8	0.81 (0.58-1.14)	-	-	-
Salmonella other	29	36.8	0.79 (0.55-1.13)	-	-	-
C. jejuni	-	-	-	148	178.4	0.83 (0.71-0.97)*
C. coli	-	-	-	14	13.0	1.08 (0.64-1.82)
Age group						
16-19 years	3	3.0	1.00 (0.32-3.12)	5	3.6	1.37 (0.57-3.30)
20-29 years	33	37.4	0.88 (0.63-1.24)	45	52.1	0.86 (0.64-1.16)
30-39 years	16	24.8	0.65 (0.40-1.05)	42	44.1	0.95(0.70-1.29)
40-49 years	13	17.9	0.73 (0.42-1.25)	21	39.9	0.53 (0.34-0.81)**
≥50 years	14	19.8	0.71 (0.42-1.19)	49	51.6	0.95 (0.72-1.26)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S17. SIRs of salmonellosis and campylobacteriosis in the 'computer programming and consultancy' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	59	71.5	0.82 (0.64-1.06)	122	191.9	0.64 (0.53-0.76)***
Female	21	20.4	1.03 (0.67-1.58)	27	48.5	0.56 (0.38-0.81)**
Serovar/species						
S. Typhimurium	20	25.2	0.80 (0.51-1.23)	-	-	-
S. Enteritidis	22	30.9	0.71 (0.46-1.08)	-	-	-
Salmonella other	38	35.9	1.06 (0.77-1.46)	-	-	-
C. jejuni	-	-	-	136	224.4	0.61 (0.51-0.72)***
C. coli	-	-	-	13	16.1	0.81 (0.47-1.39)
Age group						
16-19 years	2	2.0	0.98 (0.24-3.91)	2	3.6	0.55 (0.14-2.20)
20-29 years	34	37.6	0.90 (0.65-1.26)	59	74.9	0.79 (0.61-1.02)
30-39 years	25	25.5	0.98 (0.66-1.45)	37	66.6	0.56 (0.40-0.77)***
40-49 years	14	15.9	0.88 (0.52-1.49)	30	54.8	0.55 (0.38-0.78)**
≥50 years	5	10.8	0.46 (0.19-1.11)	21	40.5	0.52 (0.34-0.80)**

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S18. SIRs of salmonellosis and campylobacteriosis in the 'financial service activities' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	70	77.2	0.91 (0.72-1.15)	115	149.1	0.77 (0.64-0.93)**
Female	55	68.6	0.80 (0.62-1.04)	85	107.2	0.79 (0.64-0.98)*
Serovar/species						
S. Typhimurium	20	32.4	0.62 (0.40-0.96)*	-	-	-
S. Enteritidis	47	60.7	0.77 (0.58-1.03)	-	-	-
Salmonella other	58	52.8	1.10 (0.85-1.42)	-	-	-
C. jejuni	-	-	-	190	238.8	0.80 (0.69-0.92)**
C. coli	-	-	-	10	17.5	0.57 (0.31-1.06)
Age group						
16-19 years	3	2.7	1.11 (0.36-3.45)	2	2.9	0.70 (0.17-2.78)
20-29 years	36	40.5	0.89 (0.64-1.23)	40	49.0	0.82 (0.60-1.11)
30-39 years	35	39.1	0.89 (0.64-1.25)	44	64.2	0.69 (0.51-0.92)*
40-49 years	22	33.7	0.65 (0.43-0.99)*	51	68.6	0.74 (0.56-0.98)*
≥50 years	29	29.8	0.97 (0.68-1.40)	63	71.7	0.88 (0.69-1.13)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S19. SIRs of salmonellosis and campylobacteriosis in the 'activities auxiliary to financial services and insurance activities' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	30	29.2	1.03 (0.72-1.47)	37	54.5	0.68 (0.49-0.94)*
Female	29	29.3	0.99 (0.69-1.42)	35	45.6	0.77 (0.55-1.07)
Serovar/species						
S. Typhimurium	4	13.9	0.29 (0.11-0.77)*	-	-	-
S. Enteritidis	30	24.2	1.24 (0.87-1.78)	-	-	-
Salmonella other	25	20.5	1.22 (0.83-1.81)	-	-	-
C. jejuni	-	-	-	66	93.2	0.71 (0.56-0.90)**
C. coli	-	-	-	6	6.9	0.87 (0.39-1.94)
Age group						
16-19 years	1	3.4	0.29 (0.04-2.09)	1	3.4	0.30 (0.04-2.09)
20-29 years	27	20.5	1.32 (0.90-1.92)	13	25.8	0.50 (0.29-0.87)*
30-39 years	14	13.6	1.03 (0.61-1.74)	19	22.9	0.83 (0.53-1.30)
40-49 years	8	11.0	0.73 (0.36-1.45)	21	24.2	0.87 (0.57-1.33)
≥50 years	9	10.0	0.90 (0.47-1.74)	18	23.9	0.75 (0.48-1.20)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S20. SIRs of salmonellosis and campylobacteriosis in the 'activities of head offices; management consultancy activities' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	73	83.2	0.88 (0.70-1.10)	133	176.8	0.75 (0.63-0.89)**
Female	61	58.2	1.05 (0.82-1.35)	58	102.1	0.57 (0.44-0.73)***
Serovar/species						
S. Typhimurium	28	34.9	0.80 (0.55-1.16)	-	-	-
S. Enteritidis	50	55.4	0.90 (0.68-1.19)	-	-	-
Salmonella other	56	51.2	1.09 (0.84-1.42)	-	-	-
C. jejuni	-	-	-	171	259.9	0.66 (0.57-0.76)***
C. coli	-	-	-	20	19.1	1.05 (0.68-1.62)
Age group						
16-19 years	13	6.9	1.88 (1.09-3.24)*	8	8.6	0.93 (0.46-1.85)
20-29 years	48	48.4	0.99 (0.75-1.32)	54	73.5	0.74 (0.56-0.96)*
30-39 years	22	31.5	0.70 (0.46-1.06)	38	58.9	0.64 (0.47-0.89)**
40-49 years	26	26.5	0.98 (0.67-1.44)	37	63.7	0.58 (0.42-0.80)**
≥50 years	25	28.2	0.89 (0.60-1.31)	54	74.2	0.73 (0.56-0.95)*

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S21. SIRs of salmonellosis and campylobacteriosis in the 'activities of membership organizations' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	33	40.1	0.82 (0.58-1.16)	56	67.0	0.84 (0.64-1.09)
Female	41	40.9	1.00 (0.74-1.36)	57	73.0	0.78 (0.60-1.01)
Serovar/species						
S. Typhimurium	16	20.8	0.77 (0.47-1.25)	-	-	-
S. Enteritidis	26	32.5	0.80 (0.54-1.18)	-	-	-
Salmonella other	32	27.7	1.15 (0.82-1.63)	-	-	-
C. jejuni	-	-	-	107	130.1	0.82 (0.68-0.99)*
C. coli	-	-	-	6	9.9	0.61 (0.27-1.35)
Age group						
16-19 years	13	11.9	1.09 (0.63-1.88)	10	11.2	0.89 (0.48-1.66)
20-29 years	20	22.1	0.90 (0.58-1.40)	26	33.7	0.77 (0.53-1.13)
30-39 years	12	12.9	0.93 (0.53-1.64)	19	22.1	0.86 (0.55-1.35)
40-49 years	8	14.2	0.56 (0.28-1.12)	17	28.1	0.60 (0.38-0.97)*
≥50 years	21	20.0	1.05 (0.69-1.61)	41	44.9	0.91 (0.67-1.24)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S22. SIRs of salmonellosis and campylobacteriosis in the 'education' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	169	190.7	0.89 (0.76-1.03)	382	376.7	1.01 (0.92-1.12)
Female	259	311.6	0.83 (0.74-0.94)**	475	557.9	0.85 (0.78-0.93)***
Serovar/species						
S. Typhimurium	93	113.9	0.82 (0.67-1.00)	-	-	-
S. Enteritidis	161	200.6	0.80 (0.69-0.94)**	-	-	-
Salmonella other	174	187.8	0.93 (0.80-1.07)	-	-	-
C. jejuni	-	-	-	795	870.0	0.91 (0.85-0.98)*
C. coli	-	-	-	62	64.6	0.96 (0.75-1.23)
Age group						
16-19 years	20	12.9	1.55 (1.00-2.40)	22	17.6	1.25 (0.83-1.90)
20-29 years	113	148.4	0.76 (0.63-0.92)**	244	234.8	1.04 (0.92-1.18)
30-39 years	80	83.0	0.96 (0.77-1.20)	134	153.0	0.88 (0.74-1.04)
40-49 years	93	104.0	0.89 (0.73-1.10)	146	183.4	0.80 (0.68-0.94)**
≥50 years	122	154.0	0.79 (0.66-0.95)*	311	345.8	0.90 (0.80-1.01)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S23. SIRs of salmonellosis and campylobacteriosis in the 'crop and animal production and hunting' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	76	77.5	0.98 (0.78-1.23)	108	121.1	0.89 (0.74-1.08)
Female	29	41.7	0.70 (0.48-1.00)	33	60.1	0.55 (0.39-0.77)**
Serovar/species						
S. Typhimurium	40	35.8	1.12 (0.82-1.52)	-	-	-
S. Enteritidis	47	47.5	0.99 (0.74-1.32)	-	-	-
Salmonella other	18	35.9	0.50 (0.32-0.80)**	-	-	-
C. jejuni	-	-	-	134	168.8	0.79 (0.67-0.94)**
C. coli	-	-	-	7	12.4	0.56 (0.27-1.18)
Age group						
16-19 years	33	36.9	0.90 (0.64-1.26)	50	44.5	1.12 (0.85-1.48)
20-29 years	35	38.8	0.90 (0.65-1.26)	41	52.5	0.78 (0.58-1.06)
30-39 years	11	14.7	0.75 (0.41-1.35)	13	22.2	0.59 (0.34-1.01)
40-49 years	16	13.9	1.15 (0.71-1.88)	14	27.6	0.51 (0.30-0.86)*
≥50 years	10	14.9	0.67 (0.36-1.25)	23	34.4	0.67 (0.44-1.01)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S24. SIRs of salmonellosis and campylobacteriosis in the 'construction of buildings' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	58	56.1	1.03 (0.80-1.34)	123	151.2	0.81 (0.68-0.97)*
Female	6	8.1	0.74 (0.33-1.65)	18	18.5	0.97 (0.61-1.54)
Serovar/species						
S. Typhimurium	22	18.2	1.21 (0.80-1.84)	-	-	-
S. Enteritidis	22	22.0	1.00 (0.66-1.52)	-	-	-
Salmonella other	20	24.0	0.83 (0.54-1.29)	-	-	-
C. jejuni	-	-	-	136	158.0	0.86 (0.73-1.02)
C. coli	-	-	-	5	11.7	0.43 (0.18-1.02)
Age group						
16-19 years	2	3.1	0.64 (0.16-2.56)	4	5.4	0.75 (0.28-1.99)
20-29 years	29	22.8	1.27 (0.88-1.83)	35	42.5	0.82 (0.59-1.15)
30-39 years	12	11.8	1.01 (0.58-1.79)	26	30.3	0.86 (0.58-1.26)
40-49 years	8	10.2	0.78 (0.39-1.56)	28	34.4	0.81 (0.56-1.18)
≥50 years	13	16.1	0.81 (0.47-1.39)	48	57.1	0.84 (0.63-1.11)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S25. SIRs of salmonellosis and campylobacteriosis in the 'wholesale trade' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	284	288.8	0.98 (0.88-1.10)	517	592.7	0.87 (0.80-0.95)**
Female	119	142.8	0.83 (0.70-1.00)*	199	253.0	0.79 (0.68-0.90)**
Serovar/species						
S. Typhimurium	103	109.9	0.94 (0.77-1.14)	-	-	-
S. Enteritidis	168	168.6	1.00 (0.86-1.16)	-	-	-
Salmonella other	132	153.1	0.86 (0.73-1.02)	-	-	-
C. jejuni	-	-	-	665	788.1	0.84 (0.78-0.91)***
C. coli	-	-	-	51	57.6	0.89 (0.67-1.17)
Age group						
16-19 years	37	32.2	1.15 (0.83-1.59)	41	41.2	1.00 (0.73-1.35)
20-29 years	137	147.8	0.93 (0.78-1.10)	189	223.3	0.85 (0.73-0.97)*
30-39 years	94	96.2	0.98 (0.80-1.20)	164	181.8	0.90 (0.77-1.05)
40-49 years	68	78.4	0.87 (0.68-1.10)	174	195.2	0.89 (0.77-1.03)
≥50 years	67	76.9	0.87 (0.69-1.11)	148	204.2	0.72 (0.62-0.85)***

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S26. SIRs of salmonellosis and campylobacteriosis in the 'land transport' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	137	153.2	0.89 (0.76-1.06)	315	293.6	1.07 (0.96-1.20)
Female	20	32.0	0.63 (0.40-0.97)*	50	51.4	0.97 (0.74-1.28)
Serovar/species						
S. Typhimurium	40	45.1	0.89 (0.65-1.21)	-	-	-
S. Enteritidis	65	75.3	0.86 (0.68-1.10)	-	-	-
Salmonella other	52	64.7	0.80 (0.61-1.05)	-	-	-
C. jejuni	-	-	-	337	321.3	1.05 (0.94-1.17)
C. coli	-	-	-	28	23.7	1.18 (0.81-1.71)
Age group						
16-19 years	7	9.6	0.73 (0.35-1.53)	12	9.3	1.28 (0.73-2.26)
20-29 years	48	49.7	0.97 (0.73-1.28)	66	65.1	1.01 (0.80-1.29)
30-39 years	25	33.5	0.75 (0.50-1.10)	47	51.9	0.91 (0.68-1.21)
40-49 years	26	38.2	0.68 (0.46-1.00)	77	77.0	1.00 (0.80-1.25)
≥50 years	51	54.2	0.94 (0.72-1.24)	163	141.7	1.15 (0.99-1.34)

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S27. SIRs of salmonellosis and campylobacteriosis in the 'services to buildings and landscape activities' sector

	Salmonellosis			Campylobacteriosis		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Gender						
Male	77	69.3	1.11 (0.89-1.39)	106	117.2	0.90 (0.75-1.09)
Female	107	106.9	1.00 (0.83-1.21)	141	170.4	0.83 (0.70-0.98)*
Serovar/species						
S. Typhimurium	46	44.7	1.03 (0.77-1.37)	-	-	-
S. Enteritidis	81	71.9	1.13 (0.91-1.40)	-	-	-
Salmonella other	57	59.7	0.95 (0.74-1.24)	-	-	-
C. jejuni	-	-	-	225	267.5	0.84 (0.74-0.96)**
C. coli	-	-	-	22	20.0	1.10 (0.72-1.67)
Age group						
16-19 years	23	23.4	0.98 (0.65-1.48)	17	23.7	0.72 (0.45-1.15)
20-29 years	46	60.8	0.76 (0.57-1.01)	75	84.7	0.89 (0.71-1.11)
30-39 years	38	30.6	1.24 (0.90-1.70)	45	49.4	0.91 (0.68-1.22)
40-49 years	37	31.3	1.18 (0.86-1.63)	63	62.7	1.00 (0.78-1.29)
≥50 years	40	30.2	1.33 (0.97-1.81)	47	66.9	0.70 (0.53-0.93)*

Obs: observed numbers; Exp: expected numbers; SIR: standardized incidence ratio. *p<0.05; **p<0.01; ***p<0.001.

Table S28. Risk group characteristics

	Live animals	Food production	Food sale
Number of people exposed			
Salmonellosis data (1999-2016)			
Total	240,993	2,037,210	224,051
Male	154,471	962,653	101,096
Female	86,522	1,074,557	142,955
Campylobacteriosis data (2004-2016)			
Total	172,978	1,666,621	178,427
Male	108,825	791,387	75,799
Female	64,153	875,234	102,628
Median age at entry (IQR)	27 years (20-39)	20 years (17-28)	23 years (18-36)
Median age at infection (IQR)			
Salmonellosis	26 years (20-38)	22 years (19-29)	21 years (18-33)
Campylobacteriosis	31 years (21-49)	23 years (20-34)	25 years (20-39)

Chapter 3

OCCUPATIONAL EXPOSURE AND RISK OF COLON CANCER: A NATIONWIDE REGISTRY STUDY

J.W. Duijster
L. Mughini-Gras
J.J.C. Neefjes
E. Franz

Published in: *BMJ Open,* 2021: 11(8)

Abstract

Objectives: While colon cancer (CC) risk is associated with several lifestyle-related factors, including physical inactivity, smoking and diet, the contribution of occupation to CC morbidity remains largely unclear. Growing evidence indicates that gastrointestinal infections like salmonellosis could contribute to CC development. We performed a nationwide registry study to assess potential associations between occupation (history) and CC, including also those occupations with known increased exposure to gastrointestinal pathogens like *Salmonella*.

Methods: Person-level occupational data for all residents in The Netherlands were linked to CC diagnosis data. Differences in the incidence of (overall, proximal and distal) CC among occupational sectors and risk groups were tested for significance by calculating standardized incidence ratios (SIRs) with 95% CIs using the general population as reference group. Effects of gender, age, exposure duration and latency were also assessed.

Results: Significant differences in CC incidence were observed only for a few occupational sectors, including the manufacturing of rubber and plastics, machinery and leather, the printing sector and the information service sector (SIRs 1.06–1.88). No elevated risk of CC was observed among people with increased salmonellosis risk through occupational exposure to live animals, manure or among those working in the sale of animal-derived food products (SIRs 0.93–0.95, 0.81–0.95 and 0.93–1.09 for overall, proximal and distal CC, respectively).

Conclusions: The results of this study suggest that occupation in itself provides a relatively small contribution to CC incidence. This is consistent with previous studies where a similar degree of variation in risk estimates was observed. The lack of an association with the high-risk occupations for salmonellosis might be due to higher levels of physical activity, a known protective factor for CC and other diseases, of people working in the agricultural sector, which might outweigh the potential *Salmonella*-associated risk of CC.

Background

With over a million new diagnoses, colon cancer (CC) was the third most frequent malignancy worldwide in 2018 [1]. In The Netherlands (~17 million population), the age-standardized incidence rate of CC is 1.83 per 10 000 inhabitants.1 While the survival rates of patients with CC continue to improve as a result of screening programs and targeted treatments [2], the past three decades have been characterized by an increase in the incidence of colorectal cancer among people aged <50 years in several high-income countries [3]. In The Netherlands, the annual percent change of colorectal cancer between 2001 and 2016 was 2.1 for people aged 20–39 years and 2.3 for people aged 40–49 years [4]. The reason for this increase remains largely unknown.

Apart from genetic background (i.e., inheritable CC, such as hereditary non-polyposis colorectal cancer and familial adenomatous polyposis), the main risk factors for CC comprise dietary and lifestyle factors, including the consumption of red and processed meat, smoking, alcohol consumption, obesity and limited physical activity [5, 6]. The latter factor in particular has been addressed in several studies showing that people with sedentary jobs (e.g., white-collar workers) are at increased risk of colorectal and CCs [7]. Occupational exposure to chemical compounds used in several industrial productions, such as leather, metals, plastic and rubber, as well as asbestos, has also been reported to increase the risk of CC [8–10]. Moreover, in the past decade, the role of bacterial infections in cancer development has gained momentum [11]. For the gastrointestinal system, these infections concern mainly Helicobacter pylori and *Salmonella* Typhi as causative agents of gastric cancer and gallbladder carcinoma, respectively [12, 13], as well as (severe) non-typhoidal *Salmonella* infection for CC [14–17] and colibactin secreting *Escherichia coli* strains for colorectal cancer [18, 19]. Whether repeated, low-dose exposure to *Salmonella* leading to asymptomatic or paucisymptomatic infection, for instance in occupational settings, is also a risk factor for CC, remains unclear and has not yet been investigated.

While occupational exposure to carcinogens for among others lung and skin cancers, have been extensively documented, the role of occupation in CC epidemiology is complex and ambiguous [20, 21] Moreover, apart from a large study in five northern European countries making use of multiple-year census data [22], most studies have addressed only specific occupational groups (e.g., nurses, farmers, asbestos plant workers), rather than the total employed population, and these studies did not consider the occupational risk cumulatively based on exposure history, but rather the effect of occupation at a given moment in time [23–25]. The primary aim of this nationwide registry-based cohort study was to assess the potential association between occupations with known increased exposure to zoonotic pathogens like *Salmonella* and CC incidence. We also extended the analyses to the whole spectrum of occupational exposures in The Netherlands between 1999 and 2016. Rectal

cancers were not included in the analysis, as in a previous Dutch cohort study, a significant association between non-typhoidal *Salmonella* and CC was only found for the proximal part of the colon [14]. Moreover, colon and rectal cancers differ from each other with respect to molecular carcinogenesis, clinical symptoms and risk factors, with for instance high levels of physical activity being a protective factor for colon but not rectal cancer [26].

Methods

Data registries and linkage

We assessed the association between occupation and CC risk by linking two national registries in The Netherlands. Statistics Netherlands (CBS) provided person-level, de-identified occupational history data for all Dutch residents at any moment in time, including changes in occupational group, specific functions therein, and employer, between January 1999 and December 2016. The occupational groups were coded according to the European Nomenclature of Economic Activities (NACE) based on the economic activity of a registered company providing employment [27]. The NACE data are structured in four hierarchical levels (sections, divisions, groups and classes) by a five-digit code, allowing for analyses at different levels, as described in more detail elsewhere [28]. The second data set was retrieved from The Netherlands Cancer Registry and contained 135 909 CC diagnoses between January 2000 and December 2016, of which 74 254 pertained to the proximal colon (International Classification of Diseases, 10th Revision (ICD-10) codes C180–C185) and 61 655 to the distal colon (ICD-10 codes C186, C187).

Sectors and risk goups

First, we assessed the risk of CC at the NACE-level of division where all occupations are mutually classified into 86 different divisions (hereafter referred to as 'sectors') [25]. We then defined three risk groups (based on the most detailed NACE-codes of the occupations), including occupations with contact with live animals or animal manure (e.g., farmers, veterinarians), occupations involved in the production and handling of animal-derived food products (e.g., cooks, bakers) and occupations involved in the sale of animal-derived foods (e.g., butchers). This risk group classification was in accordance with the risk groups used in a previous study assessing the occupational risk of Salmonella infection [27].

Statistical analysis

For the data analysis, individuals entered the at-risk period after 1 year of registered employment in a given occupational group of interest (i.e., a sector or a risk group) or when

reaching the age of 20 years, whichever came last. Hence, the earliest date of onset of follow-up was 1 January 2000. CC diagnosis under the age of 20 years is rare (i.e., there were only 14 CC diagnoses in people aged 10–19 years during the total study period) and this is mostly the result of inheritable factors.29 The follow-up period ended at the date of CC diagnosis (i.e., the event of interest), date of death or the end of the study period (31 December 2016), whichever came first. No censoring on emigration was applied. Individuals were allowed to be included in multiple sectors or risk groups simultaneously. We excluded people from the analysis who were diagnosed with CC before onset of employment in a sector or risk group or were diagnosed after less than 1 year of exposure. First, we calculated the incidence rates (IRs) of CC (overall and per subsite) per 10 000 person-years at risk in the total employed population, by gender, age group (<50 and ≥50 years), duration of exposure (i.e., the number of years employed (<2; 2–4; 5–9; ≥10 years)) and latency (i.e., the number of years since the onset of exposure (1–4; 5–9; ≥10 years)). Second, the risk of CC in the 86 sectors and three risk groups was compared with the risk of CC in the general Dutch population which was used as the baseline reference risk. To this end we calculated standardized incidence ratios (SIRs) of CC (overall, proximal and distal) in men and women separately (and overall) by dividing the observed number of CC diagnoses in a sector or risk group by the expected number of diagnoses based on age-matched (5-year bands), gender-matched, calendar year-matched (1-year bands) and subsite-matched CC IRs in the Dutch population. For the sectors a stratified analysis was done for the age groups <50 years and ≥50 years, while for the three risk groups, stratified analyses were done by subsite, gender, age group (20–39; 40–49; ≥50 years), duration and latency. The 95% CIs for the SIRs were calculated assuming a Poisson distribution. In the analyses, p values <0.05 were considered statistically significant. Statistical analysis was performed using Stata V.16 (StataCorp LP, College Station, USA).

Patient and public involvement

No patients were involved.

Results

Cohort description

The total cohort comprised 11 136 434 individuals with registered employment in (part of) the study period. The majority of the cohort consisted of men (54%), although the percentage of women increased over the years. CC was diagnosed in 44 778 individuals over the whole study period (2000–2016), corresponding to an overall average IR of 3.03 (95% CI

3.01 to 3.06) CC cases per 10 000 person-years at risk (Table 1). For both colon subsites, the IR was higher in men than in women (Table 1). On average, women were diagnosed with CC at a lower age (median: 57.4 years; IQR: 50.6–62.8) as compared with men (median: 61 years; IQR: 55.1–65.9). Online supplementary table S1 and S2 show the SIRs for proximal, distal and overall CC among men and women, respectively, with employment history in at least 1 of the 84 different sectors. SIRs for overall CC ranged from 0.68 to 1.45 in men and 0.66 to 2.53 in women (Supplementary tables S1 and S2). The SIRs of CC in the age group under 50 years versus above 50 years differed substantially within and between sectors (Supplementary table S3). Among sectors with at least 10 observed CC diagnoses in both age groups, SIRs differed on average 12.5% (range: 0%–48%) between the two age groups within a sector. For overall CC, a significantly increased risk was observed in men, women and/or both combined for seven sectors, whereas for proximal and distal CC, this was three and six sectors, respectively. Significantly decreased risks were observed in nine sectors for overall CC, eight for proximal CC and six for distal CC for men, women and/or both combined.

Table 1. Incidence rates (IRs) of colon cancer (overall, proximal and distal) in the employed population

	Colon cancer - overall		Proximal colon cancer		Distal colon cancer	
	N	IR* (95%CI)	N	IR* (95%CI)	N	IR* (95%CI)
All	44,778	3.03 (3.01-3.06)	21,515	1.46 (1.44-1.48)	23,263	1.58 (1.56-1.60)
Males	29,446	3.63 (3.59-3.68)	13,487	1.66 (1.64-1.69)	15,959	1.97 (1.94-2.00)
Females	15,332	2.30 (2.27-2.34)	8,028	1.21 (1.18-1.23)	7,304	1.10 (1.07-1.12)
Age group						
<50 years	5,479	0.54 (0.52-0.55)	2,820	0.28 (0.27-0.29)	2,659	0.26 (0.25-0.27)
≥50 years	39,299	8.65 (8.57-8.74)	18,695	4.12 (4.06-4.18)	20,604	4.54 (4.47-4.60)
Exposure duration						
<2 years	3,401	2.35 (2.27-2.43)	1,691	1.17 (1.12-1.23)	1,710	1.18 (1.13-1.24)
2-4 years	10,959	2.73 (2.68-2.78)	5,511	1.37 (1.34-1.41)	5,448	1.36 (1.32-1.39)
5-9 years	14,376	2.86 (2.81-2.91)	7,006	1.39 (1.36-1.43)	7,370	1.47 (1.43-1.50)
≥10 years	16,042	3.76 (3.70-3.82)	7,307	1.71 (1.67-1.75)	8,735	2.05 (2.00-2.09)
Latency†						
1-4 years	4,422	1.31 (1.27-1.35)	2,245	0.66 (0.64-0.69)	2,177	0.64 (0.62-0.67)
5-9 years	9,649	1.96 (1.93-2.00)	4,806	0.98 (0.95-1.01)	4,843	0.99 (0.96-1.01)
≥10 years	30,706	4.75 (4.70-4.80)	14,463	2.24 (2.20-2.27)	16,243	2.48-2.55)

***** Incidence rate (IR) per 10,000 person-years at risk. **†** Period between start at-risk period and CC diagnosis.

Occupations with increased risk

Significantly elevated SIRs for overall CC were found for men with employment history in manufacturing of rubber and plastics (SIR 1.14), sale and repair of motor vehicles (SIR 1.10), land transport (SIR 1.06), information service activities (SIR 1.45), (re)insurance and pension funding (SIR 1.12) and real estate activities (SIR 1.11) (Figure 1). Concerning distal CC, significantly increased risk was observed for six sections, compared with one section for proximal part of the colon (Figure 1). Within the section of rubber and plastic manufacturing, SIRs were increased for both colon subsites (proximal: SIR 1.10 (95% CI 0.92 to 1.31); distal: SIR 1.17 (95% CI 1.00 to 1.37)). For the sale and repair of motor vehicles and land transport sections, SIRs were highest for the distal part, whereas for the information service activities the higher risk concerned the proximal colon only (SIR 1.88; 95% CI 1.25 to 2.83) (Figure 1). Among men with employment history in (re)insurance and real estate, the risk was most pronounced for distal CC and among people aged ≥50 years, as compared with the general population (average SIRs 1.19 (range 1.04 to 1.35) and 1.13 (range 1.01 to 1.26), respectively). Additionally, an increased risk concerning only the distal colon was observed among those with employment history in printing and reproduction of recorded media (SIR 1.16) and manufacturing of machinery (SIR 1.17), with the highest SIRs in the older age group and among people with long-term exposure. Among women, a significant increased SIR for overall CC was only observed for those employed in manufacturing of leather, with a SIR of 2.39 (95% CI 1.39 to 4.12) for proximal CC and a SIR of 0.87 (95% CI 0.33 to 2.32) for distal CC, although the observed numbers were relatively low.

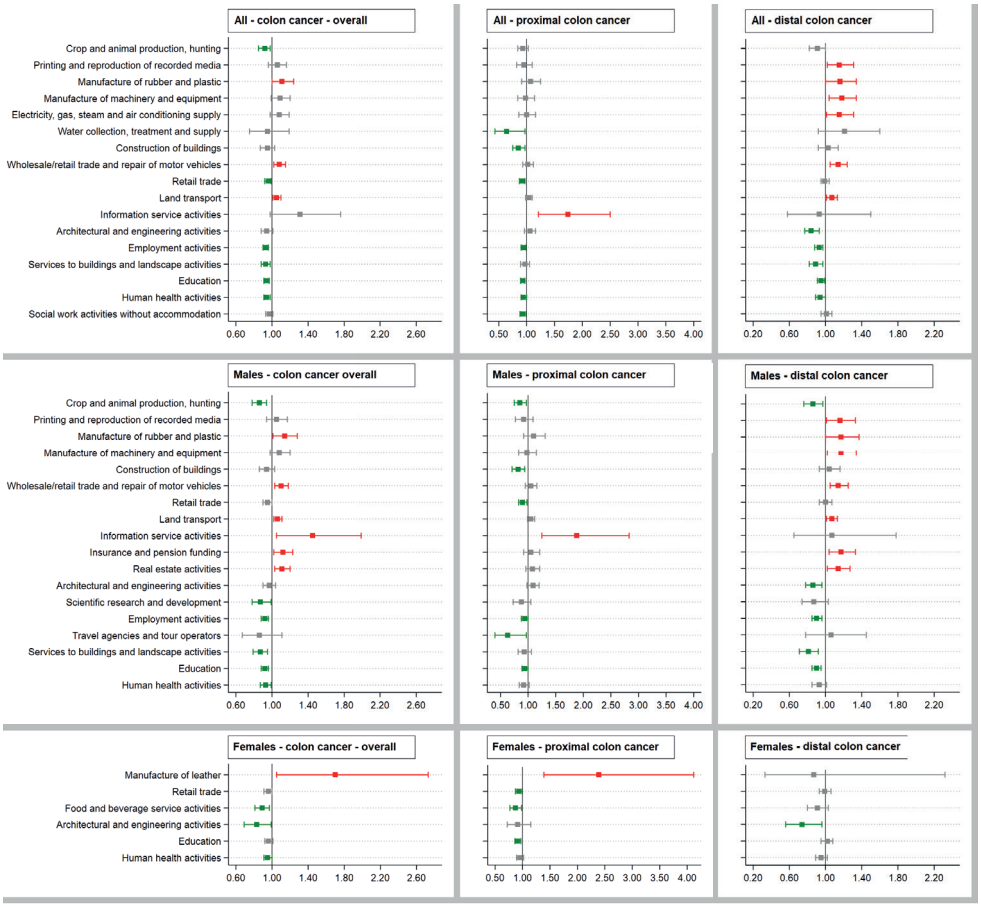


Figure 1. Standardized incidence ratios (SIRs) (squares) with 95% CIs (bars) of colon cancer (overall, proximal and distal) in the total employed population and in men and women separately per sector. Red, significantly increased SIR; green, significantly decreased SIR; grey, non-significant SIR.

Occupations with decreased risk

Among men, for 10 sectors, significantly decreased risks were observed, of which three were significant for both colon subsites (Figure 1). In women, CC risk was significantly lower for five sectors. In the agricultural sector (crop and animal production), SIRs of 0.85 (95% CI 0.75 to 0.97) and 0.86 (95% CI 0.76 to 0.97) for distal CC were found in men, whereas for women the SIRs were slightly higher than 1, though not significant. The SIR for proximal CC in men was particularly low in the age group ≥50 years (SIR 0.83; 95% CI 0.72 to 0.96), compared with the group under 50 years (SIR 0.99; 95% CI 0.71 to 1.37). The opposite was true for distal

CC where the SIR for individuals <50 years was 0.64 (95% CI 0.40 to 1.01) (≥50 years: SIR 0.88; 95% CI 0.78 to 1.00). Within the education sector and the sector of retail trade, significant lower risks of proximal CC were observed for both men and women (Figure 1). Moreover, lower risk of distal CC was found for men (SIR 0.86; 95% CI 0.78 to 0.96) and women (SIR 0.74; 95% CI 0.56 to 0.96) with employment history in architectural and engineering activities. In the sector of food and beverage service activities (e.g., cooks, waiters) the risk of overall CC and proximal CC was lower for women exclusively (SIR 0.87; 95% CI 0.77 to 0.98). Similarly, a significant, though marginal lower risk of overall CC was observed for the human healthcare sector (SIR 0.95; 95% CI 0.91 to 1.00) in women. In this sector, the risk appeared lower for overall CC in the age group ≥50 years (Supplementary table S3).

Risk groups

We also assessed specifically the incidence of CC in three groups with increased occupational exposure to zoonotic pathogens with oncogenic potential like *Salmonella*, as showed in a previous study [28]. All three groups showed a marginally decreased risk for overall CC (Table 2). Within the group involved in the sale of animal-derived food products, the SIRs were lowest for proximal CC (SIR 0.81; 95% CI 0.68 to 0.97), whereas for distal CC, the SIRs were above 1 for both men and women (Table 2).

Table 2. SIRs of colon cancer (overall, proximal and distal) per risk group as compared with the general Dutch population

	Colon cancer – overall			Proximal colon cancer			Distal colon cancer		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
LIVE ANIMALS/ ANIMAL MANURE									
All	377	398.0	0.95 (0.86-1.05)	179	189.3	0.95 (0.82-1.09)	198	208.7	0.95 (0.83-1.09)
Males	301	318.0	0.95 (0.85-1.06)	140	147.3	0.95 (0.81-1.12)	161	170.8	0.94 (0.81-1.10)
Females	76	80.0	0.95 (0.76-1.19)	39	42.0	0.93 (0.68-1.27)	37	38.0	0.97 (0.71-1.35)
Age group									
20-39 years	16	14.0	1.15 (0.70-1.87)	–	–	–	–	–	–
40-49 years	38	42.7	0.89 (0.65-1.22)	–	–	–	–	–	–
≥50 years	323	341.3	0.95 (0.85-1.06)	172	159.5	0.95 (0.81-1.11)	172	181.8	0.95 (0.81-1.10)
Exposure duration									
1-2 years	62	53.5	0.82 (0.64-1.05)	26	36.7	0.71 (0.48-1.04)	36	38.9	0.93 (0.67-1.28)
2-4 years	173	166.4	1.04 (0.90-1.21)	81	80.2	1.01 (0.81-1.26)	92	86.2	1.07 (0.87-1.31)
5-9 years	94	102.5	0.92 (0.75-1.12)	53	48.2	1.10 (0.84-1.44)	41	54.3	0.76 (0.56-1.03)
≥10 years	48	53.5	0.90 (0.68-1.19)	19	24.0	0.79 (0.50-1.24)	29	29.4	0.98 (0.68-1.42)
Latency									
1-4 years	58	62.7	0.92 (0.71-1.20)	29	31.2	0.93 (0.65-1.34)	29	31.6	0.92 (0.64-1.32)
5-9 years	104	112.1	0.93 (0.77-1.12)	44	54.4	0.81 (0.60-1.09)	60	57.7	1.04 (0.81-1.34)
≥10 years	215	223.2	0.96 (0.84-1.10)	106	103.7	1.02 (0.84-1.24)	109	119.5	0.91 (0.76-1.10)
FOOD PRODUCTION/HANDLING									
All	1457	1565.1	0.93 (0.88-0.98)**	726	780.2	0.93 (0.86-1.00)	731	787.0	0.93 (0.86-1.00)*
Males	830	875.3	0.95 (0.88-1.01)	410	412.5	0.99 (0.90-1.09)	420	462.8	0.91 (0.82-1.00)*
Females	627	689.8	0.91 (0.84-0.98)*	316	367.7	0.86 (0.77-0.96)**	311	322.0	0.97 (0.86-1.08)
Age group									
20-39 years	118	120.5	0.98 (0.82-1.17)	77	80.6	0.96 (0.76-1.19)	41	39.9	1.03 (0.76-1.40)
40-49 years	167	186.2	0.90 (0.77-1.04)	77	90.0	0.86 (0.68-1.07)	90	96.2	0.94 (0.76-1.15)
≥50 years	1172	1258.4	0.93 (0.88-0.99)*	572	609.6	0.94 (0.86-1.02)	600	648.8	0.92 (0.85-1.00)

	Colon cancer – overall			Proximal colon cancer			Distal colon cancer		
	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Exposure duration									
1-2 years	282	315.7	0.89 (0.79-1.00)	148	162.5	0.91 (0.78-1.07)	134	153.2	0.87 (0.74-1.04)
2-4 years	528	580.7	0.91 (0.83-0.99)*	257	294.4	0.87 (0.77-0.99)*	271	286.4	0.95 (0.84-1.07)
5-9 years	405	427.7	0.95 (0.86-1.04)	203	210.7	0.96 (0.84-1.11)	202	217.0	0.93 (0.81-1.07)
≥10 years	242	240.9	1.00 (0.89-1.14)	118	112.7	1.04 (0.87-1.25)	124	128.2	0.97 (0.81-1.15)
Latency									
1-4 years	210	236.5	0.89 (0.78-1.02)	105	121.4	0.86 (0.71-1.05)	105	115.1	0.91 (0.75-1.10)
5-9 years	421	434.7	0.97 (0.88-1.07)	230	221.5	1.03 (0.91-1.18)	191	213.2	0.90 (0.78-1.03)
≥10 years	826	893.9	0.92 (0.86-0.99)*	391	437.3	0.89 (0.81-0.98)*	435	456.6	0.95 (0.87-1.05)
SALE OF FOOD PRODUCTS									
All	290	304.7	0.95 (0.85-1.07)	121	149.4	0.81 (0.68-0.97)*	169	155.3	1.09 (0.94-1.27)
Males	175	186.7	0.94 (0.81-1.09)	72	87.2	0.83 (0.66-1.04)	103	99.5	1.03 (0.85-1.26)
Females	115	118.0	0.97 (0.81-1.17)	49	62.2	0.79 (0.60-1.04)	66	55.8	1.18 (0.93-1.51)
Age group									
20-39 years	15	16.0	0.94 (0.57-1.56)	–	–	–	–	–	–
40-49 years	28	33.7	0.83 (0.57-1.20)	–	–	–	–	–	–
≥50 years	247	255.1	0.97 (0.85-1.10)	98	123.0	0.80 (0.65-0.97)*	149	132.1	1.13 (0.96-1.32)
Exposure duration									
1-2 years	62	64.3	0.96 (0.75-1.24)	28	32.2	0.87 (0.60-1.26)	34	32.0	1.06 (0.76-1.49)
2-4 years	101	119.4	0.85 (0.70-1.03)	45	59.1	0.76 (0.57-1.02)	56	60.3	0.93 (0.71-1.21)
5-9 years	84	83.1	1.01 (0.82-1.25)	32	40.3	0.79 (0.56-1.12)	52	42.8	1.22 (0.93-1.60)
≥10 years	43	37.9	1.13 (0.84-1.53)	16	17.7	0.90 (0.55-1.47)	27	20.2	1.34 (0.92-1.95)
Latency†									
1-4 years	39	50.7	0.77 (0.56-1.05)	16	25.4	0.63 (0.39-1.03)	23	25.3	0.91 (0.60-1.37)
5-9 years	86	88.2	0.97 (0.79-1.20)	44	43.7	1.01 (0.75-1.35)	42	44.5	0.94 (0.70-1.28)
≥10 years	165	165.8	1.00 (0.85-1.16)	61	80.3	0.76 (0.59-0.98)*	104	85.5	1.22 (1.00-1.47)*

Obs: observed numbers; Exp: expected numbers; SIR: Standardized Incidence Rate; CI: Confidence Interval. * p<0.05; ** p<0.01; *** p <0.001. – Numbers cannot be provided due to risk of subject identification. † Period between start at-risk period and CC diagnosis.

Discussion

In this study, we linked two nationwide registries to assess potential associations between occupation (history) and CC incidence in The Netherlands in order to identify possible risk-conferring exposures in the workplace for CC development. Moreover, we looked at specific occupations for which an increased risk of infection with zoonotic pathogens like *Salmonella* has been found [28], as *Salmonella* infection has shown to promote colon carcinogenesis in both epidemiological [14] and experimental [11, 13, 30] studies. In contrast to other malignancies, the magnitude of occupation as risk factor for CC is relatively small compared with the major nutritional and lifestyle-related risk factors. It is also more difficult to quantify due to confounding factors (e.g., smoking, physical activity), which are shared between some and differ substantially between other occupational groups [20, 22, 31, 32]. This nationwide study in a high-income country covering a broad range of occupational sectors therefore wants to contribute to the existing knowledge on the occupational exposures associated with increased CC risk.

We found significantly increased risks for CC in several occupational sectors. Most of the results of this study were confirmatory in nature and mirrored previous observations available in the literature. For instance, significantly increased risks of (overall, proximal and/or distal) CC were found in multiple industrial sectors with potential exposure to chemicals, including the manufacturing of rubber and plastics, machinery and leather, as well as the printing sector. Extensive research has been done to assess the carcinogenic risk of exposures to, for example, benzene, solvents and dyes in these industries. While the causal relations between, for example, working in the rubber industry and bladder cancer and leukemia [33], and working in the leather industry and pancreatic cancer [34], are well documented, the association with CC is more ambiguous. In a meta-analysis assessing occupational exposure and CC risk, relative risks of 1.16 (95% CI 0.99 to 1.36) for the rubber and plastic industry, 1.49 (95% CI 0.90 to 2.46) for the leather industry and 1.80 (95% CI 1.20 to 2.70) for the printing sector were reported [9].

A significantly increased risk of CC among people occupationally exposed to *Salmonella* via live animals or manure or through working in the sale of animal-derived food products, was not observed here. In the past decade, a growing number of experimental studies have unraveled the pathways by which pathogenic bacteria contribute to the development of cancer in the gastrointestinal tract. On infection, non-typhoidal *Salmonella* hijacks the host cell biology by introducing several effector proteins into the host cell. Specifically, acetyltransferase AvrA suppresses the immune response and apoptosis by inhibiting the

host-signaling pathway NF-κβ while enhancing epithelial cell proliferation by β-catenin signaling-pathway activation [11, 35]. Similarly, SopB-, SopE-, SopE2- and SptP-effector proteins can facilitate transformation of pre-transformed host cells by activating the AKT-pathways and MAPK-pathways [11]. This was shown in a study of Scanu *et al*. (2015) where mouse fibroblast and gallbladder organoids underwent irreversible transformation under conditions of an inactivated p53 tumor suppressor gene and an overexpressed c-MYC oncogene [13]. Besides laboratory evidence, the risk of developing proximal CC was found to be over twofold higher in people with a registered severe *Salmonella* infection in the past (SIR 2.12; 95% CI 1.38 to 3.09) [14]. This risk was specifically higher for individuals infected with *Salmonella* Enteritidis (SIR 2.97; 95% CI 1.73 to 4.76) and people aged <60 years at time of infection (SIR 1.54; 95% CI 1.09 to 2.10) [14]. Also, we showed in an earlier study that the risk of suffering severe salmonellosis was higher among people working with live animals or animal manure (SIR 1.82; 95% CI 1.49 to 2.23) and among people working in the sale of animal-derived food products (SIR 1.55; 95% CI 1.24 to 1.93) [27]. While these prior epidemiological studies focused on severe salmonellosis, here we particularly looked at the risk of CC after possible long-term occupational exposure to *Salmonella*, not per se leading to clinically-overt salmonellosis. The risk of CC in people working with live animals or manure appeared to be slightly reduced as compared with the general population. Acquired immunity against *Salmonella* in people with frequent exposure to such pathogens could also be an explanation for the observed findings, as the bacterium is more rapidly cleared from the body leaving less time for *Salmonella* to induce cellular transformation. Acquired immunity in the occupationally exposed population has been shown for *Campylobacter* [36, 37]. As both pathogens are epidemiologically comparable in that respect, a similar mechanism can be assumed for *Salmonella*. In addition, it is possible that the *Salmonella* serovars in livestock differ from those contributing to human cell transformation. Unraveling the exact molecular mechanism by which *Salmonella* contributes to CC development could clarify this.

Extensive research into the risk of different forms of cancer among farmers has been done in the past, most of which found a reduced risk of CC as compared with non-farmers [9, 31, 38]. On the one hand, this may be related to the beneficial effect of increased physical activity (a known protective factor for CC) of people working in the agricultural sector, which might therefore outweigh other risk factors [39]. Similarly, lower smoking rates have been reported among farmers compared with other occupational groups, which might have reduced the risk of developing CC as well [40], though risk estimates only slightly differed with and without adjustment for tobacco use and alcohol consumption in a large European cohort study [20]. Although a previous study reported an increased CC incidence among

poultry farmers, this could not be confirmed here [24]. With regard to people working with raw meat, some studies reported a slightly increased (often non-significant) risk for overall CC among butchers and employees of meat-processing plants, however, a pooled analysis of multiple studies did not reveal a significant association [9, 41]. While physical activity can reduce the risk of CC up to 24% [42], we did not observe a clear overall risk difference across sectors with a higher level of occupational physical activity versus more sedentary sectors (Supplementary tables S2 and S3). Nonetheless, SIRs of CC were significantly increased for three sectors, with mainly sedentary jobs, including the sectors of information service activities, real estate and insurance (Figure 1). Conversely, significant risk deficit was found for a number of sectors with limited sedentary professions such as construction workers, farmers, teachers, the retail trade sector, the health sector and the services to buildings and landscapes sector (i.e., interior cleaning of buildings and maintenance of public parks and gardens) (Figure 1). Previous studies showed non-significant risk estimates close to 1 for both the education sector and the health sector [9, 31], whereas for construction workers a significantly reduced risk of 5%–20% was found earlier [31]. Also, for some sectors, particularly those associated with higher education and income, it is plausible that the lower CC risk is partly explained by an average healthier lifestyle, as it was previously shown that smoking rates and overweight/obesity were lower among teachers as compared with non-teachers [43].

Apart from differences in incidence ratios across sectors, we found small differences between the sexes within some sectors. The NACE-code(s) linked to an individual person are based on the economic activity of the company/organization at which he/she is employed, rather than the actual job task or individual measurements. Hence, due to this limitation, we could not disentangle possible gender disparities resulting from different job tasks of men versus women within a sector. Likewise, people might have been misclassified into a high risk group while their actual occupation does not involve exposure to zoonotic pathogens (e.g., people with an office job at a slaughterhouse company). Also, whether an individual is working part-time or full-time was not registered at the person-level in the occupational records. Hence, this might have led to an overestimation of the exposure duration of women as compared with men, as in The Netherlands over 70% of women have a part-time job [44]. Furthermore, although working environments in, for example, industries have become safer in the last decades with regard to exposure to hazardous/carcinogenic substances [45], SIRs were not consistently higher in those diagnosed with CC at an age of ≥50 years compared with the younger age group (Supplementary table S3). Probably, the study period is too small to evaluate potential causes of a risk difference between age groups within sectors.

For this study, we lacked information at the individual level about emigration. This has, to some extent, led to an underestimation of the cancer risk as a result of the overestimation of the total number of person-years at risk. Moreover, we lacked information about major risk factors, such as smoking, dietary habits, alcohol consumption and body weight, as these data are not usually routinely collected at the population level in national registries. Given that tobacco use and alcohol intake differ between occupational groups, there is evidence that adjustment for these variables could yield slightly different risk estimates [20, 46]. Likewise, consumption of red and processed meat, another risk factor for CC, is inversely correlated with income and educational level [47-49]. Yet, adjustment for these time-varying risk factors is impossible in a registry-based study with national coverage such as ours, and would require another type of study design.

In conclusion, only a few significant differences in CC incidence as a function of occupational exposures in different sectors were observed. This is unlike other forms of cancer, but is consistent with other literature on occupational risks of CC. The occupational exposures associated with increased CC risk were mainly those in the industrial sectors with potential exposure to toxic chemicals, such as the manufacturing of rubber and plastics, machinery and leather and the printing sector. These observations stress the need of continuous improvement of workplace-safety as well as more research in the future to assess whether these policies adequately reduce the incidence of cancers related to occupation. A significantly increased risk of CC among people occupationally exposed to live animals or manure or working in the sale of animal-derived food products (i.e., the groups with increased salmonellosis risk) was not observed. This may be related to both the beneficial effect of increased physical activity (a known protective factor for CC) of people working in the agricultural sector, which might outweigh other risk factors, as well as an overestimation of the number of people truly occupationally exposed to zoonotic pathogens due to the lack of detailed job content data at the individual level. Large population-based epidemiological studies based on national registries, such as the present study, have the advantage to allow for inference from available large data sets, providing an inventory of differences in CC incidence among occupational sectors that helps understanding the epidemiology of CC from a public health perspective. Yet, there are many other factors playing a role in CC development that cannot be properly controlled for in this type of studies. Therefore, understanding how different factors contribute to cancer formation can result in the design of studies with defined and coherent groups to limit the number of variables. Yet, the contribution of occupation to CC is limited regardless of the differences in the actual activity during the job.

Contributors

All authors (JWD, LMG, EF, JN) conceived and designed the study. JWD performed the data analysis together with LMG. JWD, LMG, EF an JN participated in critical interpretation of the results. JWD drafted the paper, LMG, EF and JN contributed to the drafting. All authors contributed to the revision of the manuscript.

Funding

This study was supported by the Dutch Ministry for Public Health, Welfare and Sport via the 'ZonMw' grant no. 522004001.

Ethics statements

Patient consent for publication: not required.
Ethics approval: not required.

Data availability

All data relevant to the study are included in the article or uploaded as supplementary information.

Competing interests

There are no competing interests for any author.

References

1. Ferlay J, Ervik M, Lam F. Global cancer Observatory: cancer today. Lyon, France: International Agency for Research on Cancer, 2018. https://gco.iarc.fr/today
2. IKNL. Overleving darmkanker, 2020. Available: https://www.iknl.nl/kankersoorten/darmkanker/registratie/overleving
3. Siegel RL, Torre LA, Soerjomataram I, et al. Global patterns and trends in colorectal cancer incidence in young adults. Gut 2019;68:2179–85.
4. Vuik FE, Nieuwenburg SA, Bardou M, et al. Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. Gut 2019;68:1820–6.
5. Wei EK, Giovannucci E, Wu K, et al. Comparison of risk factors for colon and rectal cancer. Int J Cancer 2004;108:433–42.
6. Liang PS, Chen T-Y, Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. Int J Cancer 2009;124:2406–15.
7. Lynch BM. Sedentary behavior and cancer: a systematic review of the literature and proposed biological mechanisms. Cancer Epidemiol Biomarkers Prev 2010;19:2691–709.
8. Offermans NSM, Vermeulen R, Burdorf A, et al. Occupational asbestos exposure and risk of esophageal, gastric and colorectal cancer in the prospective Netherlands cohort study. Int J Cancer 2014;135:1970–7.
9. Oddone E, Modonesi C, Gatta G. Occupational exposures and colorectal cancers: a quantitative overview of epidemiological evidence. World J Gastroenterol 2014;20:12431.
10. Talibov M, Sormunen J, Hansen J, et al. Benzene exposure at workplace and risk of colorectal cancer in four Nordic countries. Cancer Epidemiol 2018;55:156–61.
11. van Elsland D, Neefjes J. Bacterial infections and cancer. EMBO Rep 2018;19:e46632.
12. Kikuchi S. Epidemiology of Helicobacter pylori and gastric cancer. Gastric Cancer 2002;5:6–15.
13. Scanu T, Spaapen RM, Bakker JM, et al. Salmonella manipulation of host signaling pathways provokes cellular transformation associated with gallbladder carcinoma. Cell Host Microbe 2015;17:763–74.
14. Mughini-Gras L, Schaapveld M, Kramers J, et al. Increased colon cancer risk after severe Salmonella infection. PLoS One 2018;13:e0189721.
15. Duijster JW, Hansen JV, Franz E, et al. Association between Salmonella infection and colon cancer: a nationwide registry-based cohort study. Epidemiol Infect 2021;149:1–19.
16. Lu R, Wu S, Zhang Y-G, et al. Salmonella protein AvrA activates the STAT3 signaling pathway in colon cancer. Neoplasia 2016;18:307–16.
17. Lu R, Bosland M, Xia Y, et al. Presence of Salmonella AvrA in colorectal tumor and its precursor lesions in mouse intestine and human specimens. Oncotarget 2017;8:55104–15.
18. Pleguezuelos-Manzano C, Puschhof J, Rosendahl Huber A, et al. Mutational signature in colorectal cancer caused by genotoxic pks+ E. coli. Nature 2020;580:269–73.
19. Dziubańska-Kusibab PJ, Berger H, Battistini F, et al. Colibactin DNA-damage signature indicates mutational impact in colorectal cancer. Nat Med 2020;26:1063–9.
20. Kjaerheim K, Haldorsen T, Lynge E, et al. Variation in Nordic work-related cancer risks after adjustment for alcohol and tobacco. Int J Environ Res Public Health 2018;15:2760.
21. Lerro CC, Hofmann JN, Andreotti G, et al. Dicamba use and cancer incidence in the agricultural health study: an updated analysis. Int J Epidemiol 2020;49:1326–37.

22. Pukkala E, Martinsen JI, Lynge E, et al. Occupation and cancer - follow-up of 15 million people in five Nordic countries. Acta Oncol 2009;48:646–790.
23. Papantoniou K, Devore EE, Massa J, et al. Rotating night shift work and colorectal cancer risk in the nurses' health studies. Int J Cancer 2018;143:2709–17.
24. Boulanger M, Morlais F, Bouvier V, et al. Digestive cancers and occupational asbestos exposure: incidence study in a cohort of asbestos plant workers. Occup Environ Med 2015;72:792–7.
25. Beane Freeman LE, Deroos AJ, Koutros S, et al. Poultry and livestock exposure and cancer risk among farmers in the agricultural health study. Cancer Causes Control 2012;23:663–70.
26. Paschke S, Jafarov S, Staib L, et al. Are colon and rectal cancer two different tumor entities? A proposal to abandon the term colorectal cancer. Int J Mol Sci 2018;19:2577.
27. Eurostat. NACE Rev. 2—Statistical classification of economic activities, 2008. Available: http://ec.europa.eu/eurostat/web/nace-rev2
28. Duijster JW, Franz E, Neefjes JJC, et al. Occupational risk of salmonellosis and campylobacteriosis: a nationwide population-based registry study. Occup Environ Med 2019;76:617–24.
29. IKNL. NKR cijfers - dikkedarmkanker, 2020. Available: http://iknl.nl/nkr-cijfers
30. Zha L, Garrett S, Sun J. Salmonella Infection in Chronic Inflammation and Gastrointestinal Cancer. Diseases 2019;7:28.
31. Chow WH, Malker HS, Hsing AW, et al. Occupational risks for colon cancer in Sweden. J Occup Med 1994;36:647–51.
32. Lerro CC, Koutros S, Andreotti G, et al. Cancer incidence in the agricultural health study after 20 years of follow-up. Cancer Causes Control 2019;30:311–22.
33. Marant Micallef C, Shield KD, Baldi I, et al. Occupational exposures and cancer: a review of agents and relative risk estimates. Occup Environ Med 2018;75:604–14.
34. Rastogi SK, Kesavachandran C, Mahdi F, et al. Occupational cancers in leather tanning industries: a short review. Indian J Occup Environ Med 2007;11:3.
35. Sun J, Hobert ME, Rao AS, et al. Bacterial activation of beta-catenin signaling in human epithelia. Am J Physiol Gastrointest Liver Physiol 2004;287:G220–7.
36. de Perio MA, Niemeier RT, Levine SJ, et al. Campylobacter infection in poultry-processing workers, Virginia, USA, 2008-2011. Emerg Infect Dis 2013;19:286–8.
37. Cawthraw SA, Lind L, Kaijser B, et al. Antibodies, directed towards Campylobacter jejuni antigens, in sera from poultry abattoir workers. Clin Exp Immunol 2000;122:55–60.
38. Blair A, Freeman LB. Epidemiologic studies in agricultural populations: observations and future directions. J Agromedicine 2009;14:125–31.
39. Haggar FA, Boushey RP. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. Clin Colon Rectal Surg 2009;22:191–7.
40. McCurdy SA, Sunyer J, Zock J-P, et al. Smoking and occupation from the European community respiratory health survey. Occup Environ Med 2003;60:643–8.
41. Guo Z-L, Wang J-Y, Li Y-S, et al. Association between butchers and cancer mortality and incidence: a systematic review and meta-analysis. Medicine 2017;96:e8177.
42. Wolin KY, Yan Y, Colditz GA, et al. Physical activity and colon cancer prevention: a meta-analysis. Br J Cancer 2009;100:611–6.
43. Gilbert F, Richard J-B, Lapie-Legouis P, et al. Health behaviors: is there any distinction for teachers? A cross-sectional nationwide study. PLoS One 2015;10:e0120040.

44. Statistics Netherlands. Werkenden, 2021. Available: https://www.cbs.nl/nl-nl/visualisaties/dashboard-arbeidsmarkt/werkenden [Accessed 1 Jul 2021].
45. Jongeneel WP, Eysink PED, Theodori D. Work-Related cancer in the European Union: size impact and options for further prevention 2016.
46. Kuntz B, Kroll LE, Hoebel J, et al. [Time trends of occupational differences in smoking behaviour of employed men and women in Germany : Results of the 1999-2013 microcensus]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2018;61:1388–98.
47. Clonan A, Roberts KE, Holdsworth M. Socioeconomic and demographic drivers of red and processed meat consumption: implications for health and environmental sustainability. Proc Nutr Soc 2016;75:367–73.
48. Stoll-Kleemann S, Schmidt UJ. Reducing meat consumption in developed and transition countries to counter climate change and biodiversity loss: a review of influence factors. Reg Environ Change 2017;17:1261–77.
49. Kloosterman R, Akkermans M, Reep C. Vleesconsumptie, 2021. Available: https://www.cbs.nl/nl-nl/longread/rapportages/2021/klimaatverandering-en-energietransitie-opvattingen-en-gedrag-van-nederlanders-in-2020/6-vleesconsumptie [Accessed 1 Jul 2021].

Supplementary material

Table S1. SIRs of colon cancer (overall, proximal and distal) in males per occupational sector as compared with the general Dutch population

	Colon cancer – overall			Proximal colon cancer			Distal colon cancer		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Crop and animal production, hunting	**487**	**569**	**0.86 (0.78-0.94)*** **	**231**	**271**	**0.85 (0.75-0.97)***	**256**	**298**	**0.86 (0.76-0.97)***
Forestry and logging	13	12	1.07 (0.62-1.84)	–	–	–	–	–	–
Fishing and aquaculture	16	14.9	1.07 (0.66-1.75)	–	–	–	–	–	–
Extraction of crude petroleum and natural gas	48	51	0.94 (0.71-1.24)	23	23	0.99 (0.66-1.50)	25	28	0.89 (0.60-1.32)
Other mining and quarrying	29	25	1.16 (0.80-1.67)	15	12	1.30 (0.78-2.15)	14	14	1.04 (0.62-1.75)
Mining support activities	13	19	0.68 (0.39-1.16)	–	–	–	–	–	–
Manufacture of food products	732	768	0.95 (0.89-1.03)	357	355	1.01 (0.91-1.12)	375	413	0.91 (0.82-1.00)
Manufacture of beverages	88	90	0.98 (0.79-1.21)	40	41	0.97 (0.71-1.32)	48	49	0.98 (0.74-1.31)
Manufacture of tobacco products	42	37	1.15 (0.85-1.55)	23	17	1.37 (0.91-2.06)	19	20	0.96 (0.61-1.50)
Manufacture of textiles	132	119	1.11 (0.94-1.32)	56	55	1.02 (0.78-1.32)	76	64	1.19 (0.95-1.49)
Manufacture of wearing apparel	29	29	1.00 (0.70-1.45)	12	14	0.89 (0.51-1.57)	17	15	1.10 (0.69-1.78)
Manufacture of leather (products)	24	24	1.01 (0.67-1.50)	–	–	–	–	–	–
Manufacture of wood/cork/straw products	149	150	1.00 (0.85-1.17)	63	69	0.91 (0.71-1.17)	86	81	1.07 (0.86-1.32)
Manufacture of paper (products)	179	181	0.99 (0.85-1.14)	82	84	0.98 (0.79-1.22)	97	98	0.99 (0.81-1.21)
Printing and reproduction of recorded media	330	315	1.05 (0.94-1.17)	133	145	0.92 (0.77-1.09)	**197**	**170**	**1.16 (1.01-1.33)***
Manufacture of coke and refined petroleum products	45	56	0.81 (0.60-1.08)	18	25	0.71 (0.45-1.13)	27	30	0.89 (0.61-1.30)
Manufacture of chemicals and chemical products	444	450	0.99 (0.90-1.08)	186	206	0.90 (0.78-1.04)	258	244	1.06 (0.94-1.20)
Manufacture of basic pharmaceutical products/preparations	46	44	1.04 (0.78-1.39)	24	20	1.20 (0.81-1.80)	22	24	0.90 (0.59-1.37)

	Colon cancer – overall			Proximal colon cancer			Distal colon cancer		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Manufacture of rubber and plastic products	**286**	**252**	**1.14 (1.01-1.28)***	128	117	1.10 (0.92-1.31)	**158**	**135**	**1.17 (1.00-1.37)***
Manufacture of other non-metallic mineral products	261	271	0.96 (0.85-1.09)	116	125	0.93 (0.77-1.11)	145	146	0.99 (0.84-1.17)
Manufacture of basic metals	242	243	0.99 (0.88-1.13)	112	111	1.01 (0.84-1.21)	130	132	0.99 (0.83-1.17)
Manufacture of fabricated metal products	737	760	0.97 (0.90-1.04)	327	351	0.93 (0.84-1.04)	410	409	1.00 (0.91-1.10)
Manufacture of computer, electronic and optical products	295	320	0.92 (0.82-1.03)	136	146	0.93 (0.79-1.10)	159	174	0.92 (0.78-1.07)
Manufacture of electrical equipment	118	114	1.03 (0.86-1.24)	50	52	0.97 (0.73-1.27)	68	63	1.09 (0.86-1.38)
Manufacture of machinery and equipment	371	342	1.08 (0.98-1.20)	151	154	0.98 (0.83-1.15)	**220**	**188**	**1.17 (1.02-1.34)***
Manufacture of motor vehicles & (semi-) trailers	145	145	1.00 (0.85-1.18)	70	66	1.06 (0.83-1.33)	75	78	0.96 (0.76-1.20)
Manufacture of other transport equipment	87	81	1.07 (0.87-1.33)	34	37	0.93 (0.67-1.30)	53	45	1.19 (0.91-1.56)
Manufacture of furniture	109	110	0.99 (0.82-1.19)	52	51	1.03 (0.78-1.35)	57	60	0.95 (0.74-1.24)
Other manufacturing	1116	1078	1.04 (0.98-1.10)	507	492	1.03 (0.94-1.12)	609	586	1.04 (0.96-1.13)
Repair and installation of machinery and equipment	187	174	1.08 (0.93-1.24)	85	78	1.09 (0.88-1.35)	102	96	1.06 (0.88-1.29)
Electricity, gas, steam and air conditioning supply	356	329	1.08 (0.98-1.20)	155	150	1.03 (0.88-1.21)	201	179	1.12 (0.98-1.29)
Water collection, treatment and supply	66	67	0.98 (0.77-1.25)	20	31	0.65 (0.42-1.01)	46	37	1.26 (0.94-1.68)
Sewerage	23	28	0.83 (0.55-1.25)	11	13	0.87 (0.48-1.57)	12	15	0.79 (0.45-1.40)
Waste collection, treatment and disposal activities; material recovery	164	175	0.94 (0.80-1.09)	72	80	0.91 (0.72-1.14)	92	96	0.96 (0.78-1.18)
Construction of buildings	502	534	0.94 (0.86-1.03)	**196**	**239**	**0.82 (0.71-0.94)****	306	295	1.04 (0.93-1.16)
Civil engineering	511	487	1.05 (0.96-1.14)	224	223	1.00 (0.88-1.14)	287	264	1.09 (0.97-1.22)
Specialized construction activities	1557	1565	0.99 (0.95-1.05)	677	722	0.94 (0.87-1.01)	880	843	1.04 (0.98-1.12)
Wholesale and retail trade and repair of motorcycles/motor vehicles	**888**	**807**	**1.10 (1.03-1.18)****	397	377	1.05 (0.95-1.16)	**491**	**429**	**1.14 (1.05-1.25)****
Wholesale trade	2718	2661	1.02 (0.98-1.06)	1258	1235	1.02 (0.96-1.08)	1460	1426	1.02 (0.97-1.08)

	Colon cancer – overall			Proximal colon cancer			Distal colon cancer		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Retail trade	1340	1407	0.95 (0.90-1.00)	**604**	**670**	**0.90 (0.83-0.98)***	736	737	1.00 (0.93-1.07)
Land transport and transport via pipelines	**2041**	**1925**	**1.06 (1.02-1.11)****	934	888	1.05 (0.99-1.12)	**1107**	**1037**	**1.07 (1.01-1.13)***
Water transport	136	138	0.98 (0.83-1.16)	61	64	0.95 (0.74-1.23)	75	74	1.01 (0.80-1.27)
Air transport	352	351	1.00 (0.90-1.11)	165	163	1.01 (0.87-1.18)	187	189	0.99 (0.86-1.14)
Warehousing and support activities for transportation	534	533	1.00 (0.92-1.09)	250	245	1.02 (0.90-1.15)	284	288	0.99 (0.88-1.11)
Postal and courier services	491	503	0.98 (0.89-1.07)	229	231	0.99 (0.87-1.13)	262	272	0.96 (0.85-1.09)
Accommodation	183	165	1.11 (0.96-1.28)	94	78	1.21 (0.99-1.48)	89	87	1.03 (0.83-1.26)
Food and beverage service activities	502	495	1.01 (0.93-1.11)	254	239	1.06 (0.94-1.20)	248	256	0.97 (0.85-1.10)
Publishing activities	274	280	0.98 (0.87-1.10)	118	130	0.91 (0.76-1.08)	156	149	1.05 (0.89-1.22)
Motion picture, video and television program production, sound recording and music publishing activities	82	80	1.03 (0.83-1.28)	33	38	0.88 (0.62-1.23)	49	42	1.16 (0.88-1.54)
Programming and broadcasting activities	36	32	1.12 (0.81-1.55)	16	15	1.08 (0.66-1.75)	20	17	1.16 (0.75-1.80)
Telecommunications	289	270	1.07 (0.96-1.20)	136	124	1.10 (0.93-1.30)	153	146	1.05 (0.90-1.23)
Computer programming, consultancy and related activities	361	343	1.05 (0.95-1.17)	171	160	1.07 (0.92-1.24)	190	184	1.03 (0.90-1.19)
Information service activities	**38**	**26**	**1.45 (1.05-1.99)***	**23**	**12**	**1.88 (1.25-2.83)****	15	14	1.07 (0.65-1.78)
Financial service activities, except insurance and pension funding	1066	1052	1.01 (0.95-1.08)	468	483	0.97 (0.88-1.06)	598	569	1.05 (0.97-1.14)
(Re)insurance and pension funding	**450**	**403**	**1.12 (1.02-1.23)***	195	185	1.05 (0.92-1.21)	**255**	**217**	**1.17 (1.04-1.33)***
Activities auxiliary to financial services and insurance activities	360	342	1.05 (0.95-1.17)	156	158	0.98 (0.84-1.15)	204	184	1.11 (0.97-1.27)
Real estate activities	**618**	**557**	**1.11 (1.03-1.20)****	278	258	1.08 (0.96-1.21)	**340**	**299**	**1.14 (1.02-1.27)***
Legal and accounting activities	473	436	1.09 (0.99-1.19)	220	203	1.08 (0.95-1.24)	253	233	1.09 (0.96-1.23)
Activities of head offices; management consultancy activities	1289	1286	1.00 (0.95-1.06)	575	590	0.98 (0.90-1.06)	714	697	1.02 (0.95-1.10)
Architectural and engineering activities; technical testing and analysis	730	755	0.97 (0.90-1.04)	380	349	1.09 (0.98-1.20)	**350**	**406**	**0.86 (0.78-0.96)****
Scientific research and development	263	301	0.87 (0.78-0.99)*	121	138	0.88 (0.73-1.05)	142	163	0.87 (0.74-1.03)
Advertising and market research	240	226	1.06 (0.94-1.21)	127	107	1.18 (0.99-1.41)	113	119	0.95 (0.79-1.15)

	Colon cancer – overall			Proximal colon cancer			Distal colon cancer		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Other professional, scientific and technical activities	86	81	1.06 (0.86-1.31)	43	37	1.16 (0.86-1.57)	43	44	0.98 (0.72-1.32)
Rental and leasing activities	181	169	1.01 (0.93-1.24)	73	79	0.92 (0.73-1.16)	108	90	1.20 (0.99-1.45)
Employment activities	**2171**	**2360**	**0.92 (0.88-0.96)*** **	**1044**	**1112**	**0.94 (0.88-1.00)***	**1127**	**1249**	**0.90 (0.85-0.96)*** **
Travel agency, tour operator reservation service	59	69	0.86 (0.67-1.11)	**20**	**32**	**0.63 (0.40-0.97)***	39	37	1.06 (0.78-1.45)
Security and investigation activities	129	138	0.94 (0.79-1.11)	65	64	1.01 (0.80-1.29)	64	74	0.87 (0.68-1.11)
Services to buildings and landscape activities	**439**	**507**	**0.87 (0.79-0.95)****	221	238	0.93 (0.82-1.06)	**218**	**270**	**0.81 (0.71-0.92)****
Office administrative/support and other business support activities	99	107	0.92 (0.76-1.12)	54	50	1.09 (0.83-1.42)	45	58	0.78 (0.78-0.58)
Public administration and defense; compulsory social security	3722	3639	1.02 (0.99-1.06)	1676	1664	1.01 (0.96-1.06)	2046	1975	1.04 (0.99-1.08)
Education	**2415**	**2626**	**0.92 (0.88-0.96)*** **	**1129**	**1202**	**0.94 (0.89-1.00)***	**1286**	**1424**	**0.90 (0.85-0.95)*** **
Human health activities	**913**	**987**	**0.93 (0.87-0.99)***	417	451	0.92 (0.84-1.02)	496	536	0.93 (0.85-1.01)
Residential care activities	524	571	0.92 (0.84-1.00)	246	261	0.94 (0.83-1.07)	278	309	0.90 (0.80-1.01)
Social work activities without accommodation	376	394	0.95 (0.86-1.06)	157	179	0.88 (0.75-1.03)	219	215	1.02 (0.89-1.16)
Creative, arts and entertainment activities	90	93	0.97 (0.79-1.19)	44	43	1.02 (0.76-1.37)	46	50	0.93 (0.69-1.24)
Libraries, archives, museums and other cultural activities	114	108	1.06 (0.88-1.27)	59	49	1.20 (0.93-1.55)	55	58	0.94 (0.72-1.23)
Gambling and betting activities	40	37	1.09 (0.80-1.49)	21	17	1.22 (0.80-1.88)	19	20	0.97 (0.62-1.52)
Sports activities and amusement and recreation activities	277	300	0.92 (0.82-1.04)	139	140	0.99 (0.84-1.17)	138	160	0.86 (0.73-1.02)
Activities of membership organizations	505	483	1.04 (0.96-1.14)	233	225	1.04 (0.91-1.18)	272	258	1.05 (0.94-1.19)
Repair of computers and personal household goods	34	34	0.99 (0.71-1.39)	12	16	0.76 (0.43-1.33)	22	18	1.20 (0.79-1.82)
Other personal service activities	204	204	1.00 (0.87-1.15)	91	97	0.94 (0.77-1.16)	113	107	1.06 (0.88-1.27)
Activities of households as employers of domestic personnel	71	70	1.02 (0.81-1.29)	33	32	1.02 (0.72-1.43)	38	37	1.02 (0.74-1.41)

Obs: observed numbers; Exp: expected numbers; SIR: Standardized Incidence Ratio; CI: confidence interval. * p<0.05; ** p<0.01; *** p<0.001. – Numbers cannot be provided due to risk of subject identification.

Table S2. SIRs of colon cancer (overall, proximal and distal) in females per occupational sector as compared with the general Dutch population

	Colon cancer – overall			Proximal colon cancer			Distal colon cancer		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Crop and animal production, hunting	286	274	1.04 (0.93-1.17)	157	147	1.06 (0.91-1.25)	129	127	1.02 (0.86-1.21)
Manufacture of food products	243	252	0.97 (0.85-1.09)	124	132	0.94 (0.79-1.12)	119	120	0.99 (0.83-1.19)
Manufacture of beverages	10	11	0.94 (0.51-1.75)	–	–	–	–	–	–
Manufacture of textiles	34	44	0.77 (0.55-1.08)	14	23	0.60 (0.36-1.02)	20	21	0.96 (0.62-1.49)
Manufacture of wearing apparel	30	38	0.80 (0.56-1.14)	13	20	0.65 (0.37-1.11)	17	17	0.98 (0.61-1.58)
Manufacture of leather (products)	**17**	**10**	**1.70 (1.05-2.73)***	–	–	–	–	–	–
Manufacture of wood/cork/straw products	15	17	0.86 (0.52-1.43)	–	–	–	–	–	–
Manufacture of paper (products)	28	25	1.12 (0.78-1.63)	13	13	1.01 (0.59-1.74)	15	12	1.25 (0.75-2.07)
Printing and reproduction of recorded media	92	84	1.09 (0.89-1.34)	47	44	1.06 (0.80-1.41)	45	40	1.13 (0.84-1.51)
Manufacture of chemicals and chemical products	50	56	0.90 (0.68-1.18)	27	29	0.94 (0.65-1.38)	23	27	0.85 (0.56-1.28)
Manufacture of basic pharmaceutical products/preparations	18	19	0.93 (0.59-1.48)	–	–	–	–	–	–
Manufacture of rubber and plastic products	48	48	1.00 (0.75-1.33)	23	25	0.92 (0.61-1.38)	25	23	1.09 (0.73-1.61)
Manufacture of other non-metallic mineral products	30	27	1.12 (0.79-1.61)	19	14	1.36 (0.86-2.12)	11	13	0.87 (0.48-1.56)
Manufacture of basic metals	12	13	0.91 (0.52-1.60)	–	–	–	–	–	–
Manufacture of fabricated metal products	77	86	0.90 (0.72-1.12)	39	45	0.87 (0.63-1.19)	38	41	0.93 (0.68-1.28)
Manufacture of computer, electronic and optical products	48	51	0.94 (0.71-1.24)	21	26	0.80 (0.52-1.23)	27	25	1.08 (0.74-1.57)
Manufacture of electrical equipment	18	18	1.02 (0.64-1.62)	–	–	–	–	–	–
Manufacture of machinery and equipment	43	37	1.15 (0.85-1.55)	19	19	1.00 (0.64-1.56)	24	18	1.31 (0.88-1.96)
Manufacture of motor vehicles & (semi-) trailers	13	15	0.89 (0.52-1.53)	–	–	–	–	–	–
Manufacture of furniture	15	20	0.74 (0.45-1.23)	–	–	–	–	–	–
Other manufacturing	282	262	1.08 (0.96-1.21)	154	137	1.13 (0.96-1.32)	128	125	1.02 (0.86-1.21)
Repair and installation of machinery and equipment	18	22	0.81 (0.51-1.28)	–	–	–	–	–	–

	Colon cancer – overall			Proximal colon cancer			Distal colon cancer		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Electricity, gas, steam and air conditioning supply	31	30	1.03 (0.73-1.47)	10	16	0.64 (0.35-1.19)	21	15	1.45 (0.95-2.23)
Waste collection, treatment and disposal activities; material recovery	35	32	1.08 (0.78-1.51)	18	17	1.09 (0.69-1.73)	17	16	1.07 (0.67-1.72)
Construction of buildings	55	55	1.01 (0.77-1.31)	31	28	1.10 (0.78-1.57)	24	27	0.90 (0.60-1.35)
Civil engineering	29	30	0.96 (0.66-1.38)	16	16	1.01 (0.62-1.65)	13	15	0.90 (0.52-1.54)
Specialized construction activities	168	164	1.03 (0.88-1.19)	84	86	0.98 (0.79-.1.21)	84	789	1.08 (0.87-1.34)
Wholesale and retail trade and repair of motorcycles/motor vehicles	128	132	0.97 (0.81-1.15)	58	70	0.83 (0.64-1.08)	70	63	1.12 (0.89-1.41)
Wholesale trade	841	837	1.00 (0.94-1.07)	423	438	0.97 (0.88-1.06)	418	399	1.05 (0.95-1.15)
Retail trade	1838	1920	0.96 (0.91-1.00)	**959**	**1030**	**0.93 (0.87-0.99)***	879	889	0.99 (0.93-1.06)
Land transport and transport via pipelines	310	303	1.02 (0.91-1.14)	155	161	0.97 (0.82-1.13)	155	143	1.08 (0.93-1.27)
Water transport	23	20	1.16 (0.77-1.75)	13	11	1.24 (0.72-2.13)	10	9	1.08 (0.58-2.01)
Air transport	90	100	0.90 (0.73-1.10)	47	52	0.91 (0.68-1.21)	43	48	0.89 (0.66-1.20)
Warehousing and support activities for transportation	88	99	0.89 (0.72-1.09)	46	51	0.90 (0.67-1.20)	42	48	0.87 (0.64-1.18)
Postal and courier services	152	176	0.86 (0.74-1.01)	78	91	0.86 (0.69-1.07)	74	86	0.86 (0.69-1.09)
Accommodation	168	167	1.01 (0.87-1.17)	89	89	1.00 (0.81-1.23)	79	78	1.01 (0.81-1.26)
Food and beverage service activities	**526**	**594**	**0.89 (0.81-0.97)****	**277**	**319**	**0.87 (0.77-0.98)***	249	275	0.91 (0.80-1.03)
Publishing activities	179	172	1.04 (0.90-1.21)	105	92	1.14 (0.94-1.38)	74	80	0.93 (0.74-1.17)
Motion picture, video and television program production, sound recording and music publishing activities	39	39	1.01 (0.73-1.38)	18	21	0.88 (0.55-1.39)	21	18	1.15 (0.75-1.76)
Programming and broadcasting activities	19	20	0.94 (0.60-1.47)	–	–	–	–	–	–
Telecommunications	95	99	0.96 (0.79-1.18)	48	51	0.93 (0.70-1.24)	47	47	0.99 (0.75-1.32)
Computer programming, consultancy and related activities	73	73	1.00 (0.79-1.25)	40	37	1.09 (0.80-1.48)	33	36.4	0.91 (0.64-1.28)
Financial service activities, except insurance and pension funding	402	401	1.00 (0.91-1.10)	187	208	0.90 (0.78-1.04)	215	194	1.11 (0.97-1.27)
(Re)insurance and pension funding	158	167	0.94 (0.81-1.10)	84	86	0.98 (0.79-1.22)	74	82	0.90 (0.72-1.14)

	Colon cancer – overall			Proximal colon cancer			Distal colon cancer		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Activities auxiliary to financial services and insurance activities	190	177	1.07 (0.93-1.24)	95	92	1.03 (0.84-1.26)	95	85	1.12 (0.91-1.37)
Real estate activities	213	241	0.88 (0.77-1.01)	114	128	0.89 (0.74-1.07)	99	114	0.87 (0.72-1.06)
Legal and accounting activities	365	360	1.01 (0.91-1.12)	189	189	1.00 (0.87-1.16)	176	172	1.03 (0.88-1.19)
Activities of head offices; management consultancy activities	426	420	1.01 (0.92-1.12)	215	218	0.98 (0.86-1.13)	211	202	1.05 (0.91-1.20)
Architectural and engineering activities; technical testing and analysis	**124**	**150**	**0.83 (0.69-0.99)***	71	78	0.91 (0.72-1.15)	**53**	**72**	**0.74 (0.56-0.96)***
Scientific research and development	88	79	1.11 (0.90-1.37)	40	41	0.98 (0.72-1.33)	48	38	1.25 (0.94-1.66)
Advertising and market research	188	182	1.03 (0.89-1.19)	103	97	1.06 (0.87-1.28)	85	85	1.00 (0.81-1.24)
Other professional, scientific and technical activities	53	47	1.13 (0.87-1.49)	25	24	1.03 (0.70-1.53)	28	23	1.25 (0.86-1.80)
Veterinary activities	11	15	0.73 (0.41-1.32)	–	–	–	–	–	–
Rental and leasing activities	41	46	0.88 (0.65-1.20)	21	24	0.87 (0.57-1.33)	20	22	0.90 (0.58-1.39)
Employment activities	1328	1390	0.96 (0.91-1.01)	684	725	0.94 (0.87-1.02)	644	665	0.97 (0.90-1.05)
Travel agency, tour operator reservation service	79	80	0.99 (0.80-1.24)	40	42	0.95 (0.70-1.30)	39	38	1.04 (0.76-1.42)
Security and investigation activities	38	41	0.94 (0.68-1.29)	20	21	0.97 (0.63-1.50)	18	20	0.90 (0.57-1.43)
Services to buildings and landscape activities	764	786	0.97 (0.91-1.04)	406	411	0.99 (0.90-1.09)	358	375	0.95 (0.86-1.06)
Office administrative/support and other business support activities	82	87	0.94 (0.76-1.17)	39	45	0.86 (0.63-1.18)	43	42	1.03 (0.76-1.39)
Public administration and defense; compulsory social security	1160	1150	1.01 (0.95-1.07)	635	602	1.06 (0.98-1.14)	525	549	0.96 (0.88-1.04)
Education	1898	1973	0.96 (0.92-1.01)	**955**	**1044**	**0.91 (0.86-0.97)****	943	929	1.02 (0.95-1.08)
Human health activities	**1768**	**1853**	**0.95 (0.91-1.00)***	917	961	0.95 (0.89-1.02)	851	892	0.95 (0.89-1.02)
Residential care activities	2250	2220	1.01 (0.97-1.06)	1125	1154	0.98 (0.92-1.03)	1125	1066	1.06 (1.00-1.12)
Social work activities without accommodation	1676	1725	0.97 (0.93-1.02)	844	896	0.94 (0.88-1.01)	832	830	1.00 (0.94-1.07)
Creative, arts and entertainment activities	53	57	0.93 (0.71-1.22)	30	30	1.00 (0.70-1.44)	23	27	0.86 (0.57-1.29)

	Colon cancer – overall			Proximal colon cancer			Distal colon cancer		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Libraries, archives, museums and other cultural activities	**148**	**126**	**1.18 (1.00-1.38)***	77	67	1.15 (0.92-1.43)	71	58	1.22 (0.96-1.53)
Gambling and betting activities	19	18	1.06 (0.68-1.66)	–	–	–	–	–	–
Sports activities and amusement and recreation activities	172	171	1.01 (0.87-1.17)	77	90	0.86 (0.68-1.07)	95	81	1.18 (0.96-1.44)
Activities of membership organizations	389	396	0.98 (0.89-1.08)	205	210	0.97 (0.85-1.12)	184	186	0.99 (0.86-1.14)
Repair of computers and personal household goods	17	12	1.38 (0.86-2.21)	–	–	–	–	–	–
Other personal service activities	154	180	0.86 (0.73-1.00)	86	95	0.90 (0.73-1.11)	68	84	0.81 (0.64-1.03)
Activities of households as employers of domestic personnel	165	158	1.04 (0.89-1.21)	90	86	1.05 (0.85-1.29)	75	73	1.03 (0.82-1.30)

Obs: observed numbers; Exp: expected numbers; SIR: Standardized Incidence Ratio; CI: confidence interval. * p<0.05; ** p<0.01; *** p<0.001. – Numbers cannot be provided due to risk of subject identification.

Table S3. SIRs of overall colon cancer per age group per sector as compared with the general Dutch population

	<50 years			≥50 years		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Crop and animal production, hunting	103	108	0.95 (0.79-1.16)	**670**	**735**	**0.91 (0.85-0.98)***
Manufacture of food products	147	147	1.00 (0.85-1.18)	828	873	0.95 (0.89-1.02)
Manufacture of beverages	10	11	0.92 (0.49-1.69)	88	90	0.98 (0.80-1.21)
Manufacture of textiles	25	19	1.29 (0.87-1.91)	141	143	0.98 (0.83-1.16)
Manufacture of wood/cork/straw products	23	22	1.06 (0.71-1.60)	141	145	0.97 (0.82-1.14)
Manufacture of paper (products)	30	28	1.09 (0.76-1.55)	177	178	0.99 (0.86-1.15)
Printing and reproduction of recorded media	47	49	0.96 (0.72-1.27)	375	350	1.07 (0.97-1.19)
Manufacture of chemicals and chemical products	53	60	0.88 (0.67-1.15)	441	445	0.99 (0.90-1.09)
Manufacture of basic pharmaceutical products/preparations	12	12	0.98 (0.56-1.72)	52	51	1.01 (0.77-1.33)
Manufacture of rubber and plastic products	57	46	1.24 (0.95-1.60)	277	254	1.09 (0.97-1.23)
Manufacture of other non-metallic mineral products	33	34	0.97 (0.69-1.37)	258	264	0.98 (0.87-1.11)
Manufacture of basic metals	21	24	0.88 (0.57-1.35)	233	233	1.00 (0.88-1.14)
Manufacture of fabricated metal products	116	111	1.04 (0.87-1.25)	698	734	0.95 (0.88-1.02)
Manufacture of computer, electronic and optical products	40	49	0.81 (0.60-1.11)	303	322	0.94 (0.84-1.05)
Manufacture of electrical equipment	20	19	1.07 (0.69-1.65)	116	113	1.03 (0.85-1.23)
Manufacture of machinery and equipment	59	55	1.07 (0.83-1.39)	355	325	1.09 (0.99-1.21)
Manufacture of motor vehicles & (semi-)trailers	18	23	0.80 (0.50-1.27)	140	137	1.02 (0.87-1.21)
Manufacture of other transport equipment	12	12	1.01 (0.57-1.78)	81	78	1.04 (0.83-1.29)
Manufacture of furniture	22	19	1.13 (0.75-1.72)	102	111	0.92 (0.76-1.11)
Other manufacturing	**151**	**114**	**1.33 (1.13-1.56)*****	1247	1227	1.02 (0.96-1.07)
Repair and installation of machinery and equipment	33	30	1.11 (0.79-1.57)	172	166	1.03 (0.89-1.20)
Electricity, gas, steam and air conditioning supply	**37**	**27**	**1.39 (1.01-1.92)***	350	332	1.05 (0.95-1.17)
Waste collection, treatment and disposal activities; material recovery	32	30	1.06 (0.75-1.50)	167	177	0.94 (0.81-1.10)
Construction of buildings	61	64	0.96 (0.74-1.23)	496	525	0.94 (0.86-1.03)
Civil engineering	61	62	0.99 (0.77-1.27)	479	456	1.05 (0.96-1.15)
Specialized construction activities	218	234	0.93 (0.81-1.06)	1507	1495	1.01 (0.96-1.06)

	<50 years			≥50 years		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Wholesale and retail trade and repair of motorcycles/motor vehicles	125	118	1.06 (0.89-1.26)	**891**	**821**	**1.09 (1.02-1.16)***
Wholesale trade	526	538	0.98 (0.90-1.06)	3033	2960	1.02 (0.99-1.06)
Retail trade	629	636	0.99 (0.91-1.07)	**2549**	**2691**	**0.95 (0.91-0.98)****
Land transport and transport via pipelines	212	210	1.01 (0.88-1.16)	**2139**	**2019**	**1.06 (1.02-1.11)****
Water transport	12	16	0.77 (0.44-1.35)	147	142	1.03 (0.88-1.21)
Air transport	62	77	0.81 (0.63-1.03)	380	374	1.02 (0.92-1.12)
Warehousing and support activities for transportation	88	93	0.95 (0.77-1.17)	534	540	0.99 (0.91-1.08)
Postal and courier services	95	89	1.07 (0.88-1.31)	548	590	0.93 (0.85-1.01)
Accommodation	66	64	1.04 (0.81-1.32)	285	268	1.06 (0.95-1.20)
Food and beverage service activities	226	250	0.90 (0.79-1.03)	802	838	0.96 (0.89-1.03)
Publishing activities	55	55	1.00 (0.77-1.31)	398	397	1.00 (0.91-1.11)
Motion picture, video and television program production, sound recording and music publishing activities	18	22	0.81 (0.51-1.29)	103	97	1.07 (0.88-1.29)
Telecommunications	71	66	1.07 (0.85-1.35)	313	302	1.04 (0.93-1.16)
Computer programming, consultancy and related activities	114	117	0.98 (0.81-1.17)	320	300	1.07 (0.96-1.19)
Information service activities	14	9	1.51 (0.90-2.55)	32	26	1.24 (0.88-1.76)
Financial service activities, except insurance and pension funding	200	203	0.99 (0.86-1.13)	1268	1251	1.01 (0.96-1.07)
(Re)insurance and pension funding	85	94	0.90 (0.73-1.12)	**523**	**476**	**1.10 (1.01-1.20)***
Activities auxiliary to financial services and insurance activities	99	86	1.15 (0.95-1.40)	451	433	1.04 (0.95-1.14)
Real estate activities	76	84	0.90 (0.72-1.13)	755	714	1.06 (0.98-1.14)
Legal and accounting activities	148	152	0.97 (0.83-1.14)	690	644	1.07 (0.99-1.15)
Activities of head offices; management consultancy activities	221	241	0.92 (0.80-1.04)	1494	1465	1.02 (0.97-1.07)
Architectural and engineering activities; technical testing and analysis	140	132	1.06 (0.90-1.26)	**714**	**773**	**0.92 (0.86-0.99)***
Scientific research and development	57	53	1.07 (0.82-1.38)	294	327	0.90 (0.80-1.01)
Advertising and market research	76	79	0.96 (0.77-1.21)	352	329	1.07 (0.96-1.19)
Other professional, scientific and technical activities	31	22	1.44 (1.01-2.04)	108	106	1.02 (0.84-1.23)
Rental and leasing activities	28	39	0.72 (0.50-1.05)	194	177	1.10 (0.95-1.26)

	<50 years			≥50 years		
Occupational sector	Obs	Exp	SIR (95%CI)	Obs	Exp	SIR (95%CI)
Employment activities	**717**	**776**	**0.92 (0.86-0.99)***	**2782**	**2974**	**0.94 (0.90-0.97)*****
Travel agency, tour operator reservation service	36	31	1.17 (0.85-1.63)	102	118	0.87 (0.71-1.05)
Security and investigation activities	31	40	0.79 (0.55-1.12)	136	139	0.98 (0.83-1.16)
Services to buildings and landscape activities	230	228	1.01 (0.89-1.15)	973	1065	0.91 (0.86-0.97)
Office administrative/support and other business support activities	40	36	1.12 (0.82-1.53)	141	159	0.89 (0.75-1.05)
Public administration and defense; compulsory social security	496	497	1.00 (0.91-1.09)	4386	4293	1.02 (0.99-1.05)
Education	422	435	0.97 (0.88-1.07)	**3891**	**4164**	**0.93 (0.91-0.96)*****
Human health activities	454	459	0.99 (0.90-1.08)	2227	2381	0.94 (0.90-0.98)
Residential care activities	470	453	1.04 (0.95-1.14)	2304	2338	0.99 (0.95-1.03)
Social work activities without accommodation	317	320	0.99 (0.89-1.11)	1735	1799	0.96 (0.92-1.01)
Creative, arts and entertainment activities	23	26	0.89 (0.59-1.34)	120	124	0.97 (0.81-1.16)
Libraries, archives, museums and other cultural activities	19	23	0.84 (0.54-1.32)	**243**	**211**	**1.15 (1.02-1.31)***
Gambling and betting activities	17	12	1.45 (0.90-2.34)	42	43	0.98 (0.72-1.32)
Sports activities and amusement and recreation activities	79	75	1.06 (0.85-1.32)	370	395	0.94 (0.85-1.04)
Activities of membership organizations	91	99	0.92 (0.75-1.13)	803	781	1.03 (0.96-1.10)
Other personal service activities	60	60	1.01 (0.78-1.29)	298	323	0.92 (0.82-1.03)
Activities of households as employers of domestic personnel	25	23	1.10 (0.75-1.63)	211	205	1.03 (0.90-1.18)

Obs: observed numbers; Exp: expected numbers; SIR: Standardized Incidence Ratio; CI: confidence interval. * p<0.05; ** p<0.01; *** p<0.001.

Chapter 4

REPETITIVE EXPOSURE TO NON-TYPHOIDAL SALMONELLAE IS AN ENVIRONMENTAL RISK FACTOR FOR COLON CANCER

D.M. van Elsland*
J.W. Duijster*
J. Zhang
V. Stévenin
Y.G. Zhang
L. Zha
Y. Xia
E. Franz
J. Sun
L. Mughini-Gras
J.J.C. Neefjes

*equal contributors

Submitted to: *Cell Reports Medicine, 2022*

Abstract

During its infectious cycle, *Salmonella* exploits its host by targeting and manipulating essential signalling pathways. This may disrupt cellular integrity and induce oncogenic transformation. Chronic *S.* Typhi infection can indeed cause gallbladder cancer whereas severe non-typhoidal *Salmonella* (NTS) infection is associated with colon cancer (CC). These severe cases, however, represent only a small fraction of all NTS infections occurring in the population. To assess the overall impact of NTS infections on CC development, we performed a retrospective serological study on NTS exposure in CC patients. We observed significant positive association between the magnitude of exposure to NTS and CC risk. Furthermore, repetitive exposure to low NTS doses were shown to recapitulate the tumorigenic effect of a high NTS exposure *in vivo* and repetitive infection with NTS was shown to induce an increase in malignant transformation in predisposed fibroblast cells.

Statement of significance

We observed a higher NTS seroincidence among prospective CC patients and an oncogenic role of repetitive NTS infections in tissue culture and mouse models. As people acquire numerous NTS infections throughout their life these findings raise the importance to consider NTS as an environmental risk factor for CC.

Introduction

Salmonella enterica subspecies *enterica* is a Gram-negative bacterium including more than 2,500 different serovars that can cause gastrointestinal disease and occasionally invasive infection of variable severity. These serovars are commonly divided into two groups. The typhoidal serovars (i.e., Typhi and Paratyphi) are human restricted pathogens that can cause the severe systemic illnesses; typhoid or paratyphoid fever. The non-typhoidal *Salmonella* (NTS) serovars, of which Enteritidis and Typhimurium are among the most common ones in clinical patients, can colonize asymptomatically a broad range of animals and usually cause gastroenteritis in humans. As *S.* Typhi and Paratyphi are mainly transmitted between humans via the fecal-oral route, the vast majority of (para)typhoid fever cases occur in densely populated areas lacking access to improved sanitation [1]. Conversely, NTS infections occur worldwide, are common in developed countries and are transmitted mostly from animals to humans via food, as well as directly via animal contact or indirectly via the environment [2,3].

Both typhoidal and non-typhoidal serovars have been linked to human cancer. Globally, the incidence of typhoid fever and gallbladder cancer (GBC) show substantial geographical overlap [4, 5]. This link is further supported by histological findings of *S.* Typhi in tumors of GBC patients from geographic areas with high GBC prevalence [6]. Similar to *S.* Typhi, severe NTS infection is epidemiologically associated with increased colon cancer (CC) risk [7]. Indeed, in a large registry-based nationwide cohort study in the Netherlands, the risk of proximal CC was twice as high among people with a laboratory-diagnosed NTS infection as in the general population [7].

During host cell invasion, *Salmonella* injects over 30 different effector proteins into its host to increase its uptake, intracellular survival and egress [8] (LaRock et al., 2015). Among these effectors, AvrA, SopE, SopE2, SopB and SptP are known to mediate activation of the hosts β-catenin, MAPK and AKT signalling pathways, respectively [6, 9-12]. The activation of these pathways by *Salmonella* results in transformation of both *in vitro* and *in vivo* models harbouring pre-transformed genotypes, such as partial (heterozygote) or total (homozygote) deficiency of the tumor suppressor genes *Apc* or *Arf*, respectively, and constitutive expression of the protooncogene c-MYC [6]. *Salmonella* infection thus triggers the activation of oncogenic pathways and as such contributes to one or more steps in the multi-step oncogenic transformation of pre-transformed cells [13, 14].

The severity of a *Salmonella* infection is determined by (a) host factors, (b) the *Salmonella* virulence profile, and (c) the number of Salmonellae ingested [15]. While about 90,000 human salmonellosis cases are reported to public health authorities in Europe each year

[16], this number is based on only those cases needing medical attention, laboratory diagnosis and reporting to public health authorities. It has been estimated that, on average, for every reported salmonellosis case in Europe, approximately 57 *Salmonella* infections go unreported [16]. The reported cases therefore represent mostly severe infections, i.e. a small fraction of all *Salmonella* infections occurring in the population. This has been further supported by serological studies where the rate of the immune response-eliciting exposures to NTS was measured, showing that such exposure vastly exceeds the incidence of clinically overt salmonellosis, with people acquiring numerous mild NTS infections throughout their life [17-19].

As severe or long-lasting *Salmonella* infections may promote colon carcinogenesis by virtue of their higher chance of affecting pre-transformed cells [20], it needs to be understood whether repetitive exposure to NTS also contributes to the multistep CC formation process. To test this, we integrated a serological approach with both *in vivo* and *in vitro* analyses and show that the magnitude of exposure to NTS is epidemiologically associated to CC formation and that *in vivo* exposure to repetitive low doses of NTS contributes to CC in a similar manner as a single high NTS dose. We furthermore report that repetitive NTS infections significantly increase the proliferation of transformed cells in tissue culture experiments. As exposure to NTS is difficult to avoid, these results indicate that *Salmonella* should be considered an environmental risk factor for CC development.

Results

Increased Salmonella seroincidence is associated with increased CC risk

Previously we showed that notified *Salmonella* infections are epidemiologically associated with increased CC risk [7]. However, this study included only reported *Salmonella* infections, which represent a small fraction of the NTS infections that people can acquire throughout life [17-19]. To assess the risk of CC development as a function of the magnitude of NTS exposure, regardless of disease severity, we performed a retrospective matched cohort study on two linked data sets. The first data set was derived from a nationwide cross-sectional serological survey conducted in the Netherlands between October 1995 and December 1996, the so-called 'PIENTER-1' study (De Melker and Conyn-van Spaendonck, 1998). This study established a large serum bank with accompanying epidemiological data representative of the Dutch general population, primarily aimed at immunosurveillance to evaluate the national immunization program. The second data set, obtained from the Netherlands Cancer Registry (NCR), covers all Dutch residents and includes data on patients

diagnosed with CC in the proximal part of the colon since 1998 (ICD-10 codes: C180-C185). We focused our analyses on these colon subsites, as our previous study highlighted a significantly increased risk of cancer after NTS infection only in the proximal but not in the distal part of the colon [7].

By linking the PIENTER-1 study data to CC diagnoses in the NCR data, 36 participants in the PIENTER-1 study were found to have a diagnosed cancer in the proximal colon in the period between January 1st, 1998 and December 31st, 2017 (end of the present study period). Each of the 36 CC patients was then demographically matched with two other PIENTER-1 study participants who were not diagnosed with CC (and were still alive) during the study period (Figure 1A). The characteristics of the total cohort are shown in Supplementary Table 1. The cohort comprised 108 participants (36 CC cases and 72 CC-free individuals), consisting of 42% men and 58% women, with a median age of 63 years (mean 60 years, interquartile range [IQR] 52-68 years) at serum sampling within the PIENTER-1 study.

The serum samples of the 36 CC patients (i.e. 'case') and the 72 persons without CC (i.e. 'controls') were retrieved from the PIENTER-1 serum bank and tested for anti-*Salmonella* IgA, IgM and IgG concentrations using a validated mixed ELISA based on commercially available lipopolysaccharides of the two most common serovars, *S.* Enteritidis (O-antigens 1,9,12) and *S.* Typhimurium (O-antigens 4,5,12). These were used as capture antigens in solid phase and have been extensively validated as a means to determine the rate of infection [21]. For each sample, the concentrations of each Ig isotype were measured and expressed as optical density (OD) units (Supplementary Figure 1). These OD values were then used to estimate the seroincidence of NTS infection, i.e. the average number of NTS infections per person-year, as a measure of NTS infection pressure or force of infection in the person in question. This was done considering the established kinetics of anti-*Salmonella* IgG, IgM, and IgA serum antibody levels following NTS infection [17-19], using an established Bayesian back-calculation model available as an R package called 'seroincidence' [17-19] (Figure 1B). The overall seroincidence was found to be 0.80 (95%CI 0.62-0.98) NTS infections per person-year. When stratified by CC status, the mean seroincidence was 0.94 (95%CI 0.55-1.32) NTS infections per person-year among those who later developed CC, which is higher than the seroincidence of 0.73 (95%CI 0.57-0.89) NTS infections per person-year in the control group. This difference was, however, not statistically significant (HR 1.24, 95%CI 0.82-1.88, p=0.302) (Figure 1C, Overall; Supplementary Table 2).

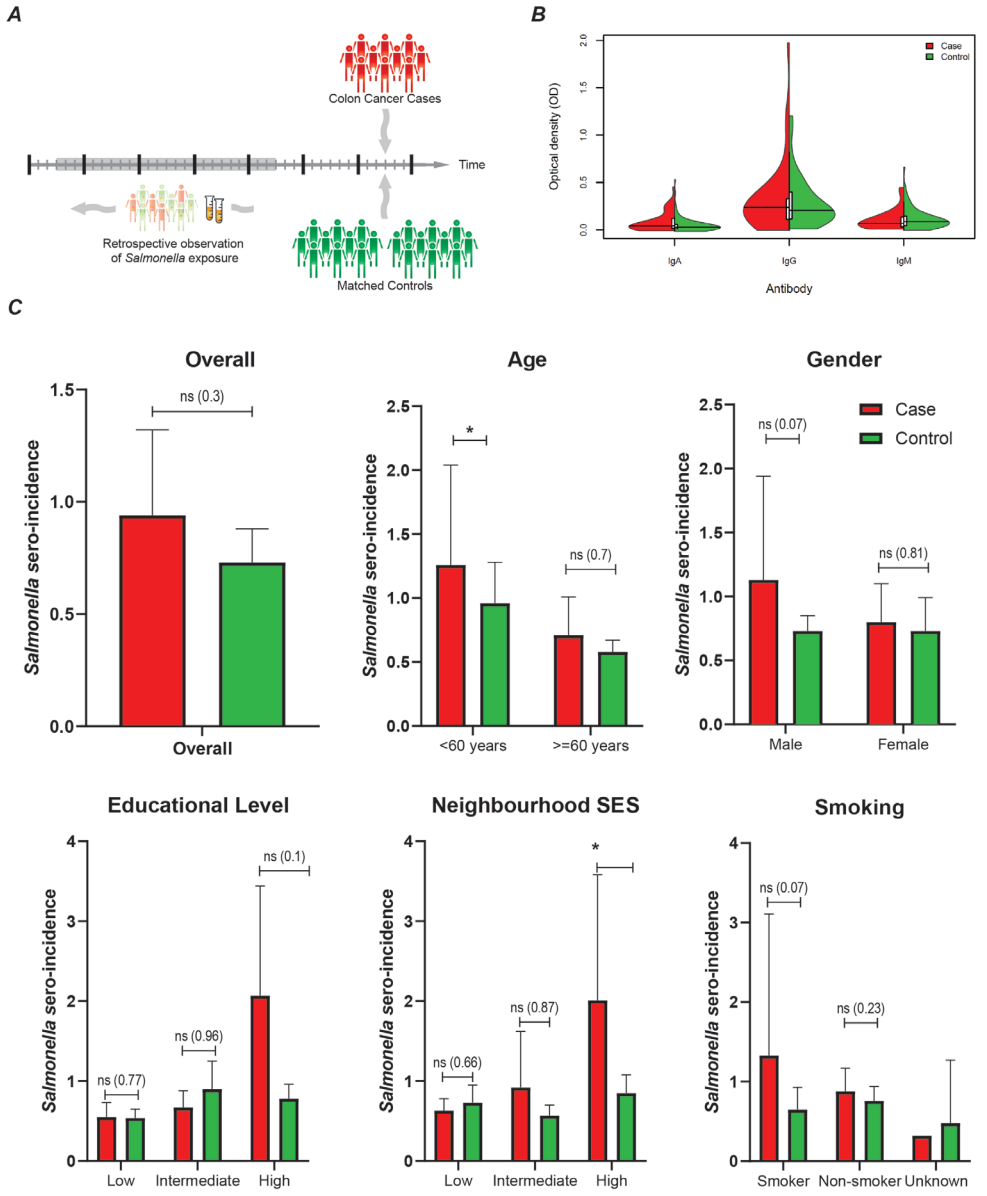


Figure 1. Increased sero-incidence rates among individuals <60 years of age at serum sampling is significantly associated with increased CC risk.

(***A***) Schematic overview of study design: 36 participants in the PIENTER-1 study with a diagnosed cancer in the proximal colon were demographically matched at a 1:2 ratio with other PIENTER-1 study participants who were not diagnosed with CC (i.e. 'controls'). The serum samples were tested for anti-*Salmonella* IgA, IgM and IgG concentrations (***B***) Concentrations of IgA, IgG and IgM anti-*Salmonella* antibodies in cases and controls, expressed in optical density (OD) values. (***C***) *Salmonella*

sero-incidence rates and colon cancer risk stratified by gender, age, ethnicity, educational level, socioeconomic status and smoking. ns= not significant; the value between brackets shows the p-value of the corresponding hazard ratio. * p-value <0.05.

In our previous study, the increased CC risk concerned specifically individuals with age <60 years at reported NTS infection, as CC risk increases dramatically with age due to a multitude other factors that may dilute the relatively smaller contribution of NTS infection [7]. We therefore stratified the present analysis by age at serum sampling and found that increased seroincidence among individuals <60 years of age at serum sampling was significantly associated with increased CC risk (HR 1.41, 95%CI 1.03-1.94, p=0.033) (Figure 1C, Age; Supplementary Table 2). Other factors like gender, educational level or smoking were not significantly associated with increased seroincidence and CC risk (Figure 1*C, Gender, Educational level, Smoking;* Supplementary Table 2). The only other factor modifying significantly the effect of increased seroincidence on CC risk was living in neighborhoods of high socioeconomic status (SES) (HR 1.32, 95%CI 1.03-1.69, p=0.027), suggesting that the link between NTS seroincidence and CC risk can be enhanced by additional environmental settings (Figure 1*C, Neighborhood SES;* Supplementary Table 2). In conclusion, the serological analyses indicated that a high NTS infection pressure (as defined by seroincidence) in age groups in which age itself can be expected to exert lower oncogenic effects than later in life, may act as risk factor for proximal CC.

Impact of repetitive low dose NTS infections on CC formation in mice

To evaluate whether repetitive NTS infections are capable to trigger cell transformation *in vivo*, a mouse study was designed to compare to role in CC formation of repetitive mild infections versus a single severe infection. Since the severity of a NTS infection is determined by the NTS genotype, host factors and the ingested dose [15], we selected the optimal NTS strain for this long-term CC mice-study through an *in vivo* mortality and morbidity screen of several human clinical NTS isolates (Supplementary Figure 2). Higher NTS doses are reported to give higher attack rates and more severe disease [15]. Mild infections were thus mimicked by infecting mice with a low NTS inoculum of 10 bacteria, whereas a severe infection was mimicked by infecting mice with a high inoculum of 10.000 bacteria; a known and well established dose of Salmonellae for mice studies [22, 23].

Mouse experiments were performed using specific pathogen–free female C57BL/6 mice in a carcinogen azoxymethane (AOM)+ inflammatory agent dextran sodium sulphate (DSS) CC model, that has been extensively used as a model system to investigate the accelerating

effect of NTS infections on the multi-step CC formation process [11]. Single high dose exposures ("Single High Sal.") were performed with single subjection to 10.000 bacteria (equivalent CFU) in a 100-µl HBSS suspension. Repetitive low dose exposures ("Multiple Low Sal.") were performed with 3 subjections to 10 bacteria (equivalent CFU) in a 100-µl HBSS suspension. In case of repetitive infections, there was two 4-week delays between exposures (Weeks 1, 4, and 8). As control, non-infected untreated mice were used, as well as non-infected AOM+DSS treated mice (Figure 2*A*).

Throughout the experiment, changes in body weight were monitored for all groups (Figure 2*B*). From week 0 to 16, overall weight increased for all these four groups. The increased rates were, however, markedly different between the untreated control group and the AOM+DSS and AOM+DSS-*Salmonella*-infected treatment groups. However, amongst the three treatment groups, no significant differences were observed. In the first 2 weeks after treatment initiation, the increasing rates of mice's body weight from the three treatment groups all slowed down. Until week 2 or 3, the average weight of all three treatment groups were significantly lower than the control group (Figure 2*B*). As no additional effects on body weight could be observed for any of the NTS exposures, we concluded that the observed body weight changes were solely the result of AOM+DSS treatment

As anticipated from previous studies, AOM+DSS treatment was dominant for tumor formation [11]. Colonic tumors were found to be formed at a similar incidence throughout all treatment groups with no significant differences in case of AOM+DSS-only treatment (tumor incidence: 76.7%; mice with tumor/total mice: 23/30), or in case of both multiple low (71.0%, 22/31) or single high (62.1%, 18/29) *Salmonella* exposures. No tumors were formed in the control group (0/10) (Figure 2*C*). Yet, tumor sizes of the mice in the groups exposed to multiple low ($p<0.001$) or single high ($p<0.01$) doses of NTS were significantly increased compared to the AOM+DSS treated group (Figure 2*C* and 2*D*). NTS infection thus appear to accelerate cancer growth in this model and multiple low doses of NTS trigger a similar tumorigenic impact on CC formation as a singular high dose of NTS. Moreover, the location of the colonic tumors was distributed from distal to proximal colon in both multiple low ($p<0.05$) and single high ($p<0.05$) NTS exposed groups compared with the AOM+DSS treated group where the tumors were restricted to the distal colon (Figure 2*E*) [11].

To further assess this attributing effect, colon tissues were subjected to Hematoxylin and Eosin (H&E) staining and pathological analysis (Figure 2I). Lesions of colon tissues of AOM+DSS treated mice revealed low grade dysplasia in the formed tumors. In comparison, both the multiple low dose and single high dose NTS-exposed mice tissues showed

high grade dysplasia and signs of invasive carcinoma. Control mice did not show any abnormalities (Figure 2*I*). We then evaluated tumor cells and their adjacent tissue growth by BrdU labelling. BrdU labelling was significantly higher in the tumors from both NTS infected groups compared to the AOM+DSS control group ($p<0.01$), with no significant difference between the single high dose and multiple low dose NTS exposed groups (Figure 2*F/G*). In the tissue adjacent to tumor, BrdU positive cells in the colon crypts were significantly higher for all treatment groups compared to the non-treated control group ($p<0.01$). Furthermore, both the low and high dose NTS exposed mice displayed significantly higher BrdU signals in colon crypts than the AOM+DSS control group ($p<0.01$) (Figure 2*F/H*). Similar to the tumor tissue, no significant difference in BrdU intensity between the singular and repetitive NTS exposed mice were observed in the colon crypt tissue (Figure 2*F/H*). These data suggest that both repeated low dose NTS infections or a single high dose NTS infection accelerate proliferation of tumor and tumor adjacent tissue. Colon tissues of mice exposed to both low dose NTS infections and a single high dose NTS infection were found to be colonized by NTS (Figure 2*J*). Tumor tissues were however colonized with significantly more bacteria in comparison to adjacent non-tumor tissues (Figure 2*K*), indicating that in case of both low and high inoculates, NTS preferentially accumulated in tumor-tissues.

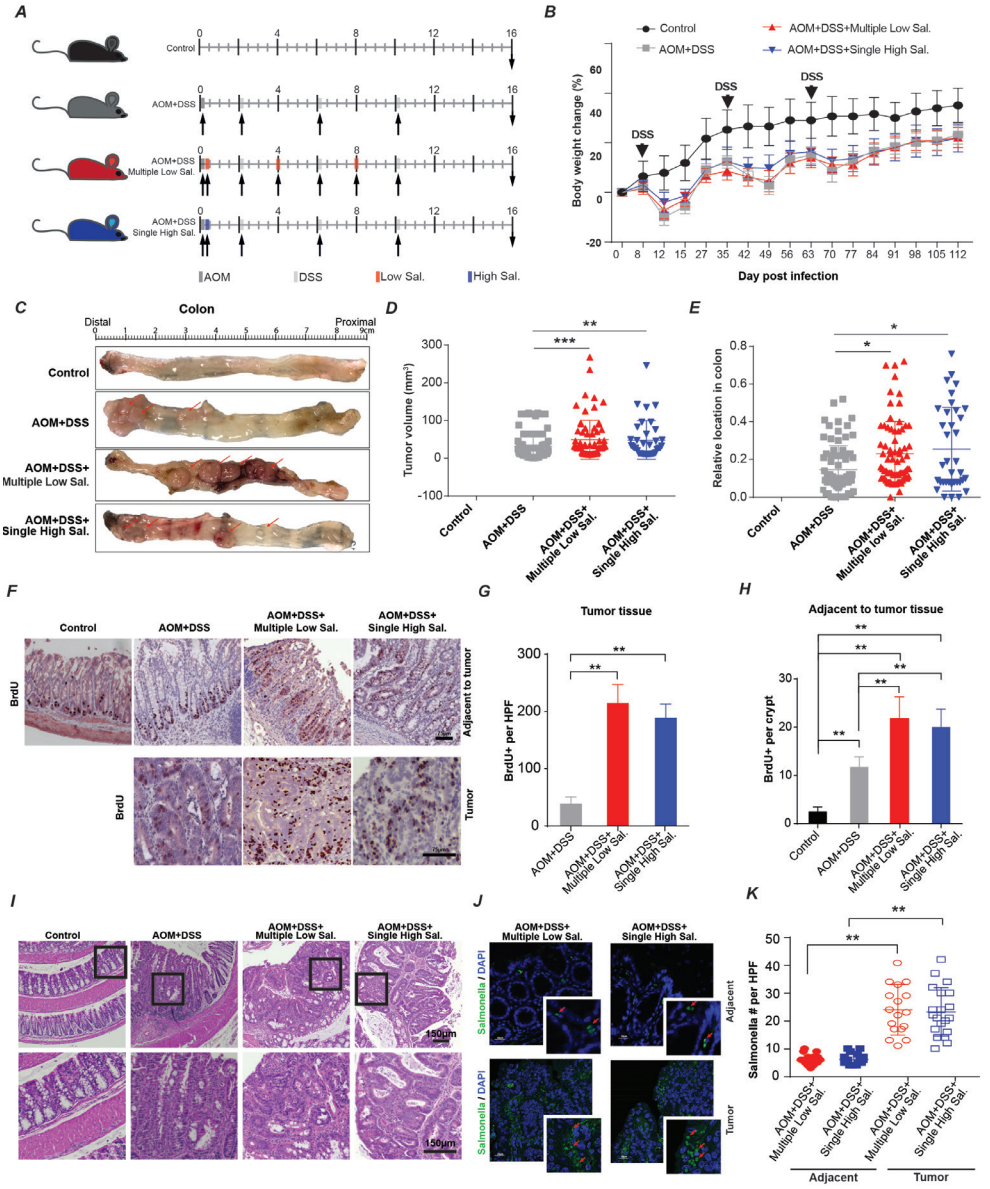


Figure 2. Repetitive Salmonella exposures have tumorigenic impact on colon cancer formation in vivo

(A) Treatment timeline of mice cohort. Mice were infected with either 10 CFU (100-µl suspension in HBSS) for the repetitive low dose of *S.* typhimurium, with 10,000 CFU (100-µl suspension in HBSS) for one single high dose, or treated with sterile HBSS (control and AOM+DSS groups) by oral gavage at day 1. The carcinogen AOM was administrated through intraperitoneal injection at day 2 for all groups, except for the control group. After 7-day recovery period, the inflammatory agent dextran

sodium sulphate (DSS) was administrated at 2% in drinking water for seven days for all groups except for the control. This DSS treatment was repeated at 5 and 9 weeks. In case of repetitive low-dose infections, oral gavage with 10 CFU (100-µl suspension in HBSS) of *S.* typhimurium was repeated at 4 and 8 weeks (Figure 2A). The experiment was evaluated at 16 weeks post infection. n=10, 30, 31, and 29 for control, AOM+DSS, AOM+DSS+Multipe Low Sal. and AOM+DSS+Single High Sal. group, respectively. (***B***) Percent of body weight change throughout the experiment for indicated groups of the mice cohort. Data was expressed as mean ± SD. (***C***) Colonic tumors in situ. Representative colons of indicated groups of the mice cohort at 16 weeks post infection were illustrated. Tumors are indicated by red arrows. (***D***) The tumor volume of indicated groups of within the mice cohort. The data was expressed as mean ±SD; one-way ANOVA, **p<0.01, ***p<0.001. (***E***) The tumor distribution of indicated groups within the mice cohort. The data was expressed as mean ± SD; one-way ANOVA, *p<0.05. (***F***) Immunohistochemistry staining of BrdU in normal colon and colonic tumors of the mice cohort. Scale bars 75 µm. (***G***) Quantification of BrdU staining in tumors from the indicated groups in the mice cohort. The data was expressed as mean ± SD; one-way ANOVA, **p<0.01, n=6 per group. (***H***) Quantification of BrdU staining in normal tissue from the indicated groups in the mice cohort. The data was expressed mean ± SD; one-way ANOVA, **p<0.01, n=6 per group. (***I***) Representative images of normal control colon tissue and colon tumor tissues of indicated groups of the mice cohort. *Control* tissue section from different parts of the colon of control mice. Representative section of a tumor in the colon of *AOM+DSS* mice shows a mucosal lesion with low grade dysplasia showing minor gland distortion and nuclear pseudostratification without significant atypia. Some normal colon tissue is still visible on the edge of the lesion. Representative sections through tumors of both *the single high* and *multiple low Salmonella* exposed mice show high grade dysplasia with mayor gland distortion, cribriform growth and intraluminal cell debris. There is more obvious nuclear atypia with pseudostratification and hyperchromasia. Scale bars 150 µm. (***J***) *Salmonella* invasion in the colon tissue. Localization of *Salmonella* (red arrow) in adjacent normal tissue and colonic tumor tissue was assessed by immunofluorescence staining with *Salmonella*-specific antibody. (***K***) *Salmonella* invasion in the colon tissue. The number of *Salmonella* was counted per High Pure Field (HPF). Data was expressed as mean ± SD; one-way ANOVA, ***p<0.001. n=5-6 per group.

Repetitive NTS infection accelerates growth of pre-transformed cells

We have established a minimal tissue culture model for monitoring *Salmonella*-induced transformation [6]. This model includes Mouse Embryonic Fibroblasts (MEF) engineered to mimic two steps towards transformation: *Arf* deficiency (resulting in TP53 inactivation) and overexpression of c-MYC (named *Arf-/-* + c-MYC). Both TP53 mutations and c-MYC overexpression was also observed in gallbladder carcinoma from patients with a history of *S*. Typhi infection (Scanu et al., 2015). To test whether repeated exposures to NTS increased the rate of transformation, *Arf-/-* + c-MYC MEFs were firstly infected with *S.* Typhimurium (MOI 5 and 25) and seeded in soft agar. As previously reported [6], the acquired capacity of the cells to grow and form colonies anchorage independently, which is an established hallmark of transformed cells, resulted from NTS-induced transformation (Figure 3*A/B; Arf-/-* + c-MYC, MOI 5, MOI 25). As control, non-infected *Arf-/-* + c-MYC MEFs were included that failed to induce colony formation (Figure 3*A/B; Arf-/-* + c-MYC, non-infected). Several colonies of NTS-infected *Arf-/-* + c-MYC MEFs were then isolated from soft-agar and cultured under normal

2D cell culture conditions. Throughout culturing no remaining NTS was observed in these cells, as reported previously [6]. These procedures resulted in the establishment of *Arf-/-* + c-MYC MEFs cell lines with a history of NTS infection (hereafter referred to as transformed MEFs). Following re-seeding of the transformed cells, a subset of cells remained able to form colonies in soft agar, as reported previously [5] (Figure 3*A/B).* To evaluate the effect of repeated NTS exposures on cell transformation, transformed MEFs were re-infected prior to re-seeding in soft-agar, which yielded significantly more colonies (Figure 3*A/B;* transformed, *comparing non-infected with MOI 5 and MOI 25*). This increase was NTS-dose dependent, as an MOI of 25 resulted in significantly more colonies than an MOI of 5 (Figure 3*A/B;* transformed, *comparing MOI 5 to MOI 25*). Remarkably, colonies of the transformed cells were also larger following a re-infection, indicating that these colonies proliferated faster than the reseeded *Arf-/-* + c-MYC MEFs or non-infected transformed MEFs (Figure 3*C*). Increased transformation upon repeated infections was found to be consistent amongst various subsets of the NTS-transformed *Arf-/-* + c-MYC MEFs-isolates (Supplementary Figure 3).

NTS preferably infects (pre-)transformed cells in vitro

As observed in our mouse cohort, tumor tissues where significantly more colonized by NTS in comparison to adjacent non-tumor tissues. In line with these observations it has been reported that NTS preferentially accumulates in tumors when compared to other organs a week after systemic injection [24]. Moreover, specific targeting of host cells by NTS has been reported for particular morphological and microenvironmental features [25]. To evaluate whether NTS specifically targets (pre-)transformed host cells, we compared NTS infection efficiency of MEF cell lines that harbored either one (*Arf-/-*) or two pre-transforming-mutations (*Arf* -/- and c-MYC) to the NTS infection efficiency of transformed MEFs. Intracellular bacterial counts of transformed MEFs were found to be significantly higher than the intracellular bacterial counts of the pre-transformed *Arf-/-* and *Arf-/-* + c-MYC MEFs, indicating that transformed MEFs are more susceptible to NTS invasion than pre-transformed MEFs (Figure 3*D*). Moreover, intracellular bacterial counts of *Arf-/-* + c-MYC MEFs were significantly higher than the intracellular bacterial counts of *Arf-/-* MEFs, also correlating infection efficiency to transformation state (Figure 3*D*). The selectivity of NTS for infecting transformed MEFs was further confirmed in a mixed culture of pre-transformed *Arf-/-* MEFs and transformed MEFs. Increasing the proportion of transformed MEFs within the overall cell population correlated with a similar increase in total intracellular NTS numbers, further illustrating that transformed MEFs are infected by NTS with higher efficiency (Figure 3*E*). Fluorescence microscopy of NTS infected *Arf-/-* c-MYC MEFs and transformed MEFs confirmed higher numbers of

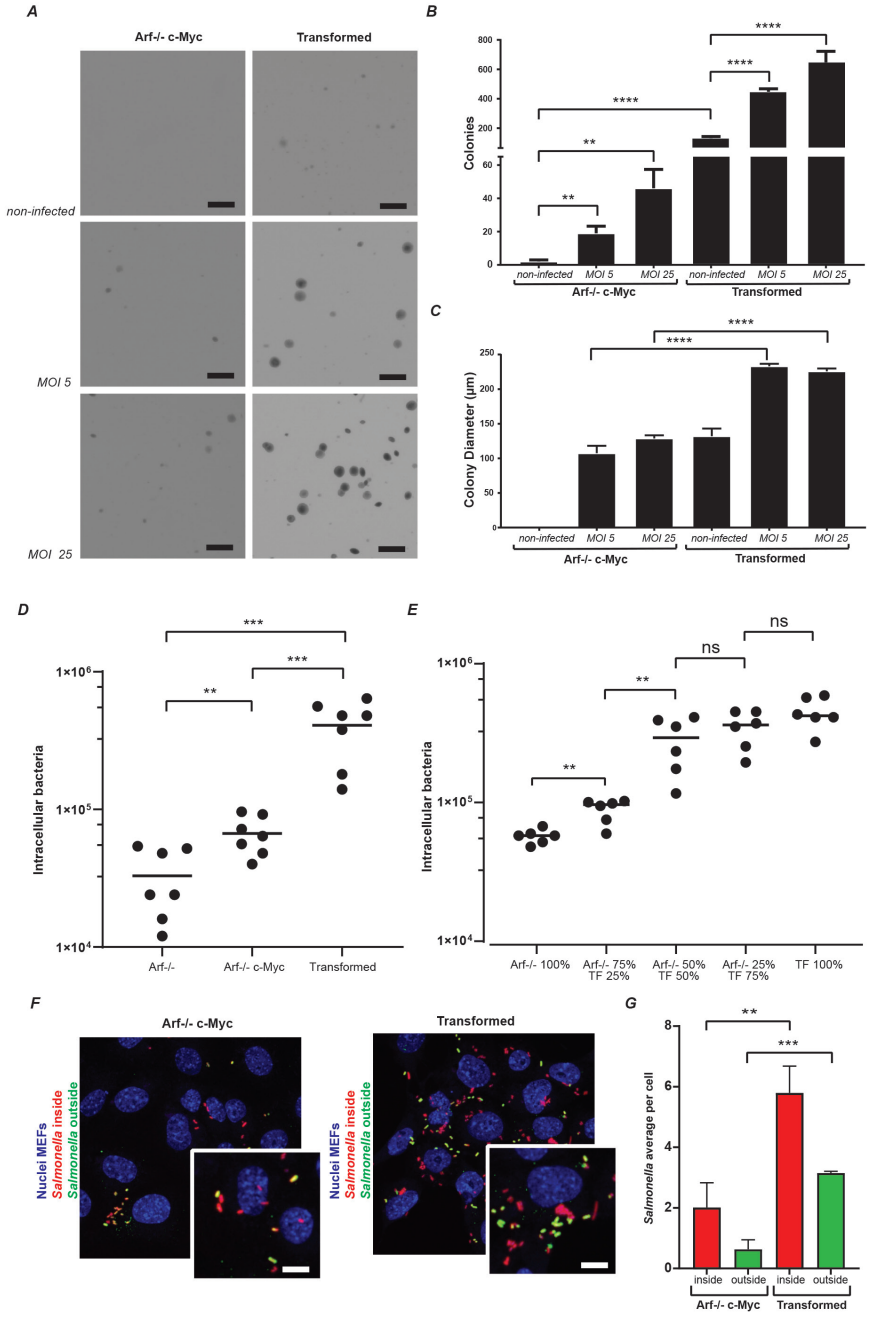


Figure 3: *Salmonella* preferentially infects (pre-)transformed cells, and repetitive *Salmonella* infections increase cellular transformation *in vitro*

(A) Representative images of anchorage-independent growth assay of *Arf*-deficient MEFs overexpressing c-MYC that have not been previously exposed to *Salmonella* (top panels: *Arf-/-* c-MYC), and transformed *Arf*-deficient MEFs overexpressing c-MYC that have been previously exposed to *Salmonella* (bottom panels: *Transformed*). *Arf-/-* c-MYC or transformed MEFs either non-infected, infected with an MOI of 5, or infected with an MOI of 25 are indicated in the left, middle and right panel, respectively. Images represent at least three independent experiments with technical triplicates. Scale bar 750 µm. **(*B*)** Average number of soft-agar colonies per well of *Arf-/-* c-MYC and transformed MEFs overexpressing c-MYC that have been either non-infected, infected with an MOI of 5, or infected with an MOI of 25. Results derive from at least three independent experiments with technical triplicates; one-way ANOVA, **p<0.01, ****p<0.0001. **(*C*)** Average colony diameter of anchorage-independent growth of naïve and transformed *Arf*-deficient MEFs overexpressing c-MYC that have been either non-infected, infected with an MOI of 5, or infected with an MOI of 25; one-way ANOVA, ****p<0.0001 **(*D*)** CFU counts of intracellular bacteria in *Arf-/-*, *Arf-/-* c-MYC and transformed MEFs after infection with *Salmonella* at MOI 25 for 1 hour **(*E*)** CFU counts of intracellular bacteria in mixed populations of *Arf-/-*and transformed MEFs after infection with *Salmonella* at MOI 25 for 1 hour **(*F*)** Representative images of intra-(inside) and extracellular (outside) *Salmonella* bacteria in *Arf-/-* c-MYC and transformed MEFs after infection with *Salmonella* at MOI 25 for 1 hour. Scale bar 10 µm. **(*G*)** Quantification of intra-(inside) and extracellular (outside) *Salmonella* bacteria in *Arf-/-* c-MYC and transformed MEFs after infection with *Salmonella* at MOI 25 for 1 hour; one-way ANOVA, **p<0.01, ***p<0.001.

Salmonellae in transformed MEFs compared to pre-transformed MEFs. To distinguish intracellular NTS from cell surface bound NTS (i.e., not eliminated during washing steps), we used an NTS strain constitutively expressing a dsRed fluorophore and immunolabelled the extracellular salmonellae. Noteworthily, NTS counts at the cell surface of transformed MEFs were also higher in comparison to pre-transformed MEFs, suggesting that the transformed state increased the host cell-adherence of NTS (Figure 3*F/G*). Together, our in vitro data demonstrate a privileged tropism of NTS for cells with the highest level of transformation.

Discussion

Environmental factors are important drivers of CC [26], and colonic microbiota play an important role in both the development and progression of CC [27-29]. In addition to microbial factors that directly affect the genomic integrity of epithelial cells, such as genotoxins [30,31], NTS infections have shown to induce oncogenic transformation of pre-transformed cells upon targeting and manipulation of essential signalling pathways [6]. In long-lasting and severe NTS infections, NTS are more likely to encounter a pre-transformed cell increasing the risk of oncogenic transformation. This has already been suggested by the epidemiological association between severe NTS infections and increased CC risk [7]. It then follows that the sum of multiple NTS infections, which people are known to acquire throughout life [17-19], are similarly conceivable to induce CC.

To test this hypothesis, we assessed the risk of CC development as a function of the magnitude of NTS exposure using a retrospective serological study. This provided unique insights into the frequency of NTS infections and revealed an increased risk of developing CC among people with increased NTS seroincidence before 60 years of age, so before advanced age becomes a dominant risk factor in itself. Indeed, after 60 years of age, CC risk increases considerably due to a multitude of factors that would dilute the relatively smaller effect of NTS infection [7].

The magnitude of exposure to NTS was found to be significantly associated with CC risk among people with a high SES. A higher SES is often associated with a more sedentary occupation (i.e. the so called 'white collar' professions), which is a known risk factor for CC [32]. Hence, it seems that colon carcinogenesis fueled by increased exposure to NTS interacts with other drivers of CC. Specifically, depending on the presence or absence of other risk factors for CC like age and lifestyle (which SES is a proxy for), an effect of NTS infection on CC risk is more evident.

Previously high doses of NTS were shown to contribute to CC formation in pre-transformed mice models [6, 11]. Here we observed that multiple low doses of NTS accelerate *in vivo* tumor formation in a similar manner as a single high dose, thereby indicating that repetitive low dose exposures to NTS triggers tissue and tumor proliferation in a comparable manner as a single high dose NTS exposure. Moreover, after both repetitive low dose infections and a single high dose infection, NTS similarly colonized mice's colon tissues at the end of the study, as deduced by pathology. This could be attributed to our observation that (pre)-transformed cells are more efficiently infected by NTS, whereby a repetitive low inoculum could suffice to target and colonize (pre-)transformed tissues with similar efficiency as a high inoculum. NTS persistence at the end of the *in vivo* study suggests acceptance of the infection without sign of inflammation as deduced from the fact that no significant differences in serum cytokines and chemokines were observed between the groups at that point (Supplementary Figure 4). This is in line with previous studies, in which the NTS infection induces high immune cytokines in the blood of the mice after one week, but then drops to the normal level at 10 weeks post infection [33]. These outcomes furthermore demonstrate that NTS infection did not induce immunomodulatory mechanisms in mice. These mechanism thus do not seem to be involved in the here observed NTS-induced CC formation.

The higher NTS seroincidence among prospective CC patients and the oncogenic role of recurrent low dose NTS infections observed in tissue culture and mouse models identify recurrent low dose NTS infections as a cumulative risk factor for CC development. Low dose

NTS infections can be easily obtained from many sources. Indeed, numerous NTS serovars colonize animals and environmental reservoirs, with *S*. Enteritidis transmission being essentially foodborne, whereas *S*. Typhimurium is virtually ubiquitous [32]. While exposure via food can in principle be prevented, elimination of environmental exposure to NTS is practically impossible. Like sunlight, mild and recurrent NTS infections may represent a hitherto unknown environmental risk factor for CC that cannot be avoided and this may be the case for other cancers and bacterial species as well [34].

Methods

Sero-epidemiological study design

A retrospective matched cohort study was performed based on two linked data sets. The first data set derived from a nationwide cross-sectional serological survey conducted in the Netherlands in October 1995-December 1996, the 'PIENTER-1' study [35] (De Melker and Conyn-van Spaendonck, 1998). The design and rationale of PIENTER-1 are described in detail elsewhere [35] (De Melker and Conyn-van Spaendonck, 1998). In brief, a two-stage cluster sampling design, with 48 municipalities nested in five study-defined regions and age-stratified random sampling applied within these municipalities, was performed. In total, 18,217 people were invited to complete an epidemiological questionnaire and to donate a blood sample. Informed consent was obtained for all participants. Data on the neighbourhood socio-economic status (SES: classified as low, intermediate, and high, based on a standardized index including income, occupation, and education) per postal code area was obtained from Statistics Netherlands (www.cbs.nl). In total, 9948 persons provided a serum sample. The second data set, maintained by the Dutch Association of Comprehensive Cancer Centres (IKNL) (www.iknl.nl), was derived from the Netherlands Cancer Registry (NCR). This registry covers all residents in the Netherlands, the data are more than 95% complete, and includes data on patients diagnosed with CC (ICD-O-3 codes: C180-C189) since 1990. These data also include the colon subsite (proximal, distal) in which the tumour has been found.

Statistics Netherlands (CBS) acted as a trusted third party for data anonymization and linkage by adding a Record Identification Number (RIN) as unique identifier for each individual in the two data sets. Birth date, gender, residence location, and date of registration formed the basis for the derivation of the RIN numbers. To this end, CBS used a reference database containing all mutations due to death or relocation in the Dutch population, including a complete housing history of all Dutch residents. After the RIN numbers were added, all

personal identifiers were removed. Based on RIN numbers, the participants of the PIENTER-1 study were linked to the NCR data on patients with diagnosed CC.

All data sets were cleared from duplicates. CC patients with a date of diagnosis falling after the end of the study period (December 31st, 2017) were censored. As a previous study highlighted a significantly increased risk of cancer only in the proximal part of the colon after reported NTS infection [7], we excluded cases with cancer in the distal part of the colon. After linking the PIENTER-1 records to those of the CC patients in the NCR data set, 36 matches were found, i.e. 36 participants in PIENTER-1 who were diagnosed with cancer in the proximal colon in the period between January 1st, 1998 and December 31st, 2017 (end of the present study period). Each of these 36 CC patients was matched at 1:2 ratio to other PIENTER-1 participants who were not diagnosed with any CC and did not die during the study period. Matching was based on age (±1 year), gender, self-reported educational level (low = primary, lower vocational or lower secondary education; intermediate = intermediate vocational, intermediate secondary or higher secondary education; high = higher vocational and university education), and smoking behavior (smoker, no-smoker, unknown), as reported in the PIENTER-1 study.

Median follow-up time (i.e. time between entry into the cohort and CC diagnosis for the cases or censoring for the matched controls) was 13 years (mean 12 years, IQR 6-16 years), amounting to 1293 person-years at risk in total. The median age at exit from the cohort (i.e. CC diagnosis for the cases or censoring for the matched controls) was 75 years (mean 72 years, IQR 65-80 years). The cohort was mainly composed by persons with a low to intermediate educational level and living predominantly in neighborhoods of low socio-economic status (SES) (Table 1).

Serological analyses and seroincidence calculation

The serum samples of the 36 CC cases and the 72 persons without CC (i.e. 'controls') were retrieved from the PIENTER-1 serum bank and tested for anti-*Salmonella* IgA, IgM, and IgG concentrations using a validated mixed ELISA based on commercially available lipopolysaccharides (SIGMA, Copenhagen) of the two most common serovars, namely Enteritidis (O-antigens 1,9,12) and Typhimurium (O-antigens 4,5,12), as capture antigens in solid phase. A detailed description of this ELISA and its validation has been published previously [21]. For each sample, the concentrations of each Ig isotype were measured separately and expressed as optical density (OD) units. These OD values were then used to estimate, for each sample, the seroincidence of NTS infection, i.e. the average number of NTS infections per person-year as a measure of NTS infection pressure (or force of

infection) in the person in question. This was done using the Bayesian 'back-calculation' model provided for in the R package called 'seroincidence', which has been described in detail elsewhere [17-19] and has been adopted as the standard seroincidence calculator by the European Centre for Disease Control (ECDC) (https://ecdc.europa.eu/en/publications-data/seroincidence-calculator-tool). In brief, the model is based on the kinetics of IgG, IgM, and IgA observed during an 18-month follow-up study with repeated bleeding of 302 Danish adult patients with stool culture-confirmed NTS infections. The model used these data as reference values for peak levels and decay rates of IgG, IgM, and IgA concentrations over time after *Salmonella* infection so that the Ig values measured in a sample can be modelled as a function of time since last seroconversion, taking into account inter-individual variation, thereby estimating an annual seroincidence for any observed set of Ig values in a single sample. This model has been used in several studies on immuno-dynamics of NTS [18, 19] and has been adapted to *Campylobacter* [36- 39] and *Yersinia enterocolitica* [40] as well.

Statistical analysis of sero-epidemiological data

The goal of the analysis was to assess whether NTS seroincidence was a significant predictor of CC. Cox proportional hazards models with attained age as the time scale were used to calculate hazard ratios (HR) and 95% confidence intervals (95%CI) for CC (failure event) as a function of the NTS seroincidence (continuous predictor variable). Follow-up started at cohort entry (i.e. serum sampling at the PIENTER-1 study) and ended at CC diagnosis for both the cases and their matched controls (censoring). As the follow-up time was equal for the members of each matched set, the Breslow method for ties in follow-up time produced HRs that corresponded to risk ratios [41]. A clustered sandwich estimator for variance was used to account for the matched sets, which shown to yield robust estimates of variance for hypothesis testing [42] and generally produce results comparable to frailty models [43].

Stratified analyses were performed according to age at sampling (defined as <60 *vs*. ≥ 60 years, as this was the mean age in our sample and a previous study showed that the potential effect of NTS infection on CC development is unlikely to be observed after that age given the prominent role of other risk factors that may 'dilute' the effect of the infection) [7], as well as gender, neighbourhood SES, educational level, and smoking status, to assess whether there were modifications of the effect of NTS infection pressure on CC risk according to these strata. The two-way interactions between seroincidence and the aforementioned variables were assessed in separate models adjusted for the other variables. However, to avoid collinearity due to the strong association between educational level and neighbourhood SES, only one of these two factors was included as covariate based on the best model fit

as revealed by the Akaike's information criterion. Proportional hazard assumptions were verified using graphical and residual-based methods and found to be met. Statistical analysis was performed using STATA 15 (StataCorp, LP, College Station, Texas, USA).

Animals and ethics statement

Female and male C57BL/6 aged 6-8 weeks mice were obtained from the Jackson Laboratory (Bar Harbor, ME, USA). All animal work was approved by University of Illinois at Chicago Committee on Animal Resources (AAC 18-216). Euthanasia was accomplished via sodium pentobarbital (100 mg per kg body weight) I.P., followed by cervical dislocation. All methods were carried out in accordance with the approved guidelines by the Committee on Animal Resources.

Bacterial strains for animal model and growth condition

Six clinical isolates, including *Salmonella* Typhimurium 1090200009, *Salmonella* Enteritidis 1090301578, *Salmonella* Enteritidis 1091100412, *Salmonella* Typhimurium 1090601671, *Salmonella* Typhimurium 1090404321, and *Salmonella* Enteritidis 1091302626, were used for the morbidity and mortality animal studies. The clinical isolate *Salmonella* Typhimurium 1090404321 was used for the long-term colon cancer mouse model study. Non-agitated microaerophilic bacterial cultures were prepared by inoculating 0.01 mL of a stationary-phase culture to 10mL Buffered Peptone Water (Sigma-Aldrich, St. Louis, MO, USA) followed by overnight incubation (~18 hours) at 37℃.

NTS-infected CC mouse model

Mice experiments were performed using a specific pathogen–free male and female C57BL/6 AOM+DSS CC model, that has been extensively used as a model system to investigate the accelerating effect of NTS infections to the multi-step CC formation process [11]. Animal experiments were performed with 50 male and 50 female C57BL/6 mice aging 6-7 weeks old (The Jackson Laboratory, Bar Harbor, ME, USA) [11,12]. After setting-down for one week in the animal facility, the mice were infected with either a) a single high dose of $1x10^4$ CFU (100-µl suspension in HBSS) *S*. Typhimurium, b) repetitive low doses of $1x10^1$ CFU (100-µl suspension in HBSS) *S*. Typhimurium or c) treated with sterile HBSS (control) by oral gavage, as previously described [7, 11]. After NTS gavage, the carcinogen AOM was administrated through intraperitoneal injection with the dose based on body weight (10 mg/kg) [11]. After a 7-day recovery period, the inflammatory agent dextran sodium sulphate (DSS) was administrated at 2% in drinking water for seven days. This DSS treatment was repeated at 5 and 9 weeks. In the group of repetitive low dose infections, oral gavage with

$1x10^1$ CFU (100-µl suspension in HBSS) of *S*. Typhimurium was repeated at 4 and 8 weeks. Throughout the experiment, mice were weighed and monitored regularly. At 16 weeks post NTS infection, tumors and tissue samples were collected. Tumor counts and measurements were performed in a blinded fashion under a stereo-dissecting microscope (Nikon SMZ1000, Melville, NY, USA). The tumor volume (V) was calculated with caliper measurements using formulas $V= (W^2 \times L)/2$ as described before [44].

Histological testing

Tissues were fixed in 10% neutral buffered formaldehyde for 4-12 hours, then transferred into 70% ethanol and processed by standard techniques. Sections (4µm) were stained with hematoxylin and eosin.

Immunohistochemistry

Tissues were fixed in 10% neutral-buffered formaldehyde for overnight, then transferred into 70% ethanol the next day and processed by standard techniques. Immunohistochemistry staining of target protein was performed on paraffin-embedded sections (4 µm). Briefly, the paraffin sections were baked in an oven at 56 ℃ for 30 minutes. The sides were deparaffinized and rehydrated in xylene, followed by graded ethanol washes at room temperature. Antigen retrieval was achieved by boiling the slides in a microwave oven with 0.01 M, pH 6.0 sodium citrate buffer. Then, the slides were incubated in hydrogen peroxide (3% H_2O_2 in PBS) for 10 minutes, followed by incubation in 5% fetal bovine serum/PBS for 1 hour at room temperature. Purified Anti-BrdU antibody (Abcam, Cambridge, MA, USA) was used in this study [45].

Immunofluorescence and Confocal Imaging

Fresh tumors were fixed in 10% neutral buffered formalin followed by paraffin embedding. For immunofluorescence staining, slides were incubated in 5% bovine serum albumin (BSA) with 0.1% goat serum in PBS for 1 hour at room temperature to reduce nonspecific background. The samples were incubated overnight at 4℃ with primary antibody at 1:100 dilution. The sections were then incubated with secondary antibodies and DAPI for 1 hour at room temperature, and they were examined with confocal microscope as described before [45, 46]. The mouse monoclonal antibody for S. Typhimurium 0-4 (Santa Cruz, Dallas, TX, USA) was used in this study.

Luminex immunoassays

The cytokines and chemokines in the plasma samples from the studied animals were assessed using the ProcartaPlex Mouse Cytokine/Chemokine Convenience Panel 1 26plex

(Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instruction. Briefly, after adding magnetic beads, 25µl of plasma samples were added and followed by detection of antibody and streptavidin-PE provided by the kit. The plate was read on a MAGPIXTM system platform (Millipore Sigma, Burlington, MA, USA) after adding reading buffer.

Statistical analysis of mice experiments

For the mouse model related experiments, data were expressed as mean ± SD. One-way ANOVA was performed to the statistical analysis in the animal studies. All statistical tests were two-sided, and p-values <0.05 were considered statistically significant. The statistical analyses of experimental data were performed with GraphPad Prism 5.

Bacterial strains and cell lines for in vitro experiments

S. Typhimurium strain SL1344 was a courtesy of S. Méresse [47]. Mouse Embryonic Fibroblasts (MEFs) were derived from Arf-deficient C57BL/6 mice. MEFs overexpressing c-MYC were generated by retroviral transduction using a pLZRS-GFP(ires)-HA backbone. MEFs were cultured at 37°C, 5% CO_2 in Dulbecco's Modified Eagle's Medium (DMEM) (Invitrogen) [6].

In vitro NTS infection, CFU, microscopy and anchorage-independent growth assays

NTS infection of MEFs cells was performed as described previously [9]. In brief, *S.* Typhimurium strain SL1344 was grown overnight at 37°C in LB medium, supplemented with 100 µg/mL ampicillin throughout the bacterial culturing. The next day, the bacteria were sub-cultured at a dilution of 1:33 in fresh LB medium and incubated for 2 hours at 37°C while shaking. Cells were infected with NTS at the indicated MOI in DMEM medium without antibiotics for 20 minutes at 37°C, 5% CO_2 in a tissue culture chamber and then incubated in the presence of 100 µg/mL gentamicin (GIBCO) for 1 hour to eliminate extracellular bacteria. In case of CFU or microscopy experiments cells were then lysed and plated on LB plates or fixed with 4% PFA for 10 min at room temperature, respectively. In case of anchorage-independent growth assays MEFs were cultured for another 2 hours in the presence of 10 µg/mL gentamicin. The infected MEFs were subsequently collected and resuspended in DMEM medium supplemented with 10 µg/mL gentamicin and 0.35% low melting point agarose (UltraPure™, Invitrogen) and were poured on a soft agar bottom layer consisting of 0.7% low melting point agarose in DMEM with 10 µg/mL gentamicin. Anchorage-independent cell growth and number of soft agar colonies were assessed after 1-3 weeks of incubation at 37°C, 5% CO_2 using GelCountTM (Oxford Optronix, UK). For microscopy analysis fixed slides were stained with rabbit polyclonal anti-*S.*Typhimurium LPS (Difco, Detroit, MI) and

DAPI (Life Technologies). Images were acquired using a Leica TCS SP8 (Leica Microsystems, Wetzlar, Germany) at 40x or 63x magnification.

Acknowledgements

This work was supported by KWF grant 'Bacterial food poisoning and colon cancer; a cell biological and epidemiological study' (2017-1-11001), an ERC Advanced grant ERCOPE (to JN) and the Netherlands Organisation for Health Research and Development (ZonMw) (grant number 522004001). Mouse experiments have been supported by the Oncode program from Oncode member JN and UIC Cancer Center Fund for JS. VS was supported by the LUMC MSCA-IF Seal of Excellence program. We would like to thank Rong Lu and Shari Garrett for helping with the animal studies, and Karen A. Krogfelt for help with the epidemiological study.

References

1. Mogasale V, Maskery B, Ochiai RL, Lee JS, Mogasale VV, Ramani E, et al. Burden of typhoid fever in low-income and middle-income countries: a systematic, literature-based update with risk-factor adjustment. The Lancet Global Health. 2014;*2*(10):e570-80.
2. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The global burden of nontyphoidal Salmonella gastroenteritis. Clinical infectious Diseases. 2010;50(6):882-889.
3. Knodler LA, Elfenbein JR. Salmonella enterica. Trends in Microbiology, 2019;27(11):964-5.
4. Lazcano-Ponce EC, Miquel JF, Muñoz N, Herrero R, Ferrecio C, Wistuba II, et al. Epidemiology and molecular pathology of gallbladder cancer. CA: a Cancer Journal for Clinicians. 2001;51(6):349-64.
5. Wistuba II, Gazdar AF. Gallbladder cancer: lessons from a rare tumour. Nature Reviews Cancer. 2004;4(9):695-706.
6. Scanu T, Spaapen RM, Bakker JM, Pratap CB, Wu LE, Hofland I, et al. *Salmonella* manipulation of host signaling pathways provokes cellular transformation associated with gallbladder carcinoma. Cell Host & Microbe. 2015;17(6):763-74.
7. Mughini-Gras L, Schaapveld M, Kramers J, Mooij S, Neefjes-Borst EA, Pelt WV, et al. Increased colon cancer risk after severe *Salmonella* infection. PloS One. 2018;13(1):e0189721.
8. LaRock DL, Chaudhary A, Miller SI. Salmonellae interactions with host processes. Nature Reviews Microbiology. 2015;13(4):191-205.
9. Kuijl C, Savage ND, Marsman M, Tuin AW, Janssen L, Egan DA, et al. Intracellular bacterial growth is controlled by a kinase network around PKB/AKT1. Nature. 2007;450(7170):725-30.
10. Lu R, Liu X, Wu S, Xia Y, Zhang YG, Petrof EO, et al. Consistent activation of the β-catenin pathway by *Salmonella* type-three secretion effector protein AvrA in chronically infected intestine. American Journal of Physiology-Gastrointestinal and Liver Physiology. 2012;303(10):G1113-25.
11. Lu R, Wu S, Zhang YG, Xia Y, Liu X, Zheng Y, et al. Enteric bacterial protein AvrA promotes colonic tumorigenesis and activates colonic beta-catenin signaling pathway. Oncogenesis. 2014;3(6):e105-e105.
12. Lu R, Bosland M, Xia Y, Zhang YG, Kato I, Sun J. Presence of *Salmonella* AvrA in colorectal tumor and its precursor lesions in mouse intestine and human specimens. Oncotarget. 2017;8(33):55104.
13. Gagnaire A, Nadel B, Raoult D, Neefjes J, Gorvel JP. Collateral damage: insights into bacterial mechanisms that predispose host cells to cancer. Nature Reviews Microbiology. 2017;15(2):109-28.
14. Van Elsland D, Neefjes J. Bacterial infections and cancer. EMBO reports. 2018;19(11):e46632.
15. Glynn JR, Bradley DJ. The relationship between infecting dose and severity of disease in reported outbreaks of Salmonella infections. Epidemiology & Infection. 1992;109(3):371-88.
16. Havelaar AH, Ivarsson S, Löfdahl M, Nauta MJ. Estimating the true incidence of campylobacteriosis and salmonellosis in the European Union, 2009. Epidemiology & Infection 2013;141(2):293-302.
17. Simonsen J, Strid MA, Mølbak K, Krogfelt KA, Linneberg A, Teunis P. (2008). Sero-epidemiology as a tool to study the incidence of *Salmonella* infections in humans. Epidemiology & Infection. 2008;136(7):895-902.
18. Simonsen J, Mølbak K, Falkenhorst G, Krogfelt KA, Linneberg A, Teunis PFM. Estimation of incidences of infectious diseases based on antibody measurements. Statistics in Medicine. 2009;28(14):1882-95.

19. Simonsen J, Teunis P, Van Pelt W, Van Duynhoven Y, Krogfelt KA, Sadkowska-Todys M, et al. Usefulness of seroconversion rates for comparing infection pressures between countries. Epidemiology & Infection. 2011;139(4);636-43.
20. Zha L, Garrett S, Sun J. *Salmonella* infection in chronic inflammation and gastrointestinal cancer. Diseases. 2019;7(1):28.
21. Strid MA, Dalby T, Mølbak K, Krogfelt KA. Kinetics of the human antibody response against *Salmonella* enterica Serovars Enteritidis and Typhimurium determined by lipopolysaccharide enzyme-linked immunosorbent assay. Clinical and Vaccine Immunology. 2007;14(6):741-7.
22. Wu S, Lu R, Zhang YG, Sun J. Chronic Salmonella infected mouse model. JoVE (Journal of Visualized Experiments). 2010;(39):e1947.
23. Nilsson OR, Kari L, Steele-Mortimer O. Foodborne infection of mice with *Salmonella* Typhimurium. PloS One. 2019;14(8):e0215190.
24. Forbes NS, Munn LL, Fukumura D, Jain RK. Sparse initial entrapment of systemically injected *Salmonella* typhimurium leads to heterogeneous accumulation within tumors. Cancer Research. 2003;63(17):5188-93.
25. Voznica J, Gardella C, Belotserkovsky I, Dufour A, Enninga J, Stévenin V. Identification of parameters of host cell vulnerability during *Salmonella* infection by quantitative image analysis and modeling. Infection and Immunity. 2017;86(1):e00644-17.
26. Lee-Six H, Olafsson S, Ellis P, Osborne RJ, Sanders MA, Moore L, et al. The landscape of somatic mutation in normal colorectal epithelial cells. Nature. 2019;574(7779):532-7.
27. Wong SH, Zhao L, Zhang X, Nakatsu G, Han J, Xu W, et al. Gavage of fecal samples from patients with colorectal cancer promotes intestinal carcinogenesis in germ-free and conventional mice. Gastroenterology. 2017;153(6):1621-33.
28. Thomas AM, Manghi P, Asnicar F, Pasolli E, Armanini F, Zolfo M, et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. Nature Medicine. 2019;25(4):667-78.
29. Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. Nature Medicine. 2019;25(4):679-89.
30. Dziubańska-Kusibab PJ, Berger H, Battistini F, Bouwman BA, Iftekhar A, Katainen R, et al. Colibactin DNA-damage signature indicates mutational impact in colorectal cancer. Nature Medicine. 2020;26(7):1063-9.
31. Iftekhar A, Berger H, Bouznad N, Heuberger J, Boccellato F, Dobrindt U, et al. Genomic aberrations after short-term exposure to colibactin-producing *E. coli* transform primary colon epithelial cells. Nature Communications. 2021;12(1):1-15.
32. Mughini-Gras L, Enserink R, Friesema I, Heck M, van Duynhoven Y, van Pelt W. Risk factors for human salmonellosis originating from pigs, cattle, broiler chickens and egg laying hens: a combined case-control and source attribution analysis. PloS One. 2014;9(2):e87933.
33. Lu R, Wu S, Liu X, Xia Y, Zhang YG, Sun J. Chronic effects of a Salmonella type III secretion effector protein AvrA in vivo. PloS One. 2010;5(5):e10505.
34. Moody S, Senkin S, Islam SM, Wang J, Nasrollahzadeh D, Cortez Cardoso Penha R, et al. Mutational signatures in esophageal squamous cell carcinoma from eight countries with varying incidence. Nature Genetics. 2021;53(11):1553-63.

35. De Melker HE, Conyn-van Spaendonck MAE. Immunosurveillance and the evaluation of national immunization programmes: a population-based approach. Epidemiology & Infection. 1998;121(3):637-43.
36. Ang CW, Teunis PFM, Herbrink P, Keijser J, Van Duynhoven YHTP, Visser CE, Van Pelt W. Seroepidemiological studies indicate frequent and repeated exposure to Campylobacter spp. during childhood. Epidemiology & Infection. 2011;139(9):1361-8.
37. Teunis PFM, Van Eijkeren JCH, Ang CW, van Duynhoven YTHP, Simonsen JB, Strid MA, et al. Biomarker dynamics: estimating infection rates from serological data. Statistics in Medicine 2012;31(20):2240-8.
38. Teunis PFM, Falkenhorst G, Ang CW, Strid MA, De Valk H, Sadkowska-Todys M, et al. Campylobacter seroconversion rates in selected countries in the European Union. Epidemiology & Infection. 2013;141(10):2051-7.
39. Monge S, Teunis P, Friesema I, Franz E, Ang W, van Pelt W, Mughini-Gras L. Immune response-eliciting exposure to Campylobacter vastly exceeds the incidence of clinically overt campylobacteriosis but is associated with similar risk factors: A nationwide serosurvey in the Netherlands. Journal of Infection. 2018;77(3):171-7.
40. Vaitkeviciute I, Teunis PFM, van Pelt W, Krogfelt KA. Kinetics of serum antibodies in response to infection with Yersinia enterocolitica. Epidemiology & Infection. 2019:147.
41. Cummings P, McKnight B, Weiss NS. Matched-pair cohort methods in traffic crash research. *Accident Analysis & Prevention*. 2003;*35*(1):131-41.
42. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. Journal of the American Statistical Association. 1989;84(408):1074-8.
43. Shaffer ML, Hiriote S. Analysis of time-to-event and duration outcomes in neonatal clinical trials with twin births. Contemporary Clinical Trials. 2009;30(2):150-4.
44. Faustino-Rocha A, Oliveira PA, Pinho-Oliveira J, Teixeira-Guedes C, Soares-Maia R, Da Costa RG, et al. Estimation of rat mammary tumor volume using caliper and ultrasonography measurements. Lab Animal. 2013;42(6):217-24.
45. Zhang J, Lu R, Zhang Y, Matuszek Ż, Zhang W, Xia Y, et al. tRNA queuosine modification enzyme modulates the growth and microbiome recruitment to breast tumors. Cancers. 2020;12(3):628.
46. Lu R, Zhang YG, Xia Y, Zhang J, Kaser A, Blumberg R, Sun J. Paneth cell alertness to pathogens maintained by vitamin D receptors. Gastroenterology. 2021;160(4):1269-83.
47. Méresse S, Steele-Mortimer O, Finlay BB, Gorvel JP. The rab7 GTPase controls the maturation of Salmonella typhimurium-containing vacuoles in HeLa cells. The EMBO journal, 1999;18(16):4394-4403.

Supplementary tables

Supplementary Table 1. Baseline characteristics of the study cohorts

	Individuals who developed colon cancer during the follow-up period	Individuals who did not develop colon cancer during the follow-up period*
Gender		
Male	15 (42%)	30 (42%)
Female	21 (58%)	42 (58%)
Age at entry		
<60 years	15 (42%)	30 (42%)
≥60 years	21 (58%)	42 (58%)
Educational level at entry**		
Low	14 (39%)	28 (39%)
Intermediate	14 (39%)	28 (39%)
High	8 (22%)	16 (22%)
Neighbourhood SES at entry***		
Low	24 (67%)	49 (68%)
Intermediate	5 (14%)	10 (14%)
High	7 (19%)	13 (18%)
Smoking at entry		
Smoker	6 (17%)	12 (17%)
Non-smoker	29 (80%)	58 (80%)
Unknown	1 (3%)	2 (3%)
Follow-up time		
<5 years	6 (16%)	12 (16%)
5-15 years	15 (42%)	30 (42%)
>5 years	15 (42%)	30 (42%)

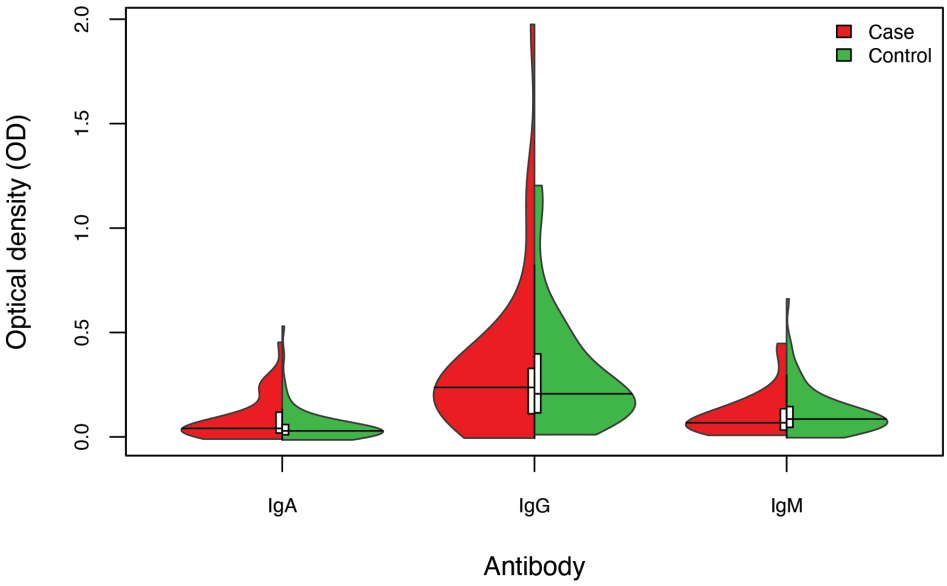
*Matched to the colon cancer patients at a 1:2 ratio based on gender, age at entry (±1 year), educational level and smoking status. **Low = primary, lower vocational or lower secondary education; intermediate = intermediate vocational intermediate secondary or higher secondary education; high = higher vocational and university education. ***Socio-economic status, classified as low, intermediate and high based on a standard index including income, occupation and education per postal code area ('neighbourhood') obtained from Statistics Netherlands (www.cbs.nl).

Supplementary Table 2. *Salmonella* sero-incidence rates and colon cancer risk by gender, age, ethnicity, educational level, socioeconomic status and smoking.

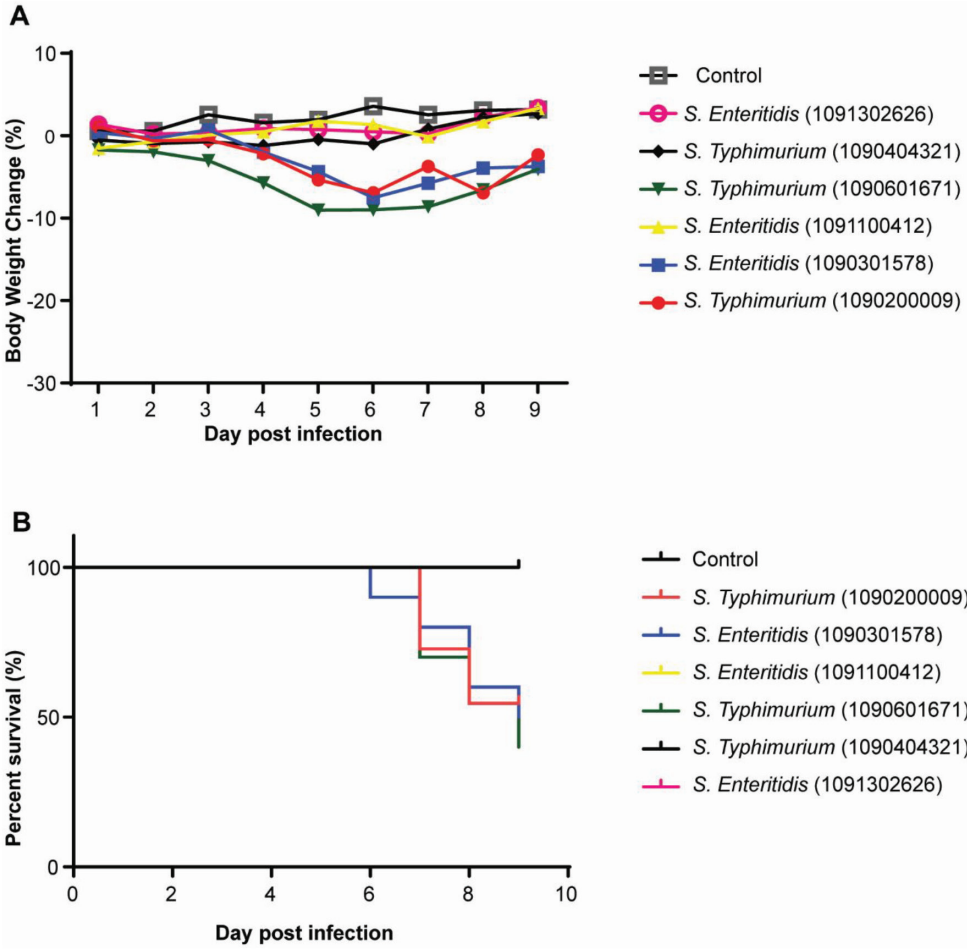
	Mean (95%CI) *Salmonella* sero-incidence among colon cancer cases	**Mean (95%CI) *Salmonella* sero-incidence among controls**	**Person-years at risk**	**Hazard ratio (95%CI)**	**P-value**
Overall	0.94 (0.55-1.32)	0.73 (0.57-0.88)	1293	1.24 (0.82-1.88)[§]	0.302
Age at entry					
<60 years	**1.26 (0.48-2.04)**	**0.96 (0.64-1.28)**	**617**	**1.41 (1.03-1.94)[‡]**	**0.033**
≥60 years	0.71 (0.40-1.01)	0.58 (0.48-0.67)	676	0.77 (0.29-2.01)[‡]	0.704
Gender					
Male	1.13 (0.32-1.94)	0.73 (0.61-0.85)	581	1.46 (0.97-2.22)[§]	0.072
Female	0.80 (0.50-1.10)	0.73 (0.47-0.99)	712	1.06 (0.68-1.64)[§]	0.810
Educational level at entry*					
Low	0.55 (0.38-0.73)	0.54 (0.42-0.65)	424	0.83 (0.23-2.94)[§]	0.771
Intermediate	0.67 (0.47-0.88)	0.90 (0.55-1.25)	564	1.01 (0.59-1.73)[§]	0.965
High	2.07 (0.71-3.44)	0.78 (0.60-0.96)	305	1.26 (0.96-1.66)[§]	0.102
Neighbourhood SES at entry**					
Low	0.63 (0.47-0.78)	0.73 (0.51-0.95)	827	0.87 (0.47-1.60)[‡]	0.656
Intermediate	0.92 (0.22-1.62)	0.57 (0.44-0.70)	202	1.03 (0.71-1.49)[‡]	0.866
High	**2.01 (0.44-3.58)**	**0.85 (0.65-1.08)**	**264**	**1.32 (1.03-1.69)[‡]**	**0.027**
Smoking at entry					
Smoker	1.33 (0.00-3.11)	0.65 (0.36-0.93)	242	1.46 (0.97-2.19)[‡]	0.068
Non-smoker	0.88 (0.58-1.17)	0.76 (0.57-0.94)	988	1.06 (0.73-1.52)[‡]	0.234
Unknown	0.32 (0.32-0.32)	0.48 (0.00-1.27)	63	-	-

*Low=primary, lower vocational or lower secondary education; intermediate=intermediate vocational intermediate secondary or higher secondary education; high=higher vocational and university education. **Socio-economic status, classified as low, intermediate and high based on a standard index including income, occupation and education per postal code area ('neighbourhood') obtained from Statistics Netherlands (www.cbs.nl). §Adjusted for all other variables in the table, except for neighbourhood SES, as it was collinear with the educational level and the inclusion of educational level in the model resulted in a better model fit (lower AIC) as compared to including neighbourhood SES. †Adjusted for all other variables in the table, except for educational level, as it was collinear with the neighbourhood SES and the inclusion of neighbourhood SES in the model resulted in a better model fit (lower AIC) as compared to including for educational level.

Supplementary figures

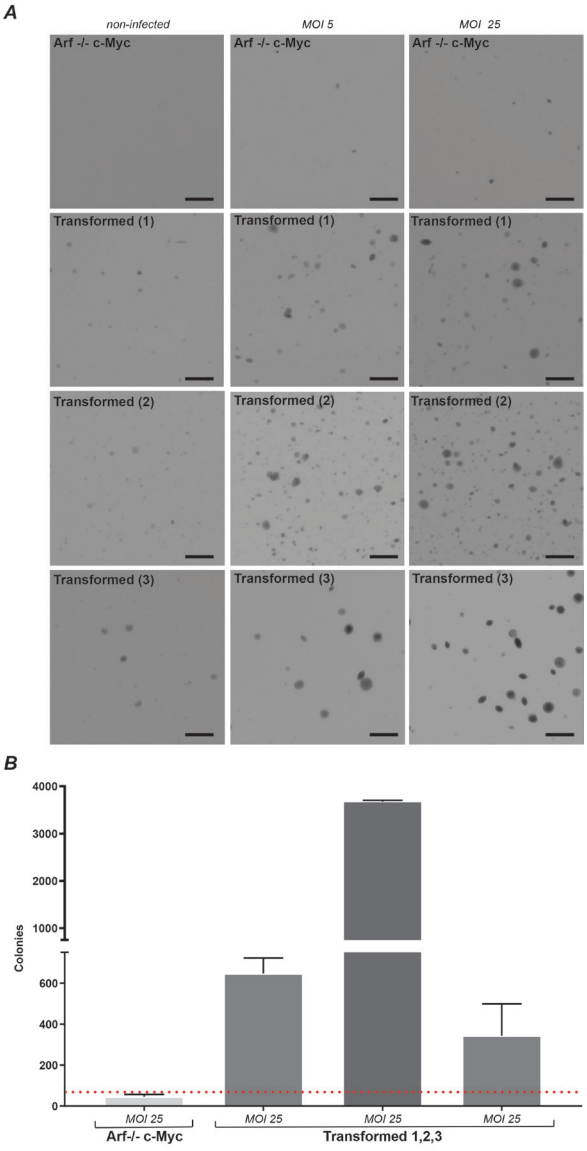


Supplementary Figure 1. Violin plot of the optical density (OD) values of IgA, IgG and IgM of the 36 cases (red) and 72 controls (green).



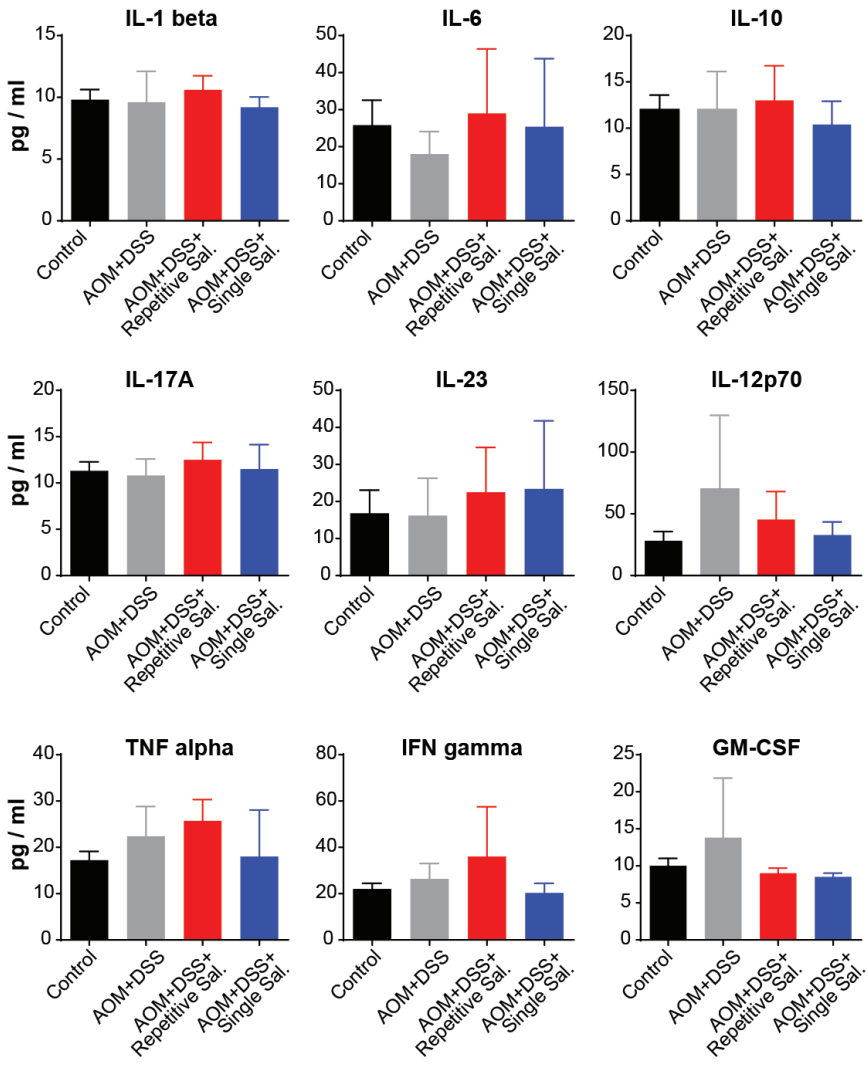
Supplementary Figure 2. Body weight and mortality of the clinical human NTS isolates.

(A) Percent of body weight change throughout the experiment for indicated groups of the mice cohort. The animals (5-female and 5-male per groups) were inoculated with indicated Salmonella isolates (1×105 bacteria per mouse). Data was expressed by mean ± SD, n=10 mice each group. ***(B)*** Mortality of the infected animals throughout the experiment for indicated groups. n=10 mice each group.



Supplementary Figure 3. Repeated *Salmonella* infection increases cellular transformation.

(*A*) Representative images of anchorage-independent growth assays of *Arf-/-* c-MYC MEFs that have not previously encountered *Salmonella* (top panels: *Arf-/-* c-MYC) and of 3 *Arf-/-* c-MYC MEFs that did previously encountered *Salmonella* (bottom 3 panels; *Transformed (1), (2) and (3)*). *Arf-/-* c-MYC or transformed MEFs either non-infected, infected with an MOI of 5 or infected with an MOI of 25 are indicated in the left, middle and right panel, respectively. Scale bar: 750 µm. **(*B*)** Average number of soft agar colonies per well of naïve and transformed *Arf-/-* c-MYC MEFs that have been either non-infected, or infected with an MOI of 25. Results are averages of three technical triplicates.



Supplementary Figure 4: Cytokine and chemokines in plasma from mouse model. The cytokines and chemokines in the serum from the experimental animals were evaluated with the Luminex kit as product's instructions. The data was expressed as mean ± SD; one-way ANOVA, n=5 per group.

Chapter 5

ASSOCIATION BETWEEN SALMONELLA INFECTION AND COLON CANCER: A NATIONWIDE REGISTRY-BASED COHORT STUDY

J.W. Duijster*
J.V. Hansen*
E. Franz
J.J.C. Neefjes
M. Frisch
L. Mughini-Gras
S. Ethelberg

*equal contributors

Published in: *Epidemiology & Infection, 2021*, 149:E56

Abstract

Laboratory data increasingly suggest that *Salmonella* infection may contribute to colon cancer (CC) development. Here, we examined epidemiologically the potential risk of CC associated with salmonellosis in the human population. We conducted a population-based cohort study using four health registries in Denmark. Person-level demographic data of all residents were linked to laboratory-confirmed non-typhoidal salmonellosis and to CC diagnoses in 1994–2016. Hazard ratios (HRs) for CC (overall/proximal/distal) associated with reported salmonellosis were estimated using Cox proportional hazard models. Potential effects of serovar, age, sex, inflammatory bowel disease and follow-up time post-infection were also assessed. We found no increased risk of CC ≥1 year post-infection (HR 0.99; 95% confidence interval (CI) 0.88–1.13). When stratifying by serovar, there was a significantly increased risk of proximal CC ≥1 year post-infection with serovars other than Enteritidis and Typhimurium (HR 1.40; 95% CI 1.03–1.90). CC risk was significantly increased in the first year post-infection (HR 2.08; 95% CI 1.48–2.93). The association between salmonellosis and CC in the first year post-infection can be explained by increased stool testing around the time of CC diagnosis. The association between proximal CC and non-Enteritidis/non-Typhimurium serovars is unclear and warrants further investigation. Overall, this study provides epidemiological evidence that notified *Salmonella* infections do not contribute significantly to CC risk in the studied population.

Introduction

Colon cancer (CC) is the third most common cancer in industrialised countries, with 1.1 million new diagnoses annually worldwide [1]. Although genetic, environmental and lifestyle-related exposures are the best-known risk factors for cancer, around 20% of the global cancer burden is estimated to be attributable to infectious agents, including bacteria [2]. Examples hereof concerning the gastrointestinal tract include Helicobacter pylori infection as risk factor for gastric cancer, and *Salmonella* Typhi infection as risk factor for gallbladder carcinoma in chronic typhoid carriers [3–5].

Several mechanisms have been identified through which bacteria can contribute to cancer formation. These include chronic inflammation, production of DNA-damaging toxins and manipulation of host cell signaling pathways [3, 4, 6]. The latter promotes bacterial uptake, intracellular survival and egress in case of *Salmonella* infection. Indeed, several *Salmonella* effector proteins have been shown to activate the major host cell signaling pathways AKT and MAPK, which are central to many signaling cascades and are often deregulated in cancers [4]. *Salmonella* is expected to contribute to carcinogenesis mainly under conditions of long-lasting infections, an intact bacterial type 3 secretion System (T3SS), and with a background of host predisposition, in which significant numbers of pre-transformed cells are present in the intestine. This has been shown in vivo by experiments demonstrating a higher risk of colon carcinoma formation after infection with wild type vs. ΔprgH mutant *S*. Typhimurium (lacking the T3SS) strains in mice genetically predisposed to cancer ($APC^{+/-}$) *vs.* normal mice [4].

Against this background of experimental data, population-based epidemiological studies addressing the association between *Salmonella* infection and CC are limited to one [7]. In a nationwide registry study in the Netherlands, an increased risk of CC was observed among patients who had a reported (severe) *Salmonella* infection between 20 and 60 years of age as compared to the baseline CC risk in the Dutch population [7]. This increased risk was significant following infection with *S*. Enteritidis and for the proximal part of the colon. Moreover, it was shown that among CC patients, the risk of having had a previously notified *Salmonella* infection was higher for individuals with pre-infectious inflammatory bowel disease (IBD), although numbers were small [7]. IBD is a known risk factor for both CC and salmonellosis, as this chronic condition is associated with recurrent episodes of gut inflammation and increased susceptibility to infection and testing [8, 9].

Salmonella is a major cause of bacterial gastroenteritis worldwide, with over 90 000 infections reported to public health authorities in Europe each year [10]. In Denmark, an annual

average of 1100 salmonellosis cases has been reported in recent years through the national surveillance system [11]. As most *Salmonella* infections are mild with self-limiting symptoms, the majority of infections go unreported. It is estimated that the true number of *Salmonella* infections (i.e. after correction for underreporting and underdiagnosis) is approximately 10 times higher than the number of infections reported in the national disease surveillance system in Denmark [12]. Each year, around 3400 people are diagnosed with CC in the Danish population [13]. Although screening programs aiming at early detection of CC typically target the older population (i.e. individuals aged >50 years), the incidence of CC in young adults has increased during the last 25 years, being a cause for concern [14].

In this study, we assessed the potential association between *Salmonella* infection and CC in Denmark. To this end, we made use of data from comprehensive health registries in Denmark to compare the incidence of CC among individuals with a previously reported salmonellosis to that of individuals without reported salmonellosis. In addition, we assessed potential effects in subgroup analyses as defined by age, sex, IBD and time since infection on the association between *Salmonella* infection and CC.

Methods

Data sources

We conducted a population-based cohort study with data from four health registries in Denmark between January 1994 and December 2016. Demographic characteristics including sex, date of birth, vital status (e.g. date of death, immigration and emigration), marital status and region of living of all people residing in Denmark were retrieved from the Danish Civil Registration System [15]. A second dataset included information on bacterial gastrointestinal infections, with recorded bacterial species and subspecies/ serovar and date of diagnosis (Danish Register on Enteric Pathogens) [16]. The presence of IBD, i.e. ulcerative colitis and Crohn's disease with date of diagnosis, was obtained from the Danish National Patient Registry [17]. The fourth dataset contained all CC diagnoses from 1978 until December 2016 reported to the Danish Cancer Registry, with date of diagnosis and tumor location (based on ICD-10 code) [18]. Data of all four sources were matched using the CPR-number, which is a unique identifier used across all national registries [15].

Study population

The cohort consisted of 7 646 978 individuals who contributed at least 1 day of follow-up between 1994 and 2016, of which 47 856 had been diagnosed with a *Salmonella* infection.

Median age at infection was 34 years (interquartile range (IQR): 14–54). S. Enteritidis (SE) (43.5%) and (monophasic) S. Typhimurium (ST) (28.6%) caused the majority of reported infections. Among the more than 400 other reported serovars (hereafter referred to as 'other serovars'), S. Infantis, S. Newport and S. Stanley were the most frequent.

Exposure and outcome definition

The exposure variable was defined as having or not having had a reported non-typhoidal *Salmonella* infection. *Salmonella* infection was categorized into infections with SE, ST or other serovars. For individuals with multiple *Salmonella* infections, only the first reported infection was considered. In analyses restricting exposure to a serovar, only the first infection of the serovar of interest was used. Considering a minimal development time of 1 year for CC formation after infection, which has been assumed previously to have a plausible relation to the infection [7, 19], people were considered at risk of CC from 1 year after reported *Salmonella* infection onwards. Hence, we defined the exposure status as a time-varying variable with three states: individuals were 'unexposed' (reference) until first reported infection, 'newly exposed' in the first year post-infection and 'exposed' from 1 year post-infection onwards. We excluded individuals with a diagnosed CC between January 1978 and December 1993, to reduce the risk that CC had developed before the *Salmonella* infection occurred. The outcome studied was CC (ICD-10 codes C180–C187). For the analysis, we looked at CC overall and by colon subsite: proximal colon (C180–C185) and distal colon (C186, C187). In the analyses of risk of cancer in one colon subsite individuals were not censored for cancers in the other subsite.

Statistical analysis

Characteristics of the study population were presented descriptively. In the survival analyses, individuals were followed from birth or 1st January 1994, whichever was last. Follow-up ended at date of cancer diagnosis, death or the end of study (31st December 2016), whichever occurred first. In addition, risk time excluded periods where individuals were temporarily or permanently living outside of Denmark. Three of the potential confounders; geographical region, marital status and IBD status were time-varying variables with five (North Jutland, South Jutland, Middle Jutland, Zealand, Capital), four (unmarried, married, divorced, widowed) and two (yes, no) levels, respectively.

We used Cox proportional hazard models to estimate hazard ratios (HRs) with 95% confidence intervals (CIs) for developing CC in individuals with a history of reported *Salmonella* infection vs. people without such history. The main comparison of interest was exposed vs. unexposed;

hence, all analyses show the HRs for this comparison. Besides, in the main analysis the HRs for 'newly exposed' vs. unexposed were displayed to address the potential effect of testing/diagnostic bias in symptomatic individuals with yet undiagnosed CC [20]. Attained age was used as the time scale for the baseline hazard function, which was stratified by sex, year of birth, geographical region, marital status and IBD to adjust for potential confounding effects. Additionally, we conducted analyses to examine whether HRs varied by sex, attained age (<50, 50–59, 60–69, 70–79 and ≥80 years), age at infection (<50, 50–59, 60–69, 70–79 and ≥80 years), follow-up time post-infection (2nd–5th year, 6th–10th year and >10 years) and IBD. The proportional hazards assumption of the main analysis (salmonellosis overall and CC overall) was assessed using a test for homogeneity of the HR in the age intervals; <50, 50–59, 60–69, 70–79 and ≥80 years of age. Incidence curves of CC (overall and by subsite) in the exposed vs. unexposed group (stratified by *Salmonella* serovar) were also generated to graphically display the comparison; incidences at all ages were calculated by weighting the number of cancers and risk days within a time span of ±5 years using a parabolic kernel. In all analyses, P-values <0.05 were considered statistically significant. In accordance with privacy legislation, small numbers were not displayed in tables. Statistical analysis was performed using the PHREG procedure of SAS (version 9.4).

Results

During a total of 124.7 million person-years of follow-up, 54 902 individuals were diagnosed with CC, at a median age of 72 years (IQR: 64–80). Among those with a CC diagnosis, 278 individuals were diagnosed with CC after salmonellosis, of which 33 occurred within the first year post-infection. The median time span between infection and CC diagnosis was 7.5 years (IQR: 3.0–13.9). In the subsite-specific analyses, 29 422 individuals were diagnosed with proximal CC and 26 108 with distal CC. Table 1 shows the number of overall CC events and incidence rates (IRs) in the exposed and unexposed groups by different subgroups. The average IR of CC in the exposed group was 47.16 per 100 000 person-years at-risk, whereas in the unexposed group the IR was 44.02.

Table 1. IRs of colon cancer (overall) of people with ('exposed') and without ('unexposed') a reported *Salmonella* infection per 100 000 person-years, by different subgroups

	Unexposed		Exposed[a]		Total	
	No. events	IR[b]	No. events	IR[b]	No. events	IR[b]
Total	54,624	44.02	278	47.16	54,902	44.04
Sex						
Female	28,124	44.90	138	46.20	28,262	44.90
Male	26,500	43.14	140	48.14	26,640	43.16
Birth year						
≤1930	21,964	225.23	71	275.19	22,035	225.36
1931-1950	26,474	100.93	159	130.97	26,633	101.06
1951-1970	5,728	16.27	42	25.03	5,770	16.31
1971-1990	415	1.36	6	4.05	421	1.38
≥1991	43	0.19	0	0.00	43	0.19
Age group						
0-49 years	2,368	2.93	17	4.48	2,385	2.94
50-59 years	5,942	36.06	34	40.67	5,976	36.08
60-69 years	13,638	103.74	79	113.83	13,717	103.80
70-79 years	18,828	217.40	88	228.57	18,916	217.45
≥80 years	13,848	278.74	60	331.49	13,908	278.94
Marital status						
Married	30,014	61.20	164	67.57	30,178	61.23
Divorced	6,011	62.28	32	65.17	6,043	62.29
Not married	4,117	7.10	24	8.94	4,141	7.11
Widowed	14,482	196.10	58	199.31	14,540	196.11
Region						
North Jutland	6,064	45.93	28	50.72	6,092	45.95
South Jutland	12,390	45.83	97	65.41	12,487	45.94
Middle Jutland	11,412	40.96	48	35.93	11,460	40.93
Zealand	9,132	49.90	46	48.12	9,178	49.89
Capital	15,626	41.47	59	37.65	15,685	41.46
IBD status						
No	51,854	42.77	249	44.56	52,103	42.78
Yes	2,770	97.97	29	94.77	2,799	97.93

IR: incidence rate. IBD: Inflammatory Bowel Disease [a] Per 100 000 person-years. [b] Including both 'newly exposed' and exposed.

Risk of colon cancer

Adjusting for sex, year of birth, region of residence, IBD and marital status, the overall risk of CC did not differ between the exposed and unexposed groups (HR: 0.99 [95% CI 0.88–1.13]) (Table 2). Similarly, no differences were observed between these groups when stratifying by colon subsite and sex. However, within 1 year post-

infection the overall risk of CC increased twofold compared to the unexposed group (HR: 2.08 [95% CI 1.48–2.93]). For the exposed group, when stratifying by serovar, an HR of 1.40 (95% CI 1.03–1.90) for cancer in the proximal colon was observed in individuals who had an infection with serovars other than SE and ST (Table 2).

Table 2. Risk of colon cancer after salmonellosis, by sex, serovar and IBD status

		CC (overall)		**Proximal colon**		**Distal colon**
	Events	HR (95%CI)[a]	Events	HR (95%CI)[a]	Events	HR (95%CI)[a]
Unexposed		ref		ref		ref
Newly exposed	33	2.08 (1.48-2.93)***	18	2.16 (1.36-3.43)**	15	1.96 (1.18-3.26)**
Exposed	245	0.99 (0.88-1.13)	144	1.09 (0.93-1.29)	102	0.87 (0.71-1.05)
Exposed vs. unexposed						
Sex						
Male	121	0.99 (0.83-1.18)	59	0.99 (0.77-1.28)	62	0.96 (0.75-1.24)
Female	124	1.00 (0.84-1.19)	85	1.18 (0.95-1.46)	40	0.75 (0.55-1.03)
***Salmonella* serovar**						
Enteritidis	137	1.03 (0.87-1.21)	74	1.04 (0.83-1.31)	63	0.99 (0.77-1.26)
Typhimurium	43	0.74 (0.55-1.00)*	28	0.90 (0.62-1.31)	15	0.54 (0.33-0.90)*
Other serovars	65	1.17 (0.91-1.49)	42	1.40 (1.03-1.90)*	24	0.91 (0.61-1.36)
IBD						
Yes	28	1.18 (0.81-1.71)	20	1.40(0.90-2.19)	8	0.80(0.40-1.61)
No	217	0.98 (0.85-1.11)	124	1.06(0.89-1.26)	94	0.87(0.71-1.07)

CC: colon cancer; HR: Hazard ratio; IBD: Inflammatory Bowel Disease; ref: Reference. [a] Adjusted for sex, year of birth, geographical region, IBD status, marital status. * p-value <0.05; ** p-value <0.01; *** p-value <0.001.

The association between *Salmonella* infection and CC did not vary by attained age (Table 3). A test for homogeneity of HRs in the five age groups yielded a P-value of 0.59. Figure 1 shows the incidence of CC (overall and per colon subsite) by attained age for the different serovars. For both proximal and distal CC, the IRs were the lowest for ST in people aged above 60 years as compared to SE and other serovars. In the age-stratified analyses, a 1.87-fold (95% CI 1.00–3.50) increased risk of distal CC was observed in the exposed group aged 0–49 years (for Salmonella overall). For the proximal colon, the highest HR, although not significant, was also observed in the age group 0–49 years among those infected with other serovars (HR: 1.75; 95% CI 0.56–5.44). The estimated association between Salmonella and CC risk did not vary much by age at infection (Table 4). The median observed ages at infection of different serovars in the total cohort were 37 years (IQR: 16–54) for SE, 30 years (IQR: 9–52) for ST and 32 years (IQR: 16–54) for other serovars.

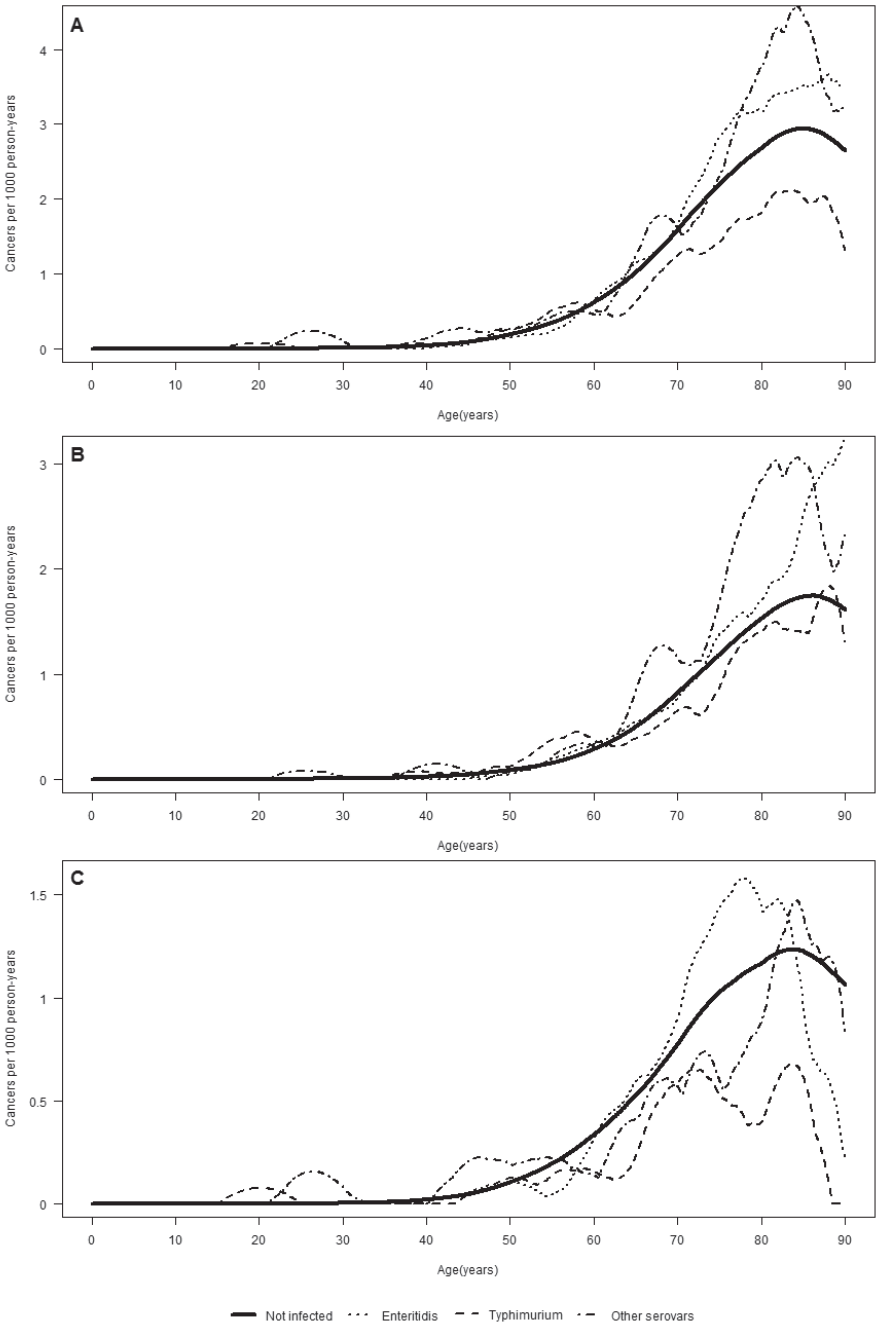


Figure 1. Incidence of overall (a), proximal (b) and distal (c) colon cancer by attained age, stratified by serovar.

Table 3. Risk of colon cancer ≥1 year after salmonellosis, by attained age group and serotype

	Colon cancer (overall)	Proximal colon	Distal colon
	HR (95%CI)[a]	HR (95%CI)[a]	HR (95%CI)[a]
***Salmonella* (total)**			
Unexposed	ref	ref	ref
<50 years	1.39 (0.85-2.27)	0.96 (0.43-2.14)	1.87 (1.00-3.50)*
50-59 years	0.90 (0.62-1.32)	1.32 (0.83-2.11)	0.54 (0.28-1.05)
60-69 years	0.97 (0.77-1.23)	1.08 (0.78-1.48)	0.89 (0.63-1.25)
70-79 years	0.93 (0.75-1.17)	0.97 (0.72-1.30)	0.87 (0.62-1.22)
≥80 years	1.09 (0.84-1.43)	1.24 (0.89-1.71)	0.83 (0.52-1.34)
***Salmonella* Enteritidis**			
Unexposed	ref	ref	ref
<50 years	0.70 (0.26-1.88)	0.34 (0.05-2.39)	1.09 (0.35-3.40)
50-59 years	0.62 (0.33-1.15)	0.95 (0.45-2.01)	0.34 (0.11-1.04)
60-69 years	1.06 (0.78-1.44)	0.99 (0.63-1.55)	1.11 (0.73-1.69)
70-79 years	1.10 (0.84-1.46)	1.00 (0.67-1.47)	1.19 (0.80-1.77)
≥80 years	1.18 (0.83-1.67)	1.34 (0.88-2.03)	0.87 (0.47-1.63)
***Salmonella* Typhimurium**			
Unexposed	ref	ref	ref
<50 years	1.39 (0.52-3.71)	1.23 (0.31-4.96)	1.57 (0.39-6.30)
50-59 years	1.24 (0.65-2.40)	1.83 (0.82-4.09)	0.75 (0.24-2.33)
60-69 years	0.68 (0.38-1.24)	0.88 (0.42-1.85)	0.49 (0.18-1.30)
70-79 years	0.60 (0.34-1.05)	0.73 (0.36-1.45)	0.43 (0.16-1.14)
≥80 years	0.59 (0.28-1.25)	0.70 (0.29-1.69)	0.42 (0.10-1.66)
Other *Salmonella* serovars			
Unexposed	ref	ref	ref
<50 years	2.61 (1.30-5.25)**	1.75 (0.56-5.44)	3.65 (1.51-8.86)**
50-59 years	1.19 (0.60-2.39)	1.62 (0.67-3.90)	0.82 (0.26-2.55)
60-69 years	1.03 (0.64-1.67)	1.45 (0.82-2.55)	0.73 (0.33-1.62)
70-79 years	0.89 (0.55-1.43)	1.14 (0.65-2.01)	0.57 (0.24-1.36)
≥80 years	1.44 (0.87-2.40)	1.57 (0.84-2.93)	1.18 (0.49-2.85)

HR: hazard ratio; ref: reference. [a] Adjusted for sex, year of birth, geographical region, IBD status, marital status.
* p-value <0.05; ** p-value <0.01; *** p-value <0.001.

Table 4. Risk of colon cancer ≥1 year after salmonellosis, by age group at infection and serotype

	Colon cancer (overall)	Proximal colon	Distal colon
	HR (95%CI)[a]	HR (95%CI)[a]	HR (95%CI)[a]
***Salmonella* (total)**			
Unexposed	ref	ref	ref
<50 years	1.10 (0.84-1.45)	1.07 (0.72-1.60)	1.12 (0.76-1.63)
50-59 years	0.91 (0.71-1.16)	1.02 (0.74-1.42)	0.77 (0.53-1.13)
60-69 years	0.95 (0.75-1.20)	1.08 (0.80-1.45)	0.80 (0.55-1.17)
70-79 years	1.11 (0.84-1.47)	1.22 (0.86-1.73)	0.92 (0.58-1.47)
≥80 years	0.91 (0.50-1.65)	1.13 (0.56-2.26)	0.57 (0.18-1.77)
***Salmonella* Enteritidis**			
Unexposed	ref	ref	ref
<50 years	0.90 (0.60-1.36)	0.91 (0.50-1.64)	0.89 (0.51-1.57)
50-59 years	0.90 (0.65-1.26)	0.87 (0.54-1.39)	0.93 (0.58-1.47)
60-69 years	1.08 (0.80-1.46)	1.08 (0.72-1.61)	1.06 (0.68-1.67)
70-79 years	1.27 (0.89-1.82)	1.31 (0.83-2.09)	1.15 (0.65-2.02)
≥80 years	1.03 (0.46-2.30)	1.19 (0.44-3.19)	0.76 (0.19-3.09)
***Salmonella* Typhimurium**			
Unexposed	ref	ref	ref
<50 years	1.16 (0.67-2.00)	1.47 (0.74-2.95)	0.86 (0.36-2.07)
50-59 years	0.88 (0.51-1.52)	1.08 (0.54-2.16)	0.67 (0.28-1.62)
60-69 years	0.51 (0.27-0.99)	0.63 (0.28-1.41)	0.36 (0.12-1.13)
70-79 years	0.52 (0.23-1.16)	0.59 (0.22-1.57)	0.41 (0.10-1.65)
≥80 years	0.64 (0.16-2.58)	1.06 (0.27-4.28)	-
Other *Salmonella* serovars			
Unexposed	ref	ref	ref
<50 years	1.52 (0.91-2.52)	1.00 (0.42-2.41)	2.00 (1.08-3.73)*
50-59 years	0.93 (0.56-1.54)	1.33 (0.73-2.40)	0.50 (0.19-1.32)
60-69 years	1.08 (0.68-1.71)	1.54 (0.91-2.61)	0.64 (0.27-1.54)
70-79 years	1.43 (0.85-2.42)	1.73 (0.93-3.23)	0.95 (0.36-2.55)
≥80 years	0.94 (0.30-2.93)	1.06 (0.26-4.24)	0.73 (0.10-5.22)

HR: hazard ratio; ref: reference. [a] Adjusted for sex, year of birth, geographical region, IBD status, marital status.
* p-value <0.05; ** p-value <0.01; *** p-value <0.001.

There was no significant effect of follow-up time post-infection on CC risk; the HRs of proximal CC for people infected with other serovars were 1.62 (95% CI 0.96–2.74), 1.47 (95% CI 0.85–2.53) and 1.20 (95% CI 0.72–1.90) at 1–5 years, 5–10 years and >10 years post-infection, respectively (Table 5). With regards to the potential effect modification of IBD on CC risk, the HR for overall CC was not significantly higher for people with underlying IBD (HR 1.18; 95% CI 0.81–1.71) (Table 2).

Table 5. Risk of colon cancer ≥1 year after salmonellosis, by serovar and time post-infection

	Colon cancer (overall)	Proximal colon	Distal colon
	HR (95%CI)[a]	HR (95%CI)[a]	HR (95%CI)[a]
***Salmonella* (total)**			
Unexposed	ref	ref	ref
1-5 years	1.09 (0.85-1.38)	1.29 (0.95-1.75)	0.88 (0.60-1.29)
5-10 years	0.96 (0.76-1.22)	1.03 (0.75-1.42)	0.86 (0.60-1.24)
>10 years	0.97 (0.80-1.16)	1.03 (0.81-1.32)	0.87 (0.65-1.16)
***Salmonella* Enteritidis**			
Unexposed	ref	ref	ref
1-5 years	1.22 (0.88-1.70)	1.30 (0.83-2.03)	1.12 (0.69-1.83)
5-10 years	0.98 (0.71-1.36)	1.02 (0.66-1.58)	0.91 (0.56-1.48)
>10 years	0.97 (0.76-1.23)	0.95 (0.68-1.33)	0.97 (0.68-1.38)
***Salmonella* Typhimurium**			
Unexposed	ref	ref	ref
1-5 years	0.86 (0.51-1.45)	0.92 (0.46-1.83)	0.78 (0.35-1.73)
5-10 years	0.45 (0.23-0.90)*	0.64 (0.29-1.42)	0.24 (0.06-0.95)*
>10 years	0.87 (0.57-1.34)	1.09 (0.64-1.83)	0.61 (0.29-1.29)
Other *Salmonella* serovars			
Unexposed	ref	ref	ref
1-5 years	1.05 (0.65-1.69)	1.62 (0.96-2.74)	0.52 (0.20-1.39)
5-10 years	1.45 (0.97-2.17)	1.47 (0.85-2.53)	1.41 (0.78-2.54)
>10 years	1.04 (0.70-1.55)	1.20 (0.72-1.99)	0.83 (0.43-1.60)

HR: hazard ratio; ref: reference. [a] Adjusted for sex, year of birth, geographical region, IBD status, marital status.
* p-value <0.05; ** p-value <0.01; *** p-value <0.001.

Discussion

We assessed the risk of CC after reported *Salmonella* infection in a 23-year follow-up of the entire Danish population. The risk of CC in individuals with reported salmonellosis was compared to the risk in individuals without a reported salmonellosis, accounting for potential confounding and modifying effects of age, sex, IBD and follow-up time post-infection. Overall, we observed no increased risk of CC among salmonellosis cases. The only significantly increased risk of CC concerned the proximal colon after infection with serovars other than SE and ST. The proximal part of the colon is the subsite of primary interest, as exposure to *Salmonella* is highest in this part of the large intestine located directly after the ileum, where *Salmonella* typically establishes infection. CC risk was highest at an attained age of <50 years for those infected with other serovars. For SE, the estimated HRs increased with increasing age, whereas for ST they decreased with increasing age, which may be due to differences in age-specific reporting between these serovars.

In a recent Dutch study, a statistically significantly increased risk of cancer in the proximal colon was found among individuals with a history of SE infection [7]. This result was not confirmed in the current study, indicating that the previously observed association between SE infection and proximal CC is not generalizable to other study populations. On the one hand, the inconsistent findings might be explained by a more complex causal mechanism than originally anticipated and the existence of situational differences, but may also represent a chance finding. Indeed, our results seem to indicate a possible scenario of increased risk of proximal CC associated with infection with a *Salmonella* serovar other than SE or ST, but it could also be the result of type I error due to multiple hypothesis testing.

For surveillance design reasons, the two studies used different types of analyses and effect measures. The Dutch *Salmonella* surveillance system covers approximately 64% of the population; therefore, the risk of CC in individuals with a reported salmonellosis was compared to the baseline CC risk in the general Dutch population, expressed as standardized incidence ratios [7]. The Danish surveillance system covers the whole population, which allowed us to compare the risk of CC in people with those without a reported salmonellosis using Cox regression. Yet, both studies used individual-level data and the inclusion criteria were comparable. Apart from the aforementioned possibility of a chance finding, other and largely unknown factors might underlie this dissimilarity including, for instance, different serovar distributions and populations exposed to them. Disease outcomes (e.g. severity, antimicrobial resistance, etc.) and epidemiology (e.g. sources, modes of transmission, high-risk groups, etc.) differ by serovar, partly due to differences in exposure but also potential factors related to virulence, invasiveness and toxins of the bacterium itself [21]. The estimated number of *Salmonella* infections in Denmark is somewhat higher compared to the Netherlands, with respectively 18.1 and 15.8 infections per 10 000 inhabitants in Denmark and the Netherlands in 2017 [12, 22]. In both the Netherlands and Denmark, SE accounts for a substantial part (25–30%) of the salmonellosis cases [11, 23]. The successful implementation of a *Salmonella* control program in the poultry production chain led to a marked reduction of domestically acquired human SE infections in Denmark since 1998, with most infections nowadays being attributable to foreign travel (78.2% in 2016) [11, 24]. In contrast, most SE infections in the Netherlands remain domestically acquired; therefore, the groups of people infected with SE might not be fully comparable in terms of, e.g. general health status, lifestyle, socio-economic status, ethnicity and possible co-morbidities. Besides, with regards to serovars other than SE and ST, different distribution and exposure patterns, as well as specific strains, might also have contributed to these differences, as the genetic makeup of the strains themselves might be associated with their ability to transform [21].

It has been shown experimentally that *Salmonella* infection of pre-transformed fibroblasts and organoids induces full cell transformation [4]. Development from a pre-malignant state to an advanced carcinoma takes several years; however, *Salmonella* infection is likely to accelerate this process [4, 25]. The results of a sub-analysis showed a twofold increased risk of CC (overall and per subsite) for individuals within the first year post-infection. Even though *Salmonella* could accelerate transformation, tumor development in less than 1 year seems implausible. We therefore consider this observation to reflect testing/diagnostic bias rather than the transformation capacity of *Salmonella*. Undiagnosed CC patients often present at their general practitioner (GP) with nonspecific symptoms resembling gastroenteritis, such as diarrhea and frequent bowel movements. In a Danish cohort, it was shown that CC patients had significantly more GP consultations in the 9 months prior to the cancer diagnosis [17]. A similar pattern was observed in another Danish cohort study that examined the risk of IBD after a *Salmonella*-positive stool test. An increased risk of IBD was observed in the first year after *Salmonella* infection; however, this was even more pronounced in the first year following a negative stool test [9]. The association we found in the first year post-infection is compatible with these prior observations. An alternative hypothesis might be that people in an early stage of cancer are more susceptible to *Salmonella* infection due to dysbiosis or other changes in the gut microbiome [26], so the association observed in the first year after infection might also reflect reverse causality. Still, both the testing/diagnostic bias and the increased susceptibility could co-exist in people with an early-stage pre-diagnosed CC.

This study has some limitations. First, mainly severe infections, outbreak-related infections or infections with a suspected foreign source are included, as most people do not present at their GP with mild and self-limiting gastrointestinal complaints. Hence, we could not assess whether multiple mild *Salmonella* infections that are undiagnosed and unreported contribute to CC risk or not. This could be the subject of another study using, for instance, serology to measure the magnitude of exposure to *Salmonella* regardless of reporting bias. Second, in the Dutch cohort, the risk of cancer was only significantly increased for enteric infections and not for invasive (bloodstream) infections [7], but we were not able to address this observation in the current study. Third, we were not able to control for some of the main risk factors for CC, such as obesity, smoking and alcohol consumption. Although considering these variables would be relevant to explain CC risk along with the studied *Salmonella* infection, this would require a different study design as these types of time-varying variables are not generally present in national health registries.

In conclusion, the current study found no unusual CC risk associated with previously reported *Salmonella* infection overall. Therefore, although there is growing experimental

evidence for a potential role of *Salmonella* in CC development, notified *Salmonella* infections do not appear to be an important driver of CC risk in the studied population. Indeed, the previously observed epidemiological association between SE infection and proximal CC was not confirmed here. The explanation for these differences, if not merely occurring by chance, is unclear.

Author contributions

Study concept and design: LMG, SE, JVH, JWD, MF. Formal/statistical analysis: JVH. Writing and/or original draft: JWD, JVH. Data collection: SE. Writing and/or review and editing: JWD, JVH, LMG, SE, EF, JJCN, MF.

Financial support

This study was supported by KWF grant 'Bacterial food poisoning and colon cancer; a cell biological and epidemiological study' (2017-1-11001). No funding bodies had any role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Conflict of interest

All authors report no potential conflicts of interest.

Ethical standards

This study was approved by the Danish Data Protection Agency (permission number: 18/04190). By Danish law, ethical approval is not required for register-based studies.

References

1. WHO. International Agency for Research on Cancer. Estimated number of new cases in 2018. Available: http://gco.iarc.fr/ [Accessed 15 Nov 2019].
2. Parkin DM. The global health burden of infection-associated cancers in the year 2002. Int J Cancer 2006;118:3030-3044.
3. Van Elsland D, Neefjes J. Bacterial infections and cancer. EMBO reports 2018;19:e46632.
4. Scanu T, Spaapen RM, Bakker JM, et al. Salmonella manipulation of host signaling pathways provokes cellular transformation associated with gallbladder carcinoma. Cell Host Microbe 2015;17:763-774.
5. Kikuchi S. Epidemiology of Helicobacter pylori and gastric cancer. Gastric Cancer 2002;5:6–15.
6. Gagnaire A, Nadel B, Raoult D, et al. Collateral damage: insights into bacterial mechanisms that predispose host cells to cancer. Nat Rev Microbiol 2017;15:109.
7. Mughini-Gras L, Schaapveld M, Kramers J, et al. Increased colon cancer risk after severe Salmonella infection. PloS One 2018;13:e0189721.
8. Dulai PS, Sandborn WJ, Gupta S. Colorectal cancer and dysplasia in inflammatory bowel disease: a review of disease epidemiology, pathophysiology, and management. Cancer Prev Res 2016;9:887-894.
9. Jess T, Simonsen J, Nielsen NM, et al. Enteric Salmonella or Campylobacter infections and the risk of inflammatory bowel disease. Gut 2011;60:318-324.
10. European Food Safety Authority and European Centre for Disease Prevention and Control (EFSA and ECDC). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. EFSa Journal 2018;16:e05500.
11. Anonymous, 2016. Annual Report on Zoonoses in Denmark 2016, National Food Institute, Technical University of Denmark. Available: *www.food.dtu.dk › media › Foedevareinstituttet › Publikationer › Pub-2017* [Accessed 15 Nov 2019].
12. Pires SM, Jakobsen LS, Ellis-Iversen J, et al. Burden of disease estimates of seven pathogens commonly transmitted through foods in Denmark, 2017. Foodborne Pathog Dis 2020;17(5):322-339.
13. Statistik om kræft i tyktarmen. Available: https://www.cancer.dk/tyktarmskraeft-coloncancer/statistik-tyktarmskraeft/ [Accessed 15 Nov 2019].
14. Vuik FE, Nieuwenburg SA, Bardou M, et al. Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. Gut 2019: gutjnl-2018-317592.
15. Schmidt M, Pedersen L, Sørensen HT. The Danish Civil Registration System as a tool in epidemiology. Eur J Epidemiol 2014;29:541-549.
16. Ethelberg S, Olsen KE, Gerner-Smidt P, Mølbak K. Household outbreaks among culture-confirmed cases of bacterial gastrointestinal disease. Am J Epidemiol 2004;159:406-412.
17. Schmidt M, Schmidt SAJ, Sandegaard JL, et al. The Danish National Patient Registry: a review of content, data quality, and research potential. Clin Epidemiol 2015;7:449.
18. Gjerstorff ML. The Danish cancer registry. Scand J Public Health 2011;39:42-45.
19. Brenner, H., Altenhofen, L., Katalinic, A., Lansdorp-Vogelaar, I., & Hoffmeister, M. Sojourn time of preclinical colorectal cancer by sex and age: estimates from the German national screening colonoscopy database. American journal of epidemiology 2011;174(10), 1140-1146.

20. Hansen PL, Hjertholm P, Vedsted P. Increased diagnostic activity in general practice during the year preceding colorectal cancer diagnosis. Int J Cancer 2015;137:615-624.
21. Kuijpers AF, Bonacic Marinovic AA, Wijnands L. et al. Phenotypic prediction: linking in vitro virulence to the genomics of 59 Salmonella enterica strains. Front Microbiol 2018;9:3182.
22. Mangen, M.J.J., Disease burden of food-related pathogens in the Netherlands, 2017. RIVM Letter report 2018-0037. 2018, National Institute for Public Health and the Environment: Bilthoven. [Accessed 16 Oct 2020].
23. National Institute for Public Health and the Environment (RIVM). Staat van Zoönosen 2017. Available: http://www.rivm.nl/bibliotheek/rapporten/2018-0112.pdf [Accessed 15 Nov 2019].
24. Mølbak K, Neimann J. Risk factors for sporadic infection with Salmonella Enteritidis, Denmark, 1997–1999. Am J Epidemiol 2002;156:654-661.
25. Jones S, Chen WD, Parmigiani G. et al. Comparative lesion sequencing provides insights into tumor evolution. Proc Natl Acad Sci 2008;105:4283-4288.
26. Garrett WS. The gut microbiota and colon cancer. Science 2019;364:1133-1135.

Chapter 6

SEVERE SALMONELLA SPP. OR CAMPYLOBACTER SPP. INFECTION AND THE RISK OF BILIARY TRACT CANCER: A POPULATION-BASED STUDY

E. de Savornin Lohman
J.W. Duijster
B. Groot Koerkamp
R. van der Post
E. Franz
L. Mughini-Gras
P. de Reuver

Published in: *Cancers,* 2020: 12(11)

Abstract

Salmonella spp. infection has shown to have oncogenic transformative effects and thereby increases the risk of certain cancers. For *Campylobacter* spp., similar effects have been demonstrated. Risk factor identification may allow for timely diagnosis and preventive treatment. To substantiate the oncogenic potential of *Salmonella* and *Campylobacter* spp., this study compared the incidence of extrahepatic biliary tract cancer (BTC) in patients with diagnosed *Salmonella* or *Campylobacter* spp. infection with BTC incidence in the Netherlands. National infectious diseases surveillance records of patients diagnosed with a laboratory-confirmed *Salmonella* or *Campylobacter* spp. infection during 1999–2016 were linked to the Netherlands Cancer Registry. Incidence of BTC in *Salmonella* and *Campylobacter* spp. patients was compared to the incidence of BTC in the general population using Standardized Incidence Ratios (SIRs). In total, 16,252 patients were diagnosed with *Salmonella* spp. and 27,668 with *Campylobacter* spp. infection. Nine patients developed BTC at a median of 46 months (13–67) after *Salmonella* spp. infection and seven at a median of 60 months (18–138) after *Campylobacter* spp. infection. SIR of BTC in salmonellosis patients was 1.53 (95% CI 0.70–2.91). In patients aged <60 years, the SIR was 1.74 (95% CI 0.36–5.04). For campylobacteriosis patients, the SIR was 0.97 (95% CI 0.39–2.00). Even though *Salmonella* or *Campylobacter* spp. infection was not significantly associated with increased BTC risk in this cohort, it remains extremely important to study potential risk factors for cancer to facilitate screening and ultimately improve prognosis of cancer patients.

Introduction

Biliary tract cancers (BTC) are rare malignancies of the distal and proximal bile ducts, the gallbladder and the cystic duct. Despite significant improvement in the overall survival of cancer patients, 5-year survival of patients with extrahepatic biliary tract cancer (i.e., gallbladder cancer, proximal and distal cholangiocarcinoma) is still only 10% [1, 2, 3]. Currently, radical surgery is the only curative treatment available. Unfortunately, surgery is not an option in the majority of patients, because BTC frequently goes undetected until the disease has progressed to an advanced, unresectable stage [4, 5].

Geography appears to be the primary risk factor for the development of non-intrahepatic BTC, and as a result, incidence rates vary significantly per region. For example, gallbladder cancer (GBC) incidence ranges from 0.9/100,000 women in the Netherlands to 35/100,000 women in Chile [6, 7]. Other risk factors for BTC include age, parasitic infections, congenital malformations of the biliary tract and primary sclerosing cholangitis, and sex [8]. However, most patients with BTC do not have any of the known risk factors apart from age [9]. Screening for and detection of risk factors in addition to geography and age could lead to significantly faster detection of BTC and a subsequent improvement in survival.

An estimated 20% of the global cancer burden can be attributed to infectious diseases [10]. The association between viral infections, such as human papilloma virus, hepatitis B and C and certain forms of cancer, has been well-established [11, 12]. This knowledge has led to the implementation of successful targeted treatment and screening programs that can facilitate prevention and early detection of these cancers and improve survival, such as the Dutch national program for cervical cancer [13]. Although less studied, bacteria also have oncogenic potential and thereby increase the risk of cancer [14]. The primary example is *Helicobacter pylori* infection, which increases the risk of gastric cancer through the secretion of toxins that mediate cell signaling, as well as chronic inflammation [15]. Similarly, *Salmonella* spp. enforce bacterial uptake by manipulating host cell signaling pathways. Specifically, host AKT and ERK pathways are activated. Both pathways are active in many cancers and are an essential step in the malignant transformation of pre-transformed cells [16]. *Salmonella* spp. infection is common and represents a known risk factor for gallbladder and colon cancers, with the former pertaining specifically to *Salmonella* typhi, the agent of typhoid fever, and the latter to non-typhoidal *Salmonella*e [16, 17]. However, the role of non-typhoidal *Salmonella* has not yet been investigated for other biliary cancers. *Campylobacter* spp. is another frequently-occurring gastrointestinal infection able to promote colon tumorigenesis by producing cytolethal distending toxins and is more frequently present in the microbiome of colorectal cancer patients, although a causal relationship between colorectal cancer and Campylobacter spp. infection has not been demonstrated [18-21].

Salmonella spp. is known to cause chronic inflammation of the bile ducts and to produce toxins with carcinogenic potential, which may lead to cancer of the extrahepatic biliary tract [22]. After an outbreak of *Salmonella* typhi in 1964, researchers found that the risk of biliary tract cancer was increased by 164 times in carriers compared to non-carriers [23]. Although non-typhoidal *Salmonella* has been associated with the development of colon cancer, its role has not been specifically investigated in biliary tract cancers other than gallbladder cancer [17]. *Campylobacter* spp. is found in abundance in the biliary microbiome of patients with BTC [24]. The potential association with non-typhoidal *Salmonella* or *Campylobacter* spp. and BTC has not been studied in large cohorts due to the rarity of BTC, especially in Western populations. In case an association is found, targeted screening for BTC in *Salmonella* spp. and *Campylobacter* spp. patients might be considered. To assess whether infection with non-typhoidal *Salmonella* or *Campylobacter* spp. is a risk factor for BTC, this study compares the incidence of BTC in patients with a registered non-typhoidal *Salmonella* or *Campylobacter* spp. infection in the past to the incidence of BTC in a Western–European population.

Results

Cohort characteristics

The final cohort consisted of 16,283 *Salmonella* spp. patients (reported between 1999–2016), 27,692 *Campylobacter* spp. patients (reported between 2002–2016) and 8506 patients with BTC (Figure 1). After linkage, nine *Salmonella* spp. patients and seven *Campylobacter* spp. patients were diagnosed with BTC ≥1 year after infection.

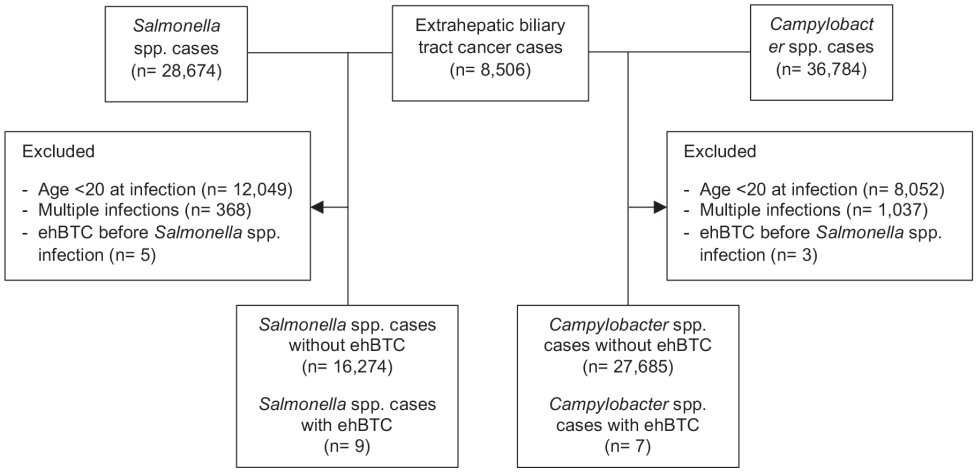


Figure 1. Cohort selection. ehBTC = extrahepatic biliary tract cancer.

Baseline characteristics of the cohorts are provided in Table 1 (salmonellosis and campylobacteriosis) and Table 2 (BTC). Median age at infection was 48.9 years (IQR: 30.0–66.0) for salmonellosis patients (66.5%, <60 years) and 48.5 years (IQR: 31.3–62.4) for campylobacteriosis patients (70.8%, <60 years). Median follow-up after infection was 7 years (IQR 3–12) in salmonellosis patients and 5 years (IQR 3–9) in campylobacteriosis patients. Median age at diagnosis was 73 years (IQR 64–80) in BTC patients. Median follow-up time from diagnosis to death or end of study in in BTC patients was 55 months.

Table 1. Baseline characteristics of the salmonellosis and campylobacteriosis patients.

	Salmonella	**Campylobacter**
Characteristic	N (%)	N (%)
Age		
<40	6167 (38%)	10,125 (37%)
40 – 49	2190 (13%)	4499 (16%)
50 – 59	2445 (15%)	4985 (18%)
60 – 69	2273 (14%)	4153 (15%)
70 - 79	2003 (12%)	2703 (10%)
80+	1172 (7%)	1227 (4%)
Sex		
Male	7640 (47%)	14,293 (52%)
Female	8612 (53%)	13,399 (48%)
Serotype/species		
(monophasic) *S*. Typhimurium	4487 (28%)	[1]
S. Enteritidis	5544 (34%)	[1]
S. (Para)Typhi	318 (2%)	[1]
Other *Salmonella* serotypes	5903 (36%)	[1]
C. jejuni	[2]	23,647 (85%)
C. coli	[2]	1910 (7%)
Other *Campylobacter* species	[2]	2135 (8%)
Type of infection		
Septicemic	884 (5%)	[3]
Enteric	13,864 (88%)	[3]
Other	1066 (7%)	[3]

[1] Not applicable for *Campylobacter.* [2] Not applicable for *Salmonella.* [3] Not registered for campylobacteriosis cases.

Table 2. Baseline characteristics of patients with BTC in the Netherlands (2000–2017).

Characteristic	N (%) / Median (95% CI)
Age	
<40	99 (1%)
40 – 49	326 (4%)
50 – 59	993 (12%)
60 – 69	2048 (24%)
70 - 79	2851 (33%)
80+	2189 (26%)
Sex	
Male	3822 (45%)
Female	4684 (55%)
Tumor location	
Gallbladder	2586 (30%)
Bile ducts, NOS	2421 (28%)
Proximal bile ducts	1846 (22%)
Distal bile ducts	1490 (18%)
Other [1]	163 (2%)
Clinical stage	
Non-metastatic	4078 (48%)
Metastatic	4428 (52%)
Treatment	
Resection	2479 (29%)
Chemotherapy	647 (8%)
Survival (months)	5.8 (5.5-6.0)

[1] Includes cystic duct and mixed types.

Patients with *Salmonella* spp. infection and BTC

Nine salmonellosis patients were diagnosed with BTC ≥1 year after salmonellosis diagnosis (Table 3). Mean time to BTC diagnosis was 47 months (range 13–81). Three of nine (33%) salmonellosis patients were ≤50 years old at time of BTC diagnosis, as opposed to the general BTC population, in which only 5.0% of patients were ≤50 years old at time of BTC diagnosis ($p < 0.001$). Four cases were diagnosed with *S*. enteritidis, three with *S*. typhimurium, and two with other *Salmonella* serovars. Eight patients had an enteric infection, one had an invasive (bloodstream) infection. Two patients had a distal cholangiocarcinoma, one patient had a proximal cholangiocarcinoma, one patient had gallbladder cancer, and five had BTC NOS (not otherwise specified).

Table 3. Baseline characteristics of patients with *Salmonella* infection and extrahepatic biliary tract cancer.

Characteristic	**N (%)**
Sex (male)	5 (56%)
Age	
<60	3 (33%)
≥60	6 (67%)
Serotype	
Enteritidis	4 (45%)
(monophasic) Typhimurium	2 (22%)
Other	3 (33%)
Interval	
<60 months	7 (78%)
≥60 months	2 (22%)
Tumor location	
Gallbladder / proximal bile ducts	2 (22%)
Distal bile ducts	2 (22%)
Extrahepatic bile ducts, NOS	5 (56%)
Base of diagnosis	
Cytology / imaging	6 (67%)
Histology	3 (33%)

Patients with *Campylobacter* spp. infection and BTC

Seven campylobacteriosis patients were diagnosed with BTC ≥1 year after diagnosis (Table 4). Mean time to BTC diagnosis was 60.6 months (range 18–138). All patients were >50 years of age at time of BTC diagnosis. Five patients had a proximal cholangiocarcinoma and two patients had gallbladder cancer.

Table 4. Baseline characteristics of patients with *Campylobacter* spp. infection and biliary tract cancer.

Characteristic	**N (%)**
Sex (male)	3 (43%)
Age	
<60	2 (29%)
≥60	5 (71%)
Interval	
<60 months	3 (43%)
≥60 months	4 (57%)
Tumor location	
Gallbladder / proximal bile ducts	[1]
Distal bile ducts	[1]
Extrahepatic bile ducts, NOS	0 (0%)
Base of diagnosis	
Cytology / imaging	2 (29%)
Histology	5 (71%)

[1] Numbers cannot be provided due to risk of subject identification.

Risk of BTC After *Salmonella* spp. or *Campylobacter* spp. infection

The SIR of BTC among the salmonellosis patients (compared to the general population) was 1.53 (95% CI 0.70–2.91, Table 5) and the absolute risk was 0.05%. Subgroup analysis in patients <60 years of age demonstrated that the SIR in this group was 1.72 (CI 0.36–5.04). Subgroup analysis according to gender revealed similar findings. In campylobacteriosis patients, the SIR was 0.97 (95% CI 0.39–2.00, Table 5) and the absolute risk was 0.03%. Subgroup analyses stratified according to gender and age revealed similar results.

Table 5. Incidence of biliary tract cancer in patients ≥1 year after laboratory confirmed infection with *Salmonella* spp. or *Campylobacter* spp., stratified by age at infection and gender.

	Observed incidence	Expected incidence	SIR	95% CI	p-value
Salmonella spp.					
All patients	9	5.875	1.53	0.70-2.91	0.280
20-60	3	1.740	1.72	0.36-5.04	0.507
Male	5	2.665	1.88	0.61-4.38	0.264
Female	4	3.289	1.22	0.33-3.11	0.835
Campylobacter spp.					
All patients	7	7.221	0.97	0.39-2.00	0.868
20-60	2	2.126	0.94	0.11-3.40	0.715
Male	3	4.025	0.75	0.15-2.18	0.857
Female	4	3.233	1.24	0.34-3.17	0.810

SIR: standardized incidence ratio.

Discussion

This study assessed whether *Salmonella* spp. or *Campylobacter* spp. infection represents a significant risk factor for BTC by comparing the incidence of BTC in patients with a history of *Salmonella* spp. or *Campylobacter* spp. infection to the (age-, gender- and calendar year-matched) incidence of BTC in the general Dutch population. Additionally, age and gender effects on the association between *Salmonella* spp. or *Campylobacter* spp. infection and BTC were investigated. No significant increase in BTC occurrence in patients who had experienced a severe *Salmonella* spp. or *Campylobacter* spp. infection was observed.

The relatively low number of *Salmonella* spp. (and *Campylobacter* spp.) infections linked to the (already rare) BTC patients found in this study was the main limitation for statistical significance, as considerable uncertainty was introduced in the estimates by such low number of outcome events. The upper limit of the SIR for BTC in salmonellosis patients was 2.7, which implies that a clinically significant effect may be present, but the study is simply insufficiently powered to detect its presence. This issue is, however, not unique to this study alone, but rather affects all studies investigating rare diseases. Experts increasingly recognize that some evidence, although maybe imprecise, may be better than no evidence at all [25].

In countries where typhoid fever is still endemic, such as the Indian subcontinent and some parts of South America, multiple epidemiological studies have shown an increased risk

for the development of BTC and especially gallbladder cancer. Besides chronic infection, an increased risk of gallstones in these populations, a higher incidence of obesity, and potential environmental pollution have been mentioned as potentially contributing to this phenomenon [22]. However, none of these factors (apart from gallstones and gallbladder cancer, which is not unique to these countries) show an extremely high correlation with the incidence of BTC. On the other hand, researchers have demonstrated a clear association with chronic *S*. typhi infection and the development of gallbladder cancer in these countries [26]. In contrast, a Chinese study investigating the correlation between chronic infection with *S*. typhi and biliary tract cancer failed to find a significant association due to a very low occurrence of such infection [27]. One may argue that association does not equal causation and that in areas with endemic typhoid fever and high rates of gallbladder cancer, other factors might be at play as well. However, even in Western countries with typically extremely low incidence of *S*. typhi infection (as typhoid fever has been eradicated in most Western countries thanks to modern sanitation), after large outbreaks of typhoid fever, an increase in number of BTC diagnoses was observed [23].

This paper focusses primarily on the incidence of BTC in non-typhoidal *Salmonella*. We hypothesized that, similar to gallbladder cancer, the increased incidence of BTC after typhoidal *Salmonella* infection would translate to increased BTC risk in non-typhoidal *Salmonella* [28]. The lack of significant correlation in non-typhoid *Salmonella* infection may be attributed to the fact that non-typhoid *Salmonella* strains are less likely to cause chronic infection and thus have lower oncogenic potential compared to their typhoid counterparts [29].

Remarkably, one third of the patients with both *Salmonella* spp. infection and BTC were under 50 years of age at time of BTC diagnosis. This proportion was significantly higher than in the general BTC population, in which only 5% is aged 50 years or younger [30]. Because the risk of BTC increases exponentially with age, we performed a subgroup analysis in all patients aged <60. Although this subgroup analysis also failed to reach significance due to the even lower numbers, the relatively high proportion of young patients suggests that *Salmonella* spp. infection at a young age might contribute to the risk of developing BTC later in life. Possibly, patients who acquire a *Salmonella* spp. infection at the age of 70 or older may die from other diseases before they develop BTC and are thus less well-represented. The median time between *Salmonella* spp. infection and BTC diagnosis was 4 years. This finding implies that the potential oncogenic effect of *Salmonella* spp. results in malignant transformation of epithelial cells in a relatively short timeframe and is concurrent with other studies [17]. Another explanation may be that patients with inflammatory bowel

disease (IBD) are at a higher risk for developing a serious *Salmonella* spp. infection. Since IBD often has an onset in early adulthood and is also a potential independent risk factor for the development of BTC, it is possible that this difference in age can be explained by the fact that the patients with *Salmonella* spp. infection also had IBD and therefore were at greater risk for developing BTC at a younger age [31].

No tendency towards increased BTC incidence after *Campylobacter* spp. infection was seen in this study. *Campylobacter* spp. and *Salmonella* spp. bacteria both release the genotoxic protein cytolethal distending toxin (CDT). However, whereas *Salmonella* spp. is linked to the development of BTC by overexpression of c-MYC in tissue samples, *Campylobacter* spp. is not [16]. Differences in bacterial mechanisms, specifically concerning the alteration of host cell signaling pathways during invasion, may account for differences in oncogenic potential between the two species.

Molecular characterization of cancers and subsequent personalization of therapy is a prime topic in current oncological research. Although the genomic landscape of BTC is incredibly diverse, multiple preclinical and clinical models show that BTC development may be associated with the alteration of several actionable genes. A particular example is the overexpression of cyclophilin-A in patients with liver-fluke-associated cholangiocarcinoma [32,33]. Identification of inflammation-associated driving mutations is an important topic as it has implications for both risk profiling and potential personalized treatment. Although molecular profiling of patients with salmonellosis and BTC was outside of the scope of this study, a study in gallbladder cancer has managed to identify the signaling pathway associated with *S*. typhi development and gallbladder cancer [16]. Further research investigating molecular alterations in infected cancer patients is paramount to increase our understanding of tumor cell transformation and cancer development.

The primary limitation of this study is the low number of *Salmonella* spp. and *Campylobacter* spp. infected patients that also developed BTC, leading to a high risk of type-2 error. Typically, patients with *Salmonella* spp. infection in the Netherlands who require medical attention, laboratory diagnosis and reporting to health authorities are severely ill. As most patients with *Salmonella* spp. infection only show mild symptoms, the actual number of *Salmonella* spp. cases in the Netherlands is much higher than reported. It is estimated that close to 1 million inhabitants developed a symptomatic *Salmonella* spp. infection in the Netherlands between 1999–2015, which is 35 times the number of cases included in this study. Campylobacteriosis cases are estimated around 81,000 in the Netherlands annually [34]. As a result, a number of patients with mild and therefore unreported *Salmonella* spp.

or *Campylobacter* spp. infection, but with a BTC diagnosis, may have been misclassified and included in the group of BTC patients without (reported) *Salmonella* spp. or *Campylobacter* spp. infection. Since the contribution of these mild infections to the risk of developing BTC is implicitly included in the baseline risk, our results may be considered as very conservative estimates of their true contribution to BTC risk. Moreover, although chronic infections are those mostly implicated in BTC formation, they could not be studied as such in this study because this information (i.e., differentiation between acute and chronic infection) is simply not available in the RIVM data set [35]. Yet, we included all reported infections, and because these infections represent the most severe ones (in terms of magnitude and duration of symptoms) occurring in the population, our analysis implicitly focused on a selection of salmonellosis and campylobacteriosis patients that showed extreme clinical manifestations. Finally, the RIVM registry only contains data on *Salmonella* spp. and *Campylobacter* spp. infection from 1999 onwards and consequently we only had a median follow-up period of 7 years. If, like in pancreatic cancer, the interval between first mutation and cancer development is over 10 years, the study period may have been insufficient to detect a correlation between infection and BTC development [36].

A major strength of this study is the cohort size and nation-wide design. Indeed, it should be acknowledged that the low number of BTC events in our cohort—despite the large surveillance data sets used—reflects mainly the rare occurrence of these tumors. The cohort analyzed in this paper is large and comprehensive, being nation-wide and covering all available years of systematic data collection. Previous studies investigating the role of bacterial infections in the development of BTC have typically drawn from case-control cohorts or small case series. Additionally, to our knowledge, this paper describes the first Western cohort of patients with *Salmonella* spp. or *Campylobacter* spp. infection and BTC [26].

Materials and methods

Data collection and linkage

Analyses were based on three linked health registries with national coverage. The first registry contains records from laboratory-confirmed human infections with *Salmonella* spp. (from 1999 onwards) and *Campylobacter* spp. (from 2002 onwards) based on the national laboratory surveillance system for gastrointestinal pathogens coordinated by the Dutch National Institute for Public Health and the Environment (RIVM) [35]. The surveillance system has an estimated coverage of the resident Dutch population of 64% for *Salmonella*

spp. and 52% for *Campylobacter* spp. infection [37]. The second registry consisted of histopathological records provided by the automated pathological archive, the nation-wide network of histopathology and cytology in the Netherlands (PALGA) [38]. The third registry was the Netherlands Cancer Registry (NCR) [39], which contains data on all newly diagnosed malignancies since 1989, covering around 95% of the Dutch population [39]. The NCR is updated through PALGA and supplemented annually by information from hospital discharge records. Statistics Netherlands (CBS, www.cbs.nl) acted as a trusted third party to anonymize and link the data sets. The CBS used the date of birth, gender and six digit postal code, which were available in all three registries, to generate a unique personal identifier (Record Identification Number (RIN)). After the RIN was generated, all personally identifying data was removed from the data sets. The researchers used the RIN to link all three data sets. Ethical Approvals were given by the CMO Arnhem-Nijmegen, code: 2017-3912 in 18 December 2017. A waiver of informed consent was provided, no informed consent form was used.

Patient Selection and Variable Definitions

All patients aged ≥20 years of age with a diagnosed non-typhoidal *Salmonella* infection from the 1 January 1999, and with a diagnosed *Campylobacter* spp. infection from the 1 January 2002, until the 31 December 2016 were identified in the RIVM database. Additionally, all patients with non-intrahepatic biliary tract cancer (ICD-O-3 location codes C239, C240, C242, C243, C244, C248, C249) were identified in the NCR database. Patients who were diagnosed with intrahepatic BTC, BTC before or within 1 year of salmonellosis/campylobacteriosis diagnosis or had less than 1 year of follow-up were excluded. In case the patient had multiple recorded *Salmonella* spp./*Campylobacter* spp. infections, only the first diagnosis was considered. Both databases were cleared from duplicates. Time at risk was defined as the number of days between 1 year after salmonellosis/campylobacteriosis diagnosis and development of BTC, death, or end of the study period (31 December 2017), whichever occurred first.

Outcomes

The primary outcome of the study was the incidence of BTC among individuals with a registered non-typhoida l*Salmonella* or *Campylobacter* spp. infection in the past as compared to the incidence of BTC in the general Dutch population. Subgroup analyses were conducted to investigate the risk of BTC in patients ≤60 years of age (at the time of infection) and by gender.

Statistical Analysis

Standardized incidence ratios (SIR) were calculated for salmonellosis and campylobacteriosis patients separately to compare the difference in incidence of BTC in patients with *Salmonella* spp. or *Campylobacter* spp. infection to an age-, gender- and calendar year-matched cohort of the general Dutch population. To this end, the observed number of BTC cases in the salmonellosis and campylobacteriosis patients was divided by the expected number of BTC cases in the matched cohort provided by the NCR. 95% confidence intervals (95% CI) for the SIRs were calculated assuming a Poisson distribution. In all analyses, p-values < 0.05 were considered statistically significant. Statistical analysis was performed using STATA version 14 (StataCorp. 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP).

Conclusions

There is accumulating evidence that pathogenic bacteria like *Salmonella* spp. play a role in cancer development, including cancers of the digestive system. However, we could not demonstrate a significantly increased occurrence of BTC among reported salmonellosis or campylobacteriosis patients as compared to the general population. Potentially, the study was either underpowered due to the low number of BTC events or *Salmonella* and *Campylobacter* spp. infections are not associated with the development of BTC in Western countries. Additional research is needed to unravel the biological mechanisms behind bacterial infections as a cause of cancer and identify potential infections that may warrant early screening and therefore facilitate early cancer detection, especially in third-world countries with high rates of (hyper)endemic bacterial infections.

Author contributions

Conceptualization, E.d.S.L., J.D., R.v.d.P., P.d.R. and L.M.G.; methodology, E.d.S.L., J.D., P.d.R., L.M.G., E.F.; software, J.D., L.M.G.; validation, J.D., L.M.G.; formal analysis, J.D., E.d.S.L.; investigation, E.d.S.L., J.D.; resources, L.M.G., E.F.; data curation, E.d.S.L., J.D., L.M.G.; writing—original draft preparation, E.d.S.L., B.G.K., P.d.R., L.M.G.; writing—review and editing, E.d.S.L., J.D., B.G.K., L.M.G., R.v.d.P., P.d.R., E.F.; visualization, E.d.S.L.; supervision, P.d.R., L.M.G., E.F., B.G.K.; project administration, E.d.S.L.; funding acquisition, P.d.R., R.v.d.P. All authors have read and agreed to the published version of the manuscript.

Funding

This study was supported by a grant from the ADP foundation (Chamber of Commerce number 55695817). The founders did not play a role in study design, data collection, data interpretation, data analysis, the writing of the manuscript or the decision to submit the article for publication.

Conflicts of interest

The authors declare no conflict of interest.

References

1. Surveillance EaERp. Biliary Tract Cancer, Survival by Stage 2018. Available online: https://www.cancer.org/cancer/bile-duct-cancer/detection-diagnosis-staging/survival-by-stage.html (accessed on 1 March 2020).
2. Ries, L.A.G.; Reichman, M.E.; Lewis, D.R.; Hankey, B.F.; Edwards, B.K. Cancer Survival and Incidence from the Surveillance, Epidemiology, and End Results (SEER) Program. Oncologist 2003, 8, 541–552.
3. Cronin, K.A.; Bs, A.J.L.; Scott, S.; Sherman, R.L.; Noone, A.-M.; Ms, N.H.; Henley, S.J.; Anderson, R.N.; Bs, A.U.F.; Ma, J.; et al. Annual Report to the Nation on the Status of Cancer, part I: National cancer statistics. Cancer 2018, 124, 2785–2800.
4. Shroff, R.T.; Kennedy, E.B.; Bachini, M.; Bekaii-Saab, T.; Crane, C.; Edeline, J.; El-Khoueiry, A.; Feng, M.; Katz, M.H.; Primrose, J.; et al. Adjuvant Therapy for Resected Biliary Tract Cancer: ASCO Clinical Practice Guideline. J. Clin. Oncol. 2019, 37, 1015–1027.
5. Public Health Agency of Canada; Statistics Canada; Canadian Cancer Society. Release notice-Canadian Cancer Statistics 2019. Health Promot. Chronic Dis. Prev. Can. 2019, 39, 255.
6. Are, C.; Ahmad, H.; Ravipati, A.; Croo, D.; Clarey, D.; Smith, L.; Price, R.R.; Butte, J.M.; Gupta, S.; Chaturvedi, A.; et al. Global epidemiological trends and variations in the burden of gallbladder cancer. J. Surg. Oncol. 2017, 115, 580–590.
7. Lohman, E.D.S.; De Bitter, T.; Verhoeven, R.H.; Van Der Geest, L.; Hagendoorn, J.; Mohammad, N.H.; Daams, F.; Klümpen, H.-J.; Van Gulik, T.; Erdmann, J.I.; et al. Trends in Treatment and Survival of Gallbladder Cancer in the Netherlands; Identifying Gaps and Opportunities from a Nation-Wide Cohort. Cancers 2020, 12, 918.
8. Tyson, G.L.; El-Serag, H.B. Risk factors for cholangiocarcinoma. Hepatology 2011, 54, 173–184.
9. Gatto, M.; Bragazzi, M.C.; Semeraro, R.; Napoli, C.; Gentile, R.; Torrice, A.; Gaudio, E.; Alvaro, D. Cholangiocarcinoma: Update and future perspectives. Dig. Liver Dis. 2010, 42, 253–260.
10. Schottenfeld, D.; Beebe-Dimmer, J. The cancer burden attributable to biologic agents. Ann. Epidemiol. 2015, 25, 183–187.
11. Wardak, S. Human Papillomavirus (HPV) and cervical cancer. Medycyna Doswiadczalna i Mikrobiologia 2016, 68, 73–84.
12. Balogh, J.; Victor, D.; Asham, E.H.; Burroughs, S.G.; Boktour, M.; Saharia, A.; Li, X.; Ghobrial, R.M.; Monsour, H.P. Hepatocellular carcinoma: A review. J. Hepatocell. Carcinoma 2016, 3, 41–53.
13. Aitken, C.A.; Van Agt, H.M.E.; Siebers, A.G.; Van Kemenade, F.J.; Niesters, H.G.M.; Melchers, W.J.G.; Vedder, J.E.M.; Schuurman, R.; Brule, A.J.C.V.D.; Van Der Linden, H.C.; et al. Introduction of primary screening using high-risk HPV DNA detection in the Dutch cervical cancer screening programme: A population-based cohort study. BMC Med. 2019, 17, 1–14.
14. Gagnaire, A.; Nadel, B.; Raoult, D.; Neefjes, J.; Gorvel, A.G.B.N.J.-P. Collateral damage: Insights into bacterial mechanisms that predispose host cells to cancer. Nat. Rev. Genet. 2017, 15, 109–128.
15. Samaras, V.; Rafailidis, P.I.; Mourtzoukou, E.G.; Peppas, G.; Falagas, M.E. Chronic bacterial and parasitic infections and cancer: A review. J. Infect. Dev. Countries 2010, 4, 267–281.
16. Scanu, T.; Spaapen, R.M.; Bakker, J.M.; Pratap, C.B.; Wu, L.-E.; Hofland, I.; Broeks, A.; Shukla, V.K.; Kumar, M.; Janssen, H.; et al. Salmonella Manipulation of Host Signaling Pathways Provokes Cellular Transformation Associated with Gallbladder Carcinoma. Cell Host Microbe 2015, 17, 763–774.

17. Mughini-Gras, L.; Schaapveld, M.; Kramers, J.; Mooij, S.; Neefjes-Borst, E.A.; Van Pelt, W.; Neefjes, J. Increased colon cancer risk after severe Salmonella infection. PLoS ONE 2018, 13, e0189721.
18. Brauner, A.; Brandt, L.; Frisan, T.; Thelestam, M.; Ekbom, A. Is there a risk of cancer development after Campylobacter infection? Scand. J. Gastroenterol. 2010, 45, 893–897.
19. Lara-Tejero, M. A Bacterial Toxin that Causes DNA Damage to Modulate Cellular Responses. Sci. World J. 2001, 1, 190–191.
20. He, Z.; Gharaibeh, R.Z.; Newsome, R.C.; Pope, J.L.; Dougherty, M.W.; Tomkovich, S.; Pons, B.; Mirey, G.; Vignard, J.; Hendrixson, D.R.; et al. Campylobacter jejuni promotes colorectal tumorigenesis through the action of cytolethal distending toxin. Gut 2018, 68, 289–300.
21. Wu, N.; Yang, X.; Zhang, R.; Li, J.; Xiao, X.; Hu, Y.; Chen, Y.; Yang, F.; Lu, N.; Wang, Z.; et al. Dysbiosis Signature of Fecal Microbiota in Colorectal Cancer Patients. Microb. Ecol. 2013, 66, 462–470.
22. Marcano-Bonilla, L.; Mohamed, E.A.; Mounajjed, T.; Roberts, L.R. Biliary tract cancers: Epidemiology, molecular pathogenesis and genetic risk associations. Chin. Clin. Oncol. 2016, 5, 61.
23. Caygill, C.; Hill, M.; Braddick, M.; Sharp, J. Cancer mortality in chronic typhoid and paratyphoid carriers. Lancet 1994, 343, 83–84.
24. Avilés-Jiménez, F.; Guitron, A.; Segura-López, F.; Méndez-Tenorio, A.; Iwai, S.; Hernández-Guerrero, A.; Torres, J. Microbiota studies in the bile duct strongly suggest a role for Helicobacter pylori in extrahepatic cholangiocarcinoma. Clin. Microbiol. Infect. 2016, 22, 178.e11–178.e22.
25. Edwards, S.; Lilford, R.J.; Braunholtz, D.; Jackson, J. Why "underpowered" trials are not necessarily unethical. Lancet 1997, 350, 804–807.
26. Nagaraja, V.; Eslick, G.D. Systematic review with meta-analysis: The relationship between chronic Salmonella typhi carrier status and gall-bladder cancer. Aliment. Pharmacol. Ther. 2014, 39, 745–750.
27. Safaeian, M.; Gao, Y.-T.; Sakoda, L.C.; Quraishi, S.M.; Rashid, A.; Wang, B.-S.; Chen, J.; Pruckler, J.; Mintz, E.; Hsing, A.W. Chronic typhoid infection and the risk of biliary tract cancer and stones in Shanghai, China. Infect. Agents Cancer 2011, 6, 6.
28. Iyer, P.; Barreto, S.G.; Sahoo, B.; Chandrani, P.; Ramadwar, M.R.; Shrikhande, S.V.; Dutt, A. Non-typhoidal Salmonella DNA traces in gallbladder cancer. Infect. Agents Cancer 2016, 11, 1–4.
29. Egal-Mor, O.; Boyle, E.C.; Grassl, G.A. Same species, different diseases: How and why typhoidal and non-typhoidal Salmonella enterica serovars differ. Front. Microbiol. 2014, 5, 391.
30. Integraal Kankercentrum Nederland: Cijfers over Kanker. Available online: https://www.iknl.nl/nkr-cijfers (accessed on 1 March 2020).
31. Huai, J.-P.; Ding, J.; Ye, X.-H.; Chen, Y.-P. Inflammatory bowel disease and risk of cholangiocarcinoma: Evidence from a meta-analysis of population-based studies. Asian Pac. J. Cancer Prev. 2014, 15, 3477–3482.
32. Massa, A.; Varamo, C.; Vita, F.; Tavolari, S.; Peraldo-Neia, C.; Brandi, G.; Rizzo, A.; Cavalloni, G.; Aglietta, M. Evolution of the Experimental Models of Cholangiocarcinoma. Cancers 2020, 12, 2308.
33. Rizzo, A.; Ricci, A.D.; Tober, N.; Nigro, M.C.; Mosca, M.; Palloni, A.; Abbati, F.; Frega, G.; De Lorenzo, S.; Tavolari, S.; et al. Second-line Treatment in Advanced Biliary Tract Cancer: Today and Tomorrow. Anticancer Res. 2020, 40, 3013–3030.
34. Havelaar, A.H.; Ivarsson, S.; Löfdahl, M.; Nauta, M. Estimating the true incidence of campylobacteriosis and salmonellosis in the European Union, 2009. Epidemiol. Infect. 2012, 141, 293–302.

35. Friesema, I.; Koppeschaar, C.; Donker, G.; Dijkstra, F.; Van Noort, S.; Smallenburg, R.; Van Der Hoek, W.; Van Der Sande, M. Internet-based monitoring of influenza-like illness in the general population: Experience of five influenza seasons in the Netherlands. Vaccine 2009, 27, 6353–6357.
36. A Iacobuzio-Donahue, C. Genetic evolution of pancreatic cancer: Lessons learnt from the pancreatic cancer genome sequencing project. Gut 2012, 61, 1085–1094.
37. Van Pelt, W.; De Wit, M.A.S.; Wannet, W.J.B.; Ligtvoet, E.J.J.; Widdowson, M.A.; Van Duynhoven, Y.T.H.P. Laboratory surveillance of bacterial gastroenteric pathogens in The Netherlands, 1991–2001. Epidemiol. Infect. 2003, 130, 431–441.
38. Casparie, M.; Tiebosch, A.T.M.G.; Burger, G.; Blauwgeers, H.; Van De Pol, A.; Van Krieken, J.H.J.M.; A Meijer, G. Pathology Databanking and Biobanking in The Netherlands, a Central Role for PALGA, the Nationwide Histopathology and Cytopathology Data Network and Archive. Cell. Oncol. 2007, 29, 19–24.
39. Van Der Willik, K.D.; Ruiter, R.; Van Rooij, F.J.; Heemst, J.V.; Hogewoning, S.J.; Timmermans, K.C.; Visser, O.; Schagen, S.B.; Ikram, M.A.; Stricker, B.H. Ascertainment of cancer in longitudinal research: The concordance between the Rotterdam Study and the Netherlands Cancer Registry. Int. J. Cancer 2019.

Chapter 7

BACTERIAL AND PARASITIC PATHOGENS AS RISK FACTORS FOR CANCERS IN THE GASTROINTESTINAL TRACT: A REVIEW OF CURRENT EPIDEMIOLOGICAL KNOWLEDGE

J.W. Duijster
E. Franz
J.J.C. Neefjes
L. Mughini-Gras

Published in: *Frontiers in Microbiology,* 2021(12), 3746

Abstract

The oncogenic potential of viral infections is well established and documented for many years already. However, the contribution of (commensal) bacteria and parasites to the development and progression of cancers has only recently gained momentum, resulting in a rapid growth of publications on the topic. Indeed, various bacteria and parasites have been suggested to play a role in the development of gastrointestinal cancer in particular. Therefore, an overview of the current epidemiological knowledge on the association between infections with bacteria and parasites and cancers of the gastrointestinal tract is needed. In this review, we summarized the methodological characteristics and main results of epidemiological studies investigating the association of 10 different bacteria (*Bacteroides fragilis*, *Campylobacter* spp., *Clostridium* spp., *Enterococcus faecalis*, *Escherichia coli*, *Fusobacterium nucleatum*, *Porphyromonas gingivalis*, nontyphoidal *Salmonella*, *Salmonella* Typhi, and *Streptococcus* spp.) and three parasites (*Cryptosporidium* spp., *Schistosoma* spp. and *Strongyloides stercoralis*) with gastrointestinal cancer. While the large body of studies based on microbiome sequencing provides valuable insights into the relative abundance of different bacterial taxa in cancer patients as compared to individuals with pre-malignant conditions or healthy controls, more research is needed to fulfil Koch's postulates, possibly making use of follow-up data, to assess the complex role of bacterial and parasitic infections in cancer epidemiology. Studies incorporating follow-up time between detection of the bacterium or parasite and cancer diagnosis remain valuable as these allow for estimation of cause-effect relationships.

Introduction

In 2020, an estimated number of over 5 million new cancers of the gastrointestinal (GI) tract were diagnosed globally [1]. For most of these cancers, the incidence is increasing, mostly as a result of increasing age and welfare characterized by factors such as changing diets and more sedentary lifestyles. Apart from the major risk factors for GI tract cancer that include genetics, age, smoking, alcohol consumption, obesity and exposure to radiation/chemicals, a potentially carcinogenic role of microorganisms is gaining momentum. Particularly, the rapid evolution of high throughput sequencing as a tool to identify/quantify the composition of the human microbiome, has led to accumulating indications for a role of commensal bacteria such as *Escherichia coli* and *Fusobacterium nucleatum* in cancer development [2]. Several mechanisms have been described by which bacteria contribute to cancer development, including induction of DNA damage by toxins and manipulation of host cell signaling pathways, thereby affecting cell proliferation, differentiation, apoptosis and immune signaling [3]. Concerning the gastrointestinal tract, the association is best established (both mechanistically and epidemiologically) for *Helicobacter pylori* as causative agent of gastric cancer (GC) and Mucosa-Associated Lymphoid Tissue (MALT) lymphoma [4, 5]. Yet, while substantial laboratory evidence already exists for the role of several pathogens in cancer development, epidemiological data on of the broader (potential) role of (opportunistic) pathogens in GI cancer development is generally dispersed and unclear.

This review paper aims to provide an overview of the current epidemiological knowledge regarding the association between bacteria and parasites and cancers in the GI tract. To this end, we reviewed epidemiological studies reporting on an association between bacterial or parasitic gastrointestinal infections and GI cancers to identify (in)consistencies in their results also in relation to the different study designs.

Methods

We systematically searched PubMed, Embase and Web of Science for articles published since 1966, 1946, and 1988 respectively until April 2021. Details about the search strategy as well as inclusion and exclusion criteria are listed in Supplementary Tables 1, 2. The search was conducted using every possible combination of keywords listed in the categories in Supplementary Table 1, including key words related to malignancies in the GI tract, microorganisms and study design or measurement indicators. We included case-control studies, cohort studies and cross-sectional studies making use of surveillance data (e.g. bacterial infection records), serological assays or data about presence/abundance of

microbial genetic material in human specimens (e.g. tumor tissue, blood, feces) in relation to GI cancer. The outcome of interest comprised all primary malignancies of the GI, including esophagus, stomach, small intestine, duodenum, colon, rectum, anus, liver, intrahepatic bile ducts, biliary tract, gallbladder and pancreas. Viruses were not included in this review. Given that the association between *H. pylori* and gastric cancer is already extensively reviewed and meta-analyzed, this was excluded from this review. For the same reason, the relation between the microbiome composition at bacterial phylum- or family-level in and GI cancer (i.e. those studies addressing commensal bacterial phyla or families rather than bacterial species) was excluded. Finally, also the relation between the presence or abundance of specific microorganisms during or post-cancer treatment, as well as experimental (*in vivo* or *in vitro*) studies were excluded (Supplementary Table 2). This primary search was supplemented by a search in MedRvix, Google Scholar and Google for pre-prints of articles and conference abstracts, using the aforementioned inclusion and exclusion criteria (Supplementary Table 2). To ensure literature saturation, we scanned the reference lists of included studies or relevant reviews identified through the search. The search results were exported and unduplicated by EndNote. PROSPERO was screened for ongoing or recently completed systematic reviews about this topic.

In the primary eligibility screening, titles and abstracts were screened against the inclusion criteria, subsequently, (potential) relevant articles were screened based on full text reports. The following data items were extracted from the included articles: first author, year of publication, country/region of study, study period, microorganism and cancer(s) of interest, study design, population size or number of cases and controls, type of test(s) (e.g. serological, culture-confirmed infection, presence of bacterial DNA) and type of material tested (e.g. blood, tumor tissue, feces), measurement indicator(s) (e.g. odds ratios, hazard ratios) and main study outcomes. For each of the included studies, the primary outcome was the overall risk estimate of the association between infection and cancer. Further, subgroup estimates (e.g. stratified by gender, age group or follow-up time) were considered secondary outcomes. Characteristics and findings of the included studies were summarized in text and tables and (in)consistencies in study outcomes within and between microorganisms and malignancies were discussed, also in light of the study designs. The definition of search terms and inclusion and exclusion criteria was done by three reviewers (JD, LMG, EF). One reviewer (JD) performed the search, screening of abstracts an full-text articles and data extraction, whereas the process thereafter was performed by four reviewers (JD, LMG, EF, JN).

Results

We identified 4,826 articles by searching the electronic databases (Supplementary Figure 1). After exclusion of duplicates and ineligible articles, 229 articles remained for full-text screening, resulting in 91 included scientific articles. In addition, the manual search (in Google [Scholar]) and screening of the reference lists of eligible studies yielded another 65 scientific articles and two conference abstracts (Supplementary Figure 1). The 158 eligible articles/abstracts cover 13 different micro-organisms, including 10 bacteria and 3 parasites (Figure 1). Most studies had a case-control design (n=101), 33 were cohort studies (22 retrospective, 11 prospective), 23 were cross-sectional studies and 1 was a case series. The majority of the studies were published during the last decade (≥2011: n=116). Study characteristics and main outcomes of the reviewed studies are elaborated in the following paragraphs.

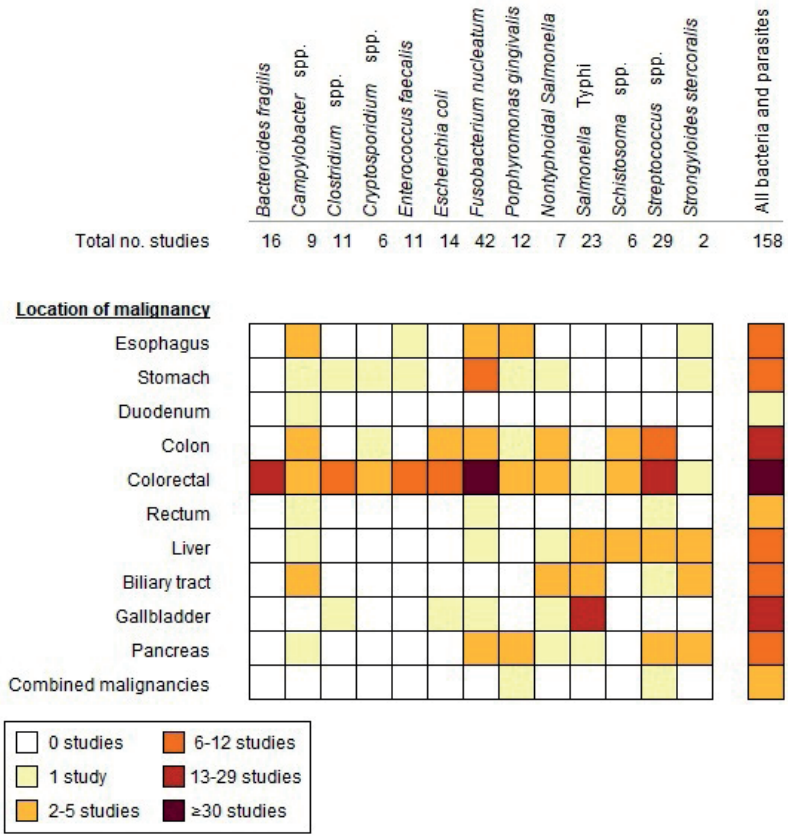


Figure 1. Number of studies included in the review for each bacterium/parasite in relation to the location of GI malignancies.

Bacteroides fragilis

Bacteroides fragilis is a commensal bacterium of the gut, subdivided into non-toxigenic *B. fragilis* (NTBF) and enterotoxigenic *B. fragilis* (ETBF). In contrast to the non-harmful NTBF, ETBF is associated with inflammatory bowel disease (IBD) and plays a role in colorectal carcinogenesis by producing pro-inflammatory cytokines and stimulating Wnt signaling [6]. Many studies report the presence or abundance of the *Bacteroides* genus (also including NTBF) in the human microbiome in relation to GI cancers. Supplementary Table 3 summarizes the characteristics and main results of 16 studies investigating the association between *B. fragilis* and colorectal cancer (CRC) [7-22]. Most (n=7) studies found an increased presence of *B. fragilis* and/or ETBF in fecal (n=3) or mucosa/tissue (n=4) samples from CRC patients versus healthy controls, albeit that the prevalences differed substantially between studies (Supplementary Table 3). One study found no significant difference in antibody titers against *B. fragilis* between CRC patients and healthy controls [11]. No studies were found assessing potential associations between antibody-levels against *B. fragilis*/ETBF and risk of cancer later in life (i.e. with follow-up time). In fact, one study reported similar antibody titers against *B. fragilis* in CRC patients as compared to healthy controls [11]. *B. fragilis* appeared also significantly enriched in precancerous conditions [8] and was significantly associated with late stages of CRC (Supplementary Table 3) [7, 12, 21]. Whilst the role of the gut microbiome (including the relative abundance of the *Bacteroides* phylum) in other GI cancers is gaining interest, no studies directly assessed the role of *B. fragilis* in GI cancers beyond CRC yet.

Campylobacter spp.

Within the genus *Campylobacter*, two zoonotic species, *Campylobacter jejuni* and *Campylobacter coli*, are the leading causes of gastroenteritis worldwide. A small portion of infections lead to long-term sequelae, including Guillain-Barré syndrome, reactive arthritis, and irritable bowel syndrome [23]. An experimental study in mice showed that infection with a human *C. jejuni* strain induced the development of colorectal tumors and changes in microbial composition through the action of cytolethal distending toxin (CDT). CDT causes double-stranded DNA breaks that may contribute to cancer formation [24]. In Supplementary Table 4 nine epidemiological studies are summarized which assess the association between *Campylobacter* and GI cancer [25-33]. In humans, *Campylobacter* spp. is significantly more abundant in the feces of CRC patients compared to healthy controls (Supplementary Table 4) [33]. Similarly, some studies observed significant higher abundance of *Campylobacter* spp. in tumor tissue versus adjacent normal tissue [25, 26] and tissue samples from cancer patients versus healthy controls (Supplementary Table 4) [31, 33]. Three cohort studies

assessed the association between *Campylobacter* infection and risk of colon cancer (CC), biliary tract cancer or multiple types of cancer later in life. All three studies used a follow-up time of more than 10 years (Supplementary Table 4) [28-30]. None of the studies found a significantly increased risk of digestive cancer after infection compared to the general population, all with standardized incidence ratios (SIRs) below or close to 1. Subgroup analysis in a Dutch cohort study revealed only a significant two-fold increased risk of CC when infected with *Campylobacter* jejuni or coli between 40-49 years of age [30]. Another species, *Campylobacter concisus*, is considered a human host-adapted species (not isolated from animals), which belongs to the commensal bacteria of the oral cavity and is associated with gingivitis, periodontitis, Barrett's esophagus (a complication of gastroesophageal reflux disease), gastroenteritis and IBD [34]. Recently, *C. concisus* has been linked to esophageal adenocarcinoma as this species was found to be significantly more abundant in patients with Barrett's esophagus, which is a precursor of esophageal cancer (Supplementary Table 4) [27, 35, 36].

Clostridium spp.

Clostridium is a genus of spore-forming bacteria with over 200 species. While the commensal species constitute 10-40% of the total gut microbiome, contributing to gut homeostasis from early infancy onwards, the bacteria are also a common cause of food-poisoning and nosocomial infections, particularly in immunocompromised patients [37, 38]. Chemotherapeutic cancer treatment alters the composition of the gut microbiome and damages the intestinal mucosa, favoring *C. difficile* colonization and the subsequent production of toxins (toxin A and B) which induce inflammation, tissue damage and cell death [39]. Whether *Clostridium* also contributes to cancer progression remains elusive. Supplementary Table 5 summarizes 11 studies focusing on *Clostridium* infection and cancer risk [9, 14, 17, 19, 40-46]. Significant increased risk of CRC was observed among individuals with a history of bacteremia with *Clostridium perfringens* or *Clostridium septicum* in a Chinese cohort. However, these numbers were low and the lag-time between bacteremia and cancer diagnosis was not reported [17]. Likewise, Justesen et al. found 22 incidences of CRC among 457 individuals with a history of bacteremia, most of them caused by *C. septicum* [9]. The majority of these cancers (n=20) were diagnosed within one year after the bacteremia (Supplementary Table 5) [9]. In another study, significant higher levels of *C. difficile* were detected in feces of CRC patients (prior to treatment) as compared to healthy controls (Supplementary Table 5) [14]. Similarly, *Clostridium hathewayi*, a relative unknown *Clostridium* species, has been proposed as potential biomarker for early detection of CRC

(rather than a causal agent), as it has been associated with overabundance in CRC patients (Supplementary Table 5) [41, 46].

Cryptosporidium spp.

Cryptosporidium spp. is a protozoan parasite causing enteric infections, generally presenting as a self-limiting watery diarrhea. Infection occurs mainly through the fecal-oral pathway or through consumption/ingestion of contaminated food and water. Over 90% of the infections are caused by the species *Cryptosporidium parvum* and *Cryptosporidium hominis* [47]. *Cryptosporidium* is known for its opportunistic behavior in immunocompromised patients (e.g. HIV-patients or individuals receiving oncological treatment) [48]. Experimental studies in mice suggest that *Cryptosporidium* infection can induce intestinal dysplasia [49, 50]. Supplementary Table 6 summarizes the characteristics of six studies assessing the association between *Cryptosporidium* and CRC (n=5), CC (n=1) or gastric cancer (n=1) in 4 different cohorts in Poland, Saudi, Lebanon and Tunisia [51-56]. In a recent meta-analysis summarizing the association between *Cryptosporidium* and cancer, most of the 19 included studies focused on all types of cancer combined (Supplementary Table 7) [48]. An over three-fold increased risk of *Cryptosporidium* infection was observed among cancer patients (all malignancies) as compared to non-cancer controls (OR 3.30; 95%CI 2.18-4.98), whereas for the digestive cancers only a site-specific estimate for CRC was given (OR 3.70; 95%CI 2.10-6.50), based on 4 studies (Supplementary Tables 6, 7) [48, 51, 53-55].

Enterococcus faecalis

Enterococcus faecalis is one of the most abundant bacterial species of the human gastrointestinal microbiome playing a major role in maintaining gut homeostasis, particularly in newborns. Some strains are widely used as probiotic in food (supplements) for their health-promoting effects. However, by virtue of its capacity to exchange/acquire virulence factors, *E. faecalis* is frequently associated with severe illness, including bloodstream infections and infective endocarditis [57]. Similarly, the contribution of *E. faecalis* in (colorectal) cancer development is controversial with some studies suggesting cancer promoting capacities while others reported protective effects [57], likely depending on the presence/absence of specific virulence factors. De Almeida et al. (2019) assessed the oncogenic potential of different *E. faecalis* strains isolated from CRC patients and healthy controls on tumor cell lines. Four strains from controls had an antiproliferative effect on three tumor cell lines, whereas four other strains (two from CRC patients and two from controls) showed no effect [58]. Upon infection of colonic epithelial cells, *E. faecalis* induces the production of

reactive oxygen species (ROS), leading to DNA damage and activation of multiple signaling pathways, thereby contributing to cancer [59]. Regarding epidemiological literature, most of the 11 studies summarized in Supplementary Table 8 compare the (relative) abundance of the bacterium in feces from cancer patients versus healthy controls [21, 58, 60-68]. Five studies reported a higher abundance of *E. faecalis* in feces from CRC patients compared to healthy controls or controls with polyps (Supplementary Table 8) [60, 62, 63, 65, 67], whereas one study reported an opposite pattern [58]. *E. faecalis* was more frequently found in tumor tissue as compared to normal adjacent tissue (Supplementary Table 8) [21, 68]. Also, patients with a history of *E. faecalis* infective endocarditis showed a higher risk of being diagnosed with CRC [61, 64].

Escherichia coli

Escherichia coli is a commensal bacterium and part of the normal human gut flora. However, within the species, different pathogenic groups exist causing various types of enteric infections [69]. Moreover, some *E. coli* strains contribute to CC development through the production of cyclomodulins; toxins that induce DNA double-stranded breaks, chromosomal instability and cell cycle arrest [70, 71]. Particularly, strains harboring the pks genomic island (pks$^+$ *E. coli*) produce colibactin, which is subject of research in relation to CC during the last decade [71]. Microbiome studies aiming at the identification of bacterial genera and abundance of bacterial species in the gut of CRC patients revealed a significantly reduced (relative) abundance of *Escherichia* in feces of CRC patients compared to healthy controls [72-74]. Supplementary Table 9 summarizes the results of 14 studies assessing the link between (mainly) oncogenic/cyclomodulin-producing *E. coli* and CRC [14, 22, 73, 75-85]. Multiple studies report overrepresentation of *E. coli* in tumor tissue or fecal samples from CRC patients compared to (paired) normal tissue or fecal samples from healthy controls [75, 76, 78, 80]. Presence of pks$^+$ *E. coli* or specific genes coding for toxins was assessed in 7 studies, most of which observed a significantly higher prevalence among CRC patients, though substantial difference in observed prevalences existed between studies (Supplementary Table 9) [22, 76, 77, 79, 81-83].

Fusobacterium nucleatum

The anaerobic bacterium *Fusobacterium nucleatum* is one of the most dominant species of the oral microbiome, displaying opportunistic behavior by causing oral inflammations such as periodontitis and gingivitis. Periodontitis is often caused by a complex of bacteria and is characterized by degradation of the soft tissue and alveolar bone around the teeth

and tooth loss [86]. The last decades, the association between *F. nucleatum* and CRC is intensively studied. Multiple mechanisms have been proposed by which *F. nucleatum* promotes carcinogenesis, including the suppression of antitumor activities of the host, promotion of tumor cell proliferation and the induction of a pro-inflammatory tumor microenvironment [87]. The earliest research assessing the abundance of *Fusobacterium* (at genus level) in relation to CC, dates from 1980, where lower numbers of *Fusobacterium* were isolated from feces of CC patients as compared to healthy controls [88]. A large number of microbiome studies have been published during the last decades in which the composition of the major phyla of commensal bacteria, including *Fusobacterium*, was examined in cancer patients versus (healthy) controls. Over 100 epidemiological studies have been identified in the primary search specifically mentioning *F. nucleatum* in the abstract. Supplementary Table 10 summarizes the characteristics and main outcomes of 42 of these epidemiological studies which met the eligibility criteria, most of which comparing *F. nucleatum* presence/abundance in feces from CRC patients versus healthy controls and of tumor tissue versus adjacent normal tissue (Supplementary Table 10) [13-15, 17, 41, 44, 46, 82, 85, 89-121]. Two meta-analyses published in 2020 reported pooled ORs of 8.3 for detection of *F. nucleatum* in colorectal specimens (feces/mucosa/tissue) and being diagnosed with CRC, and 10.06 for detection of *F. nucleatum* in CRC tissue versus healthy tissue from controls (Supplementary Table 7) [122, 123]. A similar odds was observed when comparing the presence/abundance of *F. nucleatum* in fecal samples from CRC patients versus healthy controls (OR: 9.01, 95%CI 3.39-23.95; n=7 studies) [91, 92, 94, 98, 102, 105, 122]. The odds of detecting *F. nucleatum* in CRC tissue was also significantly higher than in adjacent normal tissue (OR: 2.42, 95%CI 1.62-3.61; n=7 studies) [14, 33, 92, 93, 95, 100, 103, 122]. The association between *F. nucleatum* and CRC appeared stronger in Asian populations (OR 12.6; 95%CI 7.2-21.9) compared to American and European populations (OR 5.6; 95%CI 2.8-11.6 and OR 4.6; 95%CI 2.5-8.4 respectively) [123]. A previous review identified a relatively large difference in observed prevalences of *F. nucleatum* in CRC tissue, ranging from 13% till 75% [87]. In addition to detection of *F. nucleatum* DNA in fecal or tissue samples, few studies assessed the humoral immune response against *F. nucleatum* in relation to CRC. Significantly higher IgA and IgG titers were measured in serum from CRC patients compared to healthy controls and controls with benign colon diseases (Supplementary Table 10) [96, 119]. Antibody titers were higher for proximal versus distal CRC [119], though no association between antibody titer and CRC stage was found [96]. A large European study investigated whether higher *F. nucleatum* antibody titers represent a risk factor for CRC later in life (0.4-8.5 years after serum sampling) [104]. Antibody responses against ≥2 or ≥3 out of 11 *F. nucleatum* proteins were similar for individuals who were diagnosed with CRC after the serum sampling compared

to a control group without cancer diagnosis (17% and 9% vs 21% and 9% respectively). Also, no difference was observed when stratifying for time of serum analyses and CRC diagnoses [104]. In addition to CRC, the potential role of *F. nucleatum* in carcinogenesis of other (GI) cancers is gaining momentum. While *F. nucleatum* is an oral bacterium, it is suggested to be translocated hematogenously from the oral cavity to tumor tissues via the bloodstream during periodontitis. Here, its adhesion protein Fap2 binds to a carbohydrate (Gal□GalNAc), which is overrepresented in tumor cells of several GI cancers [124]. Ten studies investigated the association between *F. nucleatum* and other organs in the GI tract, including esophagus (n=4) [101, 109, 117, 118], stomach (n=6) [44, 101, 107, 109, 113, 114]), pancreas (n=2) [101, 112], liver and gallbladder (both n=1) [85, 101]. In two studies, presence of *F. nucleatum* was confirmed in (tumor) tissue samples of a large portion of esophageal cancer patients [117, 118], whereas presence of *F. nucleatum* in saliva of esophageal cancer patients as compared with controls appeared less consistent [109, 118]. For gastric cancer, significant associations between *F. nucleatum* presence/abundance in tumor tissue were observed as compared to adjacent normal tissue and individuals with other underlying medical conditions [44, 107, 113, 114]. For pancreatic cancer, the antibody concentrations against *F. nucleatum* in saliva and plasma were higher in cancer patients than in controls, [112], with poor evidence of *F. nucleatum* presence in pancreatic tumor tissue (Supplementary Table 10).

Porphyromonas gingivalis

The gram-negative bacterium *Porphyromonas gingivalis* is part of the oral microbiome and is considered a leading cause of severe periodontitis [86, 125]. The (chronic) inflammatory response in periodontitis is associated with several systemic diseases including cardiovascular disease, diabetes and cancer [126]. *P. gingivalis* exhibits a range of virulence factors enabling invasion of (oral) endothelial and epithelial cells, dysregulation of the immune response and inhibition of apoptosis, conditions that favor cancer initiation [126]. Recently, Liu *et al*. (2019) published a literature review about the role of *P. gingivalis* in gastrointestinal cancers, in which they observed that this bacterium is particularly associated with esophageal, colorectal and pancreatic cancers [127]. Generally the presence or abundance of *P. gingivalis* is measured through detection of bacterial DNA in oral/tissue/fecal samples or antibody serum titers (Supplementary Table 11) [32, 109, 118, 120, 128-135]. Increased orodigestive cancer mortality was reported in individuals with higher serum antibody levels against *P. gingivalis* (relative risk [RR] 2.25; 95%CI 1.23-4.14) although this study also included oral cancers and no site-specific estimates were provided [128]. With regard to esophageal cancer, in a Chinese cohort, serum IgA and IgG titers against *P. gingivalis* were significantly higher in esophageal squamous cell carcinoma (ESCC) patients compared to controls (Supplementary Table 11)

[131]. Similarly, the bacterium was detected in 61% of tumor tissues whereas none of the normal esophageal tissue contained bacterial DNA [130]. This was confirmed in a US cohort, in which the OR for presence of *P. gingivalis* DNA in an oral swab was 1.30 (95%CI 0.96-1.77) for ESCC patients versus controls (Supplementary Table 11) [Peters, 2017]. Moreover, some mechanistical evidence exists for the oncogenic potential of *P. gingivalis* in CRC development [134], though limited studies assessed the association from an epidemiological perspective (Supplementary Table 11). Ahn *et al*. (2012) reported a significant excess risk of CRC mortality in patients with periodontal disease (RR: 3.58; 95%CI 1.15–11.16); however, direct links with *P. gingivalis* serum levels were not provided [128]. Yet, elevated levels of *P. gingivalis* have been found in feces and tumor tissue samples from CRC patients (Supplementary Table 11) [134, 135]. Regarding pancreatic cancer, two studies reported a significant association between *P. gingivalis* and pancreatic cancer, either based on its presence in oral samples of cases versus controls (OR 1.60; 95%CI 1.15-2.22) [129] or based on high (>200 ng/ml) versus low (≤200 ng/ml) serum antibody titers (OR 2.14; 95%CI 1.05-4.36) [132]. However, another study reported a significant reduced abundance of *P. gingivalis* in saliva of pancreatic cancer patients (Supplementary Table 11) [32]. Although, several studies confirmed an association between periodontitis and liver cancer and *P. gingivalis* is suggested to play a role in liver diseases [127], no epidemiological studies directly assessing the association between *P. gingivalis* and liver cancer were found.

Nontyphoidal *Salmonella*

Nontyphoidal *Salmonella* (NTS) is one of the major zoonotic bacteria causing (foodborne) gastrointestinal infections. Most infections are mild and do not require medical care. NTS induces cell transformation of pre-transformed cells by activating the AKT and MAPK pathways through secretion of its effector proteins which was shown for *S. enterica* subspecies *enterica* serotype Typhimurium in an experimental setting using murine gallbladder organoids [136]. Seven epidemiological articles were found assessing the association between NTS infection and colorectal (n=3), colon cancer (n=2), biliary tract cancer (n=2) and gallbladder, gastric, liver and pancreatic cancer (latter four: n=1) (Supplementary Table 12) [29, 137-142]. The humoral immune response against *Salmonella* flagellin in CRC patients and individuals with colorectal polyps versus healthy controls was assessed in two different cohorts (Supplementary Table 12) [140]. In both cohorts, significantly higher antibody titers were observed in CRC patients versus controls [140]. In another study, the *Salmonella* effector protein AvrA, exerting a role in carcinogenesis through activation of the host β-catenin pathway, showed higher abundance in colorectal tumor tissues than in healthy mucosa [141]. Moreover, two studies focusing

exclusively on CC, used a comparable design in which the risk of developing CC after severe salmonellosis was investigated [138, 142]. In both studies, the risk of proximal CC after salmonellosis was higher than of distal CC for most subgroups (Supplementary Table 12) [138, 142]. However, a generally lower risk of CC after NTS infection was found in the Danish cohort as compared with the Dutch cohort. Whilst a significantly increased risk of proximal CC was observed after *S*. Enteritidis infection in the Dutch cohort (SIR: 1.86; 95%CI 1.28-2.61), the Danish study reported only a significant increased risk for serotypes other than Enteritidis and Typhimurium (HR: 1.40; 95%CI 1.03-1.90). In the Dutch cohort, NTS infections reported in people ≥20 years were included in the study, while the risk estimates of the Danish cohort study were based on NTS infections in all age groups. Risks of overall CC and proximal CC were particularly higher in people infected between 20 and 60 years of age in the Dutch cohort (SIRs 1.54 [95%CI 1.09-2.10] and 2.12 [95%CI 1.38-3.09] respectively) [142]. After the age of 60, the incidence of cancer increases substantially due to multiple factors, such as spontaneous mutations [143]. Hence, this may dilute the observed effect of NTS in risk estimates including the older age groups. In a Taiwanese cohort, a HR of 1.03 (95%CI 0.72-1.47) was reported for the combined risk of colon and rectum cancer, while an over 2-fold significantly increased risk of gastric cancer was observed (Supplementary Table 12) [137]. In addition, also an increased risk of developing biliary tract cancer after salmonellosis was reported in both a Dutch study (SIR: 1.53; 95%CI 0.70-2.91) and the Taiwanese cohort (HR: 2.23; 95%CI 0.83-6.05), though numbers were low in both studies [29, 137]. Iyer *et al*. (2016) demonstrated the presence of traces of *S*. Typhimurium in 12 out of 26 tumors from Indian gallbladder cancer patients (Supplementary Table 12) [139].

Typhoidal Salmonella

Salmonella Typhi is a pathogenic bacterium causing typhoid fever mainly in developing countries in Southeast-Asia, South-America and Africa. Upon invasion of the intestinal mucosa, *S. Typhi* spreads to other organs leading to colonization of the gallbladder and liver in 2-5% of the infections [144]. Subsequently, an estimated 1-4% of the infected people become chronic asymptomatic carriers, as the bacterium is able to form biofilms on gallstones [144]. A strong correlation exists between concurrent carriage of chronic *S.* Typhi and gallstones (up to 90% in endemic countries), with the latter considered a major risk factor for GBC [144]. While GBC is a (relatively) rare malignancy in Western countries, its incidence is higher in countries with endemic *S.* Typhi. Secretion of the typhoid toxin (CDT) is suggested to play a role in establishing long-term infection. Also, CDT induces DNA double-stranded breaks and is involved in activation of the MAPK and AKT pathways, ultimately

leading to transformation of pre-transformed cells [136, 144]. Numerous epidemiological studies assessing the association between *S.* Typhi infection and/or chronic carriage and cancer have been published. Supplementary Table 13 summarizes 23 studies, mostly focusing on GBC [85, 136, 139, 145-164]. The design of the studies is often based on observed differences in antibody response against *S.* Typhi between GBC patients and controls or detection rates of the bacterium in gallbladder tissue or bile. A meta-analysis from 2014 reported an overall pooled OR of 4.28 (95%CI 1.84-9.96) (Supplementary Tables 7, 13) [146-149, 151-161, 163-165]. Results were similar for serological detection (OR: 3.52; 95%CI 2.48-5.00) versus culturing methods (OR: 4.14; 95%CI 2.41-7.12) [165]. Results were corroborated by a more recent meta-analysis of Koshiol *et al*. (2016) who also reported slightly higher estimates for culturing methods as compared to antibody detection (Supplementary Tables 7, 13) [145, 147, 148, 150, 152, 154, 155, 158, 160, 162-164, 166].

Schistosoma spp.

Schistosoma is a genus of trematode worms, commonly referred to as blood flukes, which cause chronic schistosomiasis characterized by intestinal and hepatosplenic disease [167]. The species causing most infections are *Schistosoma japonicum*, endemic in parts of China, East Asia and the Philippines, *Schistosoma mansoni* mainly occurring in South America and Africa, and *Schistosoma haematobium* present in Africa and the Middle East. Schistosomiasis is associated with high morbidity and mortality levels, particularly in populations with poor sanitation and limited access to safe drinking water [167]. Infection with *S. japonicum* is assigned as a group 2b (possibly carcinogenic to humans) carcinogen by the International Agency for Research on Cancer (IARC) for its role in liver cancer, whereas the carcinogenicity of *S. haematobium* infection in bladder cancer is well established (classified as group 1 by IARC) [168]. Part of the *Schistosoma* eggs become trapped in the gut and liver where they induce inflammations and granulomas, thereby being the major drivers of *Schistosoma* pathogenicity. Moreover, *S. mansoni* soluble egg antigens activate c-Jun (proto-oncogene) and STAT3 (transcription factor), which facilitate the development and progression of HCC tumor formation [169]. In addition, the inducible nitric oxide synthase present in host cells as part of antimicrobial defense against *S. japonicum*, has a promoting effect on p53 mutations, and tumor formation and progression [170]. Disease severity is generally worse in case of co-infection of a *Schistosoma* species with hepatitis B (HBV) or hepatitis C virus (HCV), which is frequently observed in areas with high incidences of both pathogens [171]. Chronic infection with HBV/HCV can lead to liver fibrosis/cirrhosis and ultimately liver cancer. The interaction between schistosomiasis and chronic hepatitis infection and their

combined contribution to cancer formation remains to be clarified [171]. With regard to the epidemiological evidence, the earliest documentation dates from the late 1970s with several sources, mainly from China, reporting the transforming potential of *S. japonicum* infections. They describe (geographical) associations between S. *japonicum* infection and (mortality from) colorectal cancer [172]. It should be mentioned that most of these older studies are difficult to access and hardly describe the methods used. Increased risks of CC of 3.3 and 1.2 were observed among Chinese individuals with a history of *S. japonicum* infection, although the latter was not significant, whereas 8-fold and almost 4-fold increased risks were observed for rectal cancer and liver cancer respectively [173, 174]. Still, literature about *S. japonicum* and cancer is limited to a number of case reports and case series and relatively few larger case-control/cohort studies (Supplementary Table 14) [173-176]. Similarly, for *S. mansoni* a small number of epidemiological studies have been published (Supplementary Table 14) [177, 178]. Particularly for this species, quantification of the cancer risk after infection is more challenging as it remains elusive whether it directly promotes cancer development or indirect through the action of HBV/HCV co-infections [179]. In the past, HBV/HCV viruses have often been transmitted (via contaminated blood, syringes, needles) during antischistosomal parenteral therapy, particularly in Egypt [179]. None of the seven HCC patients with a history of schistosomiasis in Brazil had antibodies against HCV, whereas four were tested positive for HBV-antibodies [178]. In an Egyptian cohort, CRC of patients with *S. mansoni* schistosomiasis occurred at an earlier age and were in a more advanced stage as compared to CRC in patients not associated with schistosomiasis (Supplementary Table 14) [177].

Streptococcus spp.

Streptococcus gallolyticus subsp. *gallolyticus* (hereafter referred to as *Sgg*), formerly known as *Streptococcus bovis* biotype I is a low-abundance commensal bacterium of the gut. *Sgg* is associated with infective endocarditis and CRC, which is proposed to be related to its capacity to adhere to collagen (types I and IV), frequently present in damaged heart valves and tumors [180]. Laboratory evidence shows that *Sgg* promotes tumor development through increasing epithelial cell proliferation and upregulated β-catenin levels [181]. Many studies assessed the association between infection with (different types of) *Streptococcus* and endocarditis and CRC, albeit that the magnitude of the observed associations in these studies vary considerably, probably due to differences in study populations and methodology (Supplementary Table 15) [9, 11, 13, 17, 19, 31, 61, 182-203]. A meta-analysis published in 2011, showed that 60% of people with *S. bovis* infections undergoing colonoscopy were

diagnosed with CRC, which exceeds the CRC prevalence in the general population (10-25%) (Supplementary Tables 7, 15) [182-188, 190-192, 204]. Similarly, the risk of CRC was significantly increased among people infected with *Sgg* compared to individuals infected with *S. bovis* biotype II (pooled OR: 7.26 [95%CI 3.94-13.36]) [204]. Another, more recent, meta-analysis (Supplementary Table 7) reported pooled ORs of 14.54 (95%CI 5.66–37.35) and 2.52 (95%CI 1.14–5.58) for co-occurrence of *S. bovis* infective endocarditis and CRC and fecal carriage and CRC respectively (Supplementary Table 7) [205]. Fecal carriage of *Streptococcus* ranged from 6-46% in individuals with adenomas or CRC to 7-14% in controls (Supplementary Table 15) [180, 203], while the presence of *Streptococcal* DNA in tumor tissue varies considerably between studies (3-74%) [180]. *Sgg* infection is more frequently associated with adenomas than carcinomas [204, 206]. Whilst the majority of evidence on the association between *Streptococcus* and CRC originates from studies assessing the presence of the bacterium in feces and tumor tissue concomitant with CRC diagnosis as compared with controls, in vitro and in vivo evidence for a causal relationship between *Streptococcus* and CRC is limited [180]. Corredoira *et al*. (2015) found that within a cohort of people with a history of *S. bovis* infective endocarditis 43/54 (80%) of the individuals developed a colorectal neoplasm ([non-]advanced adenoma or carcinoma) several years after the acute infective endocarditis (mean follow-up time 60.6 months) (Supplementary Table 15) [61]. More recently, a series of large cohort studies have been published assessing the association from a serological perspective. In these studies, using data from Germany, Spain and the US, associations between several *Sgg* antigens and CRC were confirmed. Results from these studies collectively showed that an earlier infection with *Streptococcus*, as detected by antibodies, is a predictor of CRC development, also when the antibodies were detected in blood collected up to 10 years before cancer diagnosis. One study showed a particular stronger association in people aged <65 years (Supplementary Table 15) [196, 197, 199, 200]. In contrast to the numerous studies corroborating the association between *Streptococcus* and CRC, only 3 studies were found describing *Streptococcus* in relation to other GI malignancies (mainly pancreatic and liver cancers), but these concerned (sub) species other than *Sgg* (Supplementary Table 15) [190, 195, 202].

Strongyloides stercoralis

Strongyloides stercoralis is a soil-transmitted helminth mainly occurring in (sub)tropical regions where it is estimated to cause over 100 million infections annually [207]. Most *S. stercoralis* infected people are asymptomatic or have intermittent symptoms including abdominal pain, diarrhea, respiratory complaints or skin problems. Individuals with an

HTLV-1 (human T-cell lymphotropic virus type 1) coinfection or immunocompromised patients, the autoinfective cycle within the host can result in a hyperinfection, characterized by disseminated colonization affecting numerous organs with a high mortality rate [208]. This causes colitis-like intestinal symptoms, including ulcer formation, patchy inflammation, submucosal hemorrhage and eosinophilic infiltration, mimicking Crohn's disease and ulcerative colitis [209]. The mechanism by which *S. stercoralis* contributes to initiation and/or progression of malignancies is unclear. It is also unclear whether this nematode exhibits direct oncogenic potential. It is hypothesized that *S. stercoralis* stimulates replication of HTLV-1, which is known to cause adult T cell leukaemia/lymphoma. The association between *S. stercoralis* and HTLV-1 was supported by epidemiological data that which showed an over two-fold increased prevalence of *S. stercoralis* among HTLV-1 patients as compared to HTLV-free individuals. However, risk of GI cancers other than liver cancer were not elevated amongst HTLV-infected individuals [210]. Epidemiological data addressing the association between *S. stercoralis* infection and GI cancers is limited to few case-control studies (Supplementary Table 16) [210, 211] and a number of case reports addressing strongyloidiasis and gastric cancer [212] and intestinal cancer [209, 213-215]. An almost three-fold higher risk of developing cancer in the biliary tract was observed amongst patients with a *S. stercoralis* infection in a Japanese cohort (OR 2.7; 95%CI 1.1-6.3) [211]. Yet, these results could not be corroborated in a larger study several years later [210]. Moreover, no significant increased risk of other GI cancers in patients infected with *S. stercoralis* were found (Supplementary Table 16) [210, 211]. Both studies in Japan, as well as the case reports, assessed the co-occurrence of *S. stercoralis* infection and cancer, which hampers definitive conclusions about the direction of the association. In another study, strongyloidiasis was considered an opportunistic infection, as patients with GI cancer receiving chemotherapy were found to have a 6.7 times higher risk of being infected with *S. stercoralis* as compared to patients receiving treatment for other forms of cancer (OR 6.7; 95%CI 1.3-34.2) [216].

Discussion

A growing number of microbial species is being associated with the induction and progression of cancers, partly driven by the development of new diagnostic techniques allowing for a rapid and better understanding of the complex interplay between commensals, pathogens and human cells. For some (mostly pathogenic) microorganisms, the link with cancer has been studied repeatedly in different study populations for many years, whereas for others the scarce evidence is scattered and originates from relatively recent studies. In this review, we provide a comprehensive consideration of epidemiological insights into the

association between GI cancers and 13 bacteria and parasites. Figure 2 provides a graphical summary of the study characteristics. Most studies concerned *S.* Typhi, *Streptococcus* spp. and the commensals *F. nucleatum* and *B. fragilis*. Amongst studies comparing the incidence, presence or abundance of these microorganisms in cancer patients versus healthy controls or the general population, significant positive associations were observed for *B. fragilis* (6/11 studies), *Campylobacter* spp. (4/9), *Clostridium* spp. (4/8), *Cryptosporidium* spp. (3/3), *E. faecalis* (4/7), *E. coli* (5/9), *F. nucleatum* (13/24), *P. gingivalis* (7/11), NTS (3/6), *S.* Typhi (9/10), *Schistosoma* spp. (2/2), *Streptococcus* spp. (11/14), and *S. stercoralis* (1/2) (Supplementary Tables 3-6, 8-16). It is noteworthy that over half of the reviewed studies included less than 50 cancer patients, whilst only 20 percent of the studies included over 100 cancer patients (Supplementary Tables 3-6, 8-16). For relatively rare malignancies and/or infections with low incidences, the samples sizes of studies are generally rather small, hence, statistical significance is frequently not achieved.

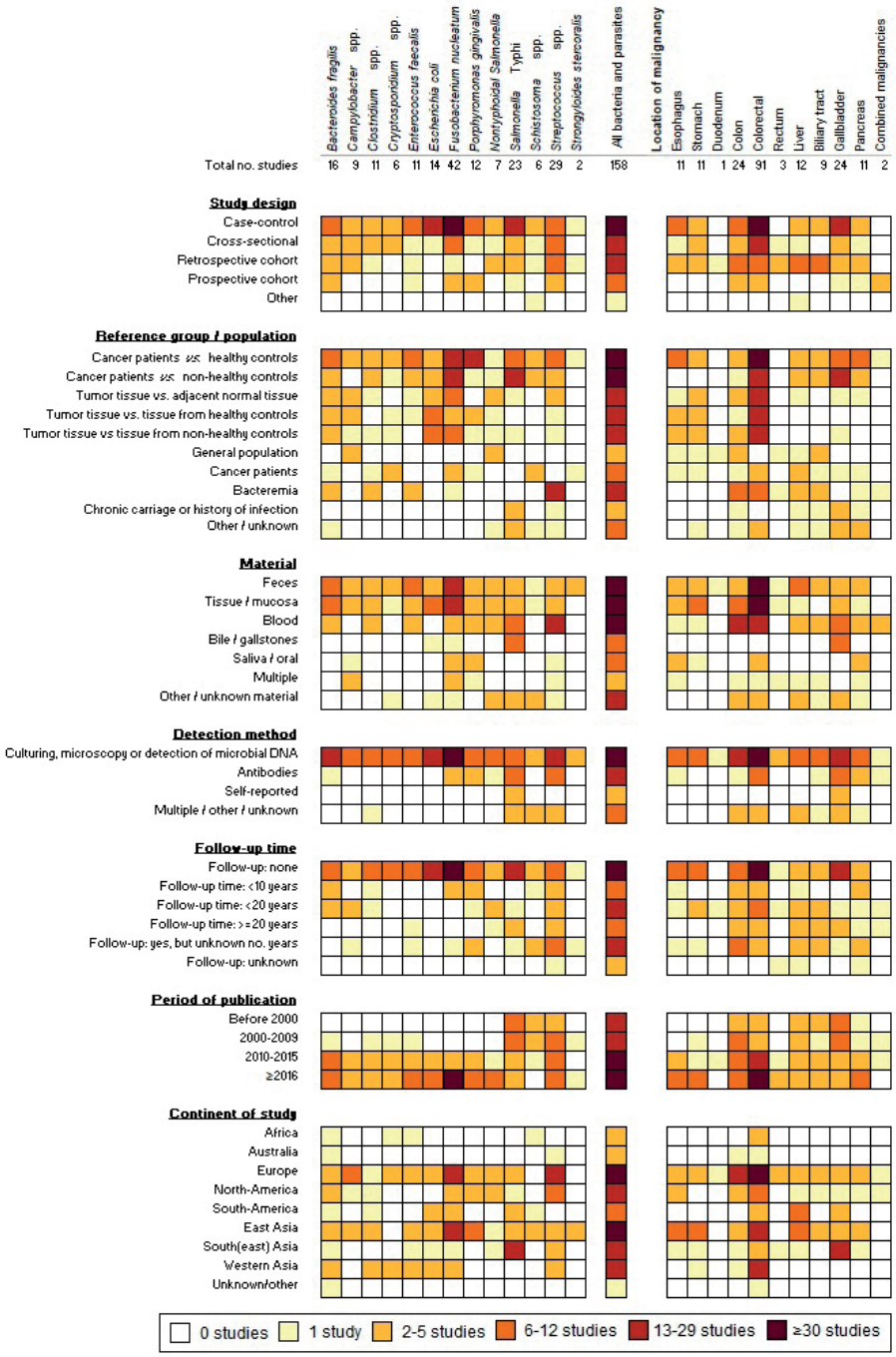


Figure 2. Graphical summary of the 158 included studies.

Forty-two studies were identified, that had a study design suitable for estimating the cause and effect relation between bacterial/parasitic infection and GI cancer by inclusion of several years of follow-up (Figure 2). These concerned mostly bacterial infections, including *Streptococcus* (n=16), *Porphyromonas* (n=5), *S.* Typhi (n=5), NTS (n=4), *Campylobacter* (n=4), *Fusobacterium* (n=5), *E. faecalis* (n=2), *B. fragilis* (n=4) and *Clostridium* (n=1) (Figure 1). For pathogenic microorganisms able to cause severe illness, such as NTS, *Campylobacter* spp. and *S. bovis*, often person-level records exist (e.g. from physicians, laboratory diagnoses or surveillance systems) that provide opportunity for linkage with cancer registry data. Subsequently, this allows for comparison of the cancer incidence among those with a registered history of infection and the cancer incidence in the general population. Maximum follow-up durations of <10 years, 10-20 years and ≥20 years were reported in 10, 13 and 6 studies, respectively, whereas details about maximum follow-up time were not listed in a substantial fraction of the articles (n=13) (Figure 2). Amongst the studies with follow-up time, five assessed whether seropositivity is a predictor of GI cancer risk later in life [104, 128, 132, 199, 200]. Significantly increased risks of up to 2-fold were observed for CRC and pancreatic cancer among individuals positive for *S. bovis* and *P. gingivalis* antigens, respectively, in the period up to 10 years before cancer diagnosis [132, 199, 200]. Conversely, seropositivity for *F. nucleatum* was not a predictor of CRC risk later in life [104].

For studying the association between commensal bacteria and cancer, a registry-study based on linkage of retrospective databases is often not feasible. However, also the number of studies using a prospective design (with a follow-up period) was limited. Instead, studies focusing on commensal bacteria, including *E. faecalis, F. nucleatum* and *B. fragilis*, primarily compare the presence or abundance of bacteria in patient material (feces, tumors, saliva) to the presence/abundance in normal tissue or samples from healthy controls, measured at one point in time, usually resulting from a medical intervention (colonoscopy, endoscopy, surgery) (Figure 2). These studies are often presented as case-control studies; however, the cross-sectional design without retrospective (risk-factor) data does not allow for assessment of the cause and effect relationship. Tjalsma *et al*. (2012) described the distinct temporal associations and separate roles of different bacteria with CRC tissue in a 'driver-passenger' model [217]. In this model, bacteria that initiate tumorigenesis by causing DNA damage and malignant transformation of epithelial stem cells are referred to as 'driver bacteria'. Subsequently, the induced intestinal alterations favor the proliferation of opportunistic 'passenger bacteria' leading to colonization of the tumor microenvironment, thereby outcompeting the original driver bacteria [217]. Among the bacteria discussed here, *Bacteroides*, *Clostridium* and members of the Enterobacteriaceae family (including *Salmonella* spp.) are considered driver bacteria according to this model, whereas *Fusobacterium* spp.

and *Streptococcus* are considered passengers [217, 218]. Although the driver-passenger model was originally developed for CRC, it might apply for other malignancies as well, as *F. nucleatum* and *P. gingivalis* are suggested passengers in pancreatic cancer [219].

Twenty-eight studies exclusively made use of control groups with underlying medical conditions, including individuals presenting with gastritis or cholelithiasis, or patients undergoing colonoscopy for gastrointestinal complaints. Although selection of individuals with medical conditions as controls is a convenient option (i.e. with regard to obtaining samples) and can provide insights into the correlation between microbial infection/presence/(relative) abundance and the presence of risk factors (such as cholelithiasis) or pre-malignant conditions (e.g. polyps), the lack of a baseline healthy reference group hampers the accurate assessment of causality [220].

For all bacteria and parasites, except *B. fragilis*, the association was studied for more than one malignancy, with colon and rectal cancer being the most frequent malignancies analyzed in the literature (CRC n=83, CC n=21, rectum n=1) (Figure 1). Whilst cancers in colon and rectum are commonly combined into one outcome or risk estimate (i.e. CRC) in epidemiological research given the similarities in anatomical structure; differences in risk factors, etiology and incidence favor the reporting of separate estimates for these subsites [221, 222]. A recent study showed that the relative abundance of commensal bacteria differs at genus level in patients with sigmoid colon cancer as compared to rectal cancer patients [222]. Similarly, part of the included studies reported separate estimates for proximal versus distal CC and EAC versus ESCC, as these cancers differ in life-style and diet related risk factors [223]. Particularly for pathogenic bacteria establishing infection in the intestine, the proximal part of the colon (i.e. closest to the ileum) might be of more interest as exposure to bacteria is highest in this part of the colon. For *B. fragilis*, *E. faecalis*, *E. coli*, *F. nucleatum* and NTS, significant differences in the estimated cancer risk or the observed presence in samples in proximal versus distal CC were reported, though not consistent across bacteria for either of the colon subsites.

Substantial differences were observed in the magnitude of the microorganism-cancer association across different countries/continents for some of the bacteria and parasites. Various factors might underly these inconsistencies, including differences in incidence of both the bacterial/parasitic infections and the specific malignancy, diagnostic performance and cancer screening programs. With respect to global cancer epidemiology, Asian, African and Latin American countries generally have higher incidences of esophageal, stomach, liver and gallbladder cancer, although there is an ongoing displacement towards cancers associated

with a higher development index (based on life expectancy, education and income), such as colorectal and pancreatic cancer, which display higher incidences in countries in Europe, Northern America and Australia/New Zealand [1]. Moreover, the cancer inducing and/or promoting capacities sometimes differ between subtypes of the same microorganism, as suggested for *E. coli*, NTS, and *Campylobacter*, consequently leading to differences in cancer risk when global distributions of microorganism subtypes vary [224].

For pathogenic bacteria and parasites, the burden of disease is usually based on the incidence, morbidity and possible long-term sequalae, expressed as disability adjusted life years (DALYs) or years of life lost (YLLs) [225]. According to the IARC, none of the bacteria and only one parasite (*Schistosoma*) discussed in this review is classified as potentially carcinogenic for humans [193]. Considering more bacteria and parasites to this IARC list would imply cancer to be recognized as a long-term sequalae of infection and as a consequence a much higher disease burden associated with the specific pathogens. This requires ongoing research to unravel the magnitude and conditions for existence of the association between bacteria/parasites and cancers, and the fraction of GI cancers attributable to these infections.

In order to establish a causal relationship between a microorganism and a disease, the four Koch's postulates need to be met. Overabundance of the microorganism in people with the disease compared to healthy individuals and isolation of the microorganism from diseased people (either directly or indirectly through serum antibodies) as defined in the first two postulates, are fulfilled in epidemiological studies. However, the last two postulates defining that the microorganism must be able to cause disease in a healthy organism and can be isolated from an experimental host, are not always met and require an experimental design. Hence, this might be a goal for the years to come. Whilst for pathogenic bacteria, measures for prevention of spread (e.g. improved kitchen hygiene, sanitation, etc.) could aid in the prevention of cancers, for commensals prevention is more complex and mainly requires life-style changes related to smoking, eating habits, alcohol consumption and physical activity, as these factors shape the composition of the microbiome [226-229].

The aim of this review was to provide an comprehensive overview of the volume of research on the several documented associations between bacteria/parasites and GI cancers. However, describing all associations in detail was beyond the scope of this review. Moreover, comparing the magnitude and strength of evidence of the cancer promoting potential between the bacteria and parasites would require thorough quality assessments and weighing of the included articles considering the different study designs and sample sizes, which would be a worthwhile future research objective. As for all literature reviews,

relevant articles might have been missed in the search of databases, which may especially be true for those associations addressed only to a limited extent in a few bodies of the scientific literature and not cited in other papers. Likewise, we neither included bacteria nor parasites for which the main body of evidence is indicative of reverse causality, nor bacteria that were exclusively addressed at phylum or family level in, for example, microbiome studies. Last, the review is subject to a degree of publication bias, as studies observing no or limited associations are less likely to be published.

Conclusion

In conclusion, this review provides a broad overview of the currently existing epidemiological literature about bacterial and parasitic infection/colonization in relation to development and progression of GI malignancies. While the rapidly growing body of studies based on microbiome sequencing provides valuable insights into the relative abundance of different bacterial taxa in cancer patients as compared to individuals with pre-malignant conditions or healthy controls potentially leading to new biomarkers for early detection of cancer, more research is needed to fulfil Koch's postulates. This involves the use of follow-up data, assessing the complex role of bacteria and parasites in cancer epidemiology and experimental data where isolated infectious species are tested under controlled (laboratory) conditions for their transforming potential. In the future, artificial intelligence could aid in the analysis and transformation of the increasing amount of research data into meaningful risk estimates.

Author contributions

JD, LM-G, EF, and JN designed and conceptualized the study. JD performed the literature search and provided a first draft of the review. All authors contributed to writing and revising the manuscript.

Funding

This work was supported by KWF grant "Bacterial food poisoning and colon cancer; a cell biological and epidemiological study" (2017-1-11001). No funding bodies had any role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a Cancer Journal for Clinicians*, 71(3), 209-249. doi: 10.3322/caac.21660
2. Chiu, C.Y. and Miller, S.A. (2019). Clinical metagenomics. *Nat. Rev. Genet*., 20(6): p. 341-355. doi: 10.1038/s41576-019-0113-7
3. Van Elsland, D. and Neefjes, J. (2018). Bacterial infections and cancer. *EMBO Rep*., 19(11): p. e46632. doi: 10.15252/embr.201846632
4. Kuo, S.-H. and Cheng, A.-L. (2013). *Helicobacter pylori* and mucosa-associated lymphoid tissue: what's new. *Hematology*, 2013(1): p. 109-117. doi: 10.1182/asheducation-2013.1.109
5. Berger, H., Marques, M. S., Zietlow, R., Meyer, T. F., Machado, J. C., and Figueiredo, C. (2016). Gastric cancer pathogenesis. *Helicobacter*, 21, 34-38. doi: 10.1111/hel.12338
6. Vacante, M., Ciuni, R., Basile, F., and Biondi, A. (2020). Gut microbiota and colorectal cancer development: A closer look to the adenoma-carcinoma sequence. *Biomedicines*, 8(11), 489. doi: 10.3390/biomedicines8110489
7. Haghi, F., Goli, E., Mirzaei, B., and Zeighami, H. (2019). The association between fecal enterotoxigenic *B. fragilis* with colorectal cancer. *BMC Cancer*, 19(1), 1-4. doi: 10.1186/s12885-019-6115-1
8. Zamani, S., Taslimi, R., Sarabi, A., Jasemi, S., Sechi, L. A., Feizabadi, M. M. (2020). Enterotoxigenic *Bacteroides fragilis*: a possible etiological candidate for bacterially-induced colorectal precancerous and cancerous lesions. *Front. Cell. Inf. Microbiol.* 9, 449. doi: 10.3389/fcimb.2019.00449
9. Justesen, U. (2020). "Bacteraemia with anaerobic bacteria and association with colorectal cancer," in *Proceeding of the ECCMID-2020*. [Conference Abstract], (Basel: European Society of Clinical Microbiology and Infectious).
10. Ma, Y., Zhang, Y., Xiang, J., Xiang, S., Zhao, Y., Xiao, M., et al. (2021). Metagenome analysis of intestinal bacteria in healthy people, patients with inflammatory bowel disease and colorectal cancer. *Front. Cell. Inf. Microbiol*., 11. doi: 10.3389/fcimb.2021.599734
11. Abdulamir, A. S., Hafidh, R. R., Mahdi, L. K., Al-jeboori, T., and Abubaker, F. (2009). Investigation into the controversial association of *Streptococcus gallolyticus* with colorectal cancer and adenoma. *BMC Cancer*, 9(1), 1-12. doi: 10.1186/1471-2407-9-403
12. Boleij, A., Hechenbleikner, E. M., Goodwin, A. C., Badani, R., Stein, E. M., Lazarev, M. G., et al. (2015). The *Bacteroides fragilis* toxin gene is prevalent in the colon mucosa of colorectal cancer patients. *Clin. Infect. Dis.,* 60(2), 208-215. doi: 10.1093/cid/ciu787
13. Bundgaard-Nielsen, C., Baandrup, U. T., Nielsen, L. P., and Sørensen, S. (2019). The presence of bacteria varies between colorectal adenocarcinomas, precursor lesions and non-malignant tissue. *BMC Cancer*, 19(1), 1-13. doi: 10.1186/s12885-019-5571-y
14. Fukugaiti, M. H., Ignacio, A., Fernandes, M. R., Ribeiro Júnior, U., Nakano, V., and Avila-Campos, M. J. (2015). High occurrence of *Fusobacterium nucleatum* and *Clostridium difficile* in the intestinal microbiota of colorectal carcinoma patients. *Braz. J. Microbiol*., 46(4), 1135-1140. doi: 10.1590/S1517-838246420140665
15. Hale, V. L., Jeraldo, P., Chen, J., Mundy, M., Yao, J., Priya, S., et al. (2018). Distinct microbes, metabolites, and ecologies define the microbiome in deficient and proficient mismatch repair colorectal cancers. *Genome Med*., 10(1), 1-13. doi: 10.1186/s13073-018-0586-6

16. Keenan, J., Aitchison, A., Greenlees, R., and Frizelle, F. (2014). Are toxin-producing strains of *Bacteroides fragilis* involved in colorectal cancer?:CR27. *Anz. J. Surg*. 84.

17. Kwong, T. N., Wang, X., Nakatsu, G., Chow, T. C., Tipoe, T., Dai, R. Z., et al. (2018). Association between bacteremia from specific microbes and subsequent diagnosis of colorectal cancer. *Gastroenterology*, 155(2), 383-390. doi: 10.1053/j.gastro.2018.04.028

18. Purcell, R. V., Pearson, J., Aitchison, A., Dixon, L., Frizelle, F. A., and Keenan, J. I. (2017). Colonization with enterotoxigenic *Bacteroides fragilis* is associated with early-stage colorectal neoplasia. *PloS One*, 12(2), e0171602. doi: 10.1371/journal.pone.0171602

19. Rahimkhani, M., Mohagheghi, M. A., and Yaraei, K. (2010). Fecal microbial flora in colorectal cancer. *Afr. J. Microbiol. Res*., 4(23), 2622-2625. doi: 10.5897/AJMR.9000323

20. Ulger Toprak, N., Yagci, A. Y. Ş. E. G. Ü. L., Gulluoglu, B. M., Akin, M. L., Demirkalem, P., Celenk, T., et al. (2006). A possible role of *Bacteroides fragilis* enterotoxin in the aetiology of colorectal cancer. *Clin. Microbiol. Infect*., 12(8), 782-786. doi: 10.1111/j.1469-0691.2006.01494.x

21. Viljoen, K. S., Dakshinamurthy, A., Goldberg, P., and Blackburn, J. M. (2015). Quantitative profiling of colorectal cancer-associated bacteria reveals associations between fusobacterium spp., enterotoxigenic *Bacteroides fragilis* (ETBF) and clinicopathological features of colorectal cancer. *PloS One*, 10(3), e0119462. doi: 10.1371/journal.pone.0119462

22. Xie, L. L., Wu, N., Zhu, Y. M., Qiu, X. Y., Chen, G. D., Zhang, L. M., et al. (2016). Expression of enterotoxigenic *Bacteroides fragilis* and polyketide synthase gene-expressing *Escherichia coli* in colorectal adenoma patients. *Zhonghua yi xue za zhi*, 96(12), 954-959. doi: 10.3760/cma.j.issn.0376-2491.2016.12.010

23. Keithlin, J., Sargeant, J., Thomas, M. K., and Fazil, A. (2014). Systematic review and meta-analysis of the proportion of *Campylobacter* cases that develop chronic sequelae. *BMC Public Health*, 14(1), 1-19. doi: 10.1186/1471-2458-14-1203

24. He, Z., Gharaibeh, R. Z., Newsome, R. C., Pope, J. L., Dougherty, M. W., Tomkovich, S., et al. (2019). *Campylobacter jejuni* promotes colorectal tumorigenesis through the action of cytolethal distending toxin. *Gut*, 68(2), 289-300. doi: 10.1136/gutjnl-2018-317200

25. Alexander, J. L., Scott, A., Mroz, A., Perdones-Montero, A., Mckenzie, J., Rees, D. N., et al. (2016). 91 Mass spectrometry imaging (MSI) of microbiome-metabolome interactions in colorectal cancer. *Gastroenterology*, 150(4), S23. doi: 10.1016/S0016-5085(16)30202-5

26. Allali, I., Delgado, S., Marron, P. I., Astudillo, A., Yeh, J. J., Ghazal, H., et al. (2015). Gut microbiome compositional and functional differences between tumor and non-tumor adjacent tissues from cohorts from the US and Spain. *Gut Microbes*, 6(3), 161-172. doi: 10.1080/19490976.2015.1039223

27. Blackett, K. L., Siddhi, S. S., Cleary, S., Steed, H., Miller, M. H., Macfarlane, S., et al. (2013). Oesophageal bacterial biofilm changes in gastro-oesophageal reflux disease, Barrett's and oesophageal carcinoma: association or causality?. *Aliment. Pharmacol. Ther*., 37(11), 1084-1092. doi: 10.1111/apt.12317

28. Brauner, A., Brandt, L., Frisan, T., Thelestam, M., and Ekbom, A. (2010). Is there a risk of cancer development after *Campylobacter* infection?. *Scand. J. Gastroenterol*., 45(7-8), 893-897. doi: 10.3109/00365521003734133

29. De Savornin Lohman, E., Duijster, J., Groot Koerkamp, B., van der Post, R., Franz, E., Mughini Gras, L., et al. (2020). Severe *Salmonella* spp. Or *Campylobacter* spp. infection and the risk of biliary tract cancer: A population-based study. *Cancers*, 12(11), 3348. doi: 10.3390/cancers12113348

30. Mughini-Gras, L. (2017). "Colon cancer risk in patients with a history of *Campylobacter* infection," in *Proceeding of the 19th International workshop CHRO 2017 on Campylobacter, Helicobacter and Related Organisms*. 2017 [Conference Abstract], (Nantes), 250.

31. Wang, Y., Zhang, C., Hou, S., Wu, X., Liu, J., and Wan, X. (2020). Analyses of potential driver and passenger bacteria in human colorectal cancer. *Cancer Manag. Res*., 12, 11553. doi: 10.2147/CMAR.S275316

32. Wei, A. L., Li, M., Li, G. Q., Wang, X., Hu, W. M., Li, Z. L., et al. (2020). Oral microbiome and pancreatic cancer. *World J. Gastroenterol*., 26(48), 7679. doi: 10.3748/wjg.v26.i48.7679

33. Wu, N., Yang, X., Zhang, R., Li, J., Xiao, X., Hu, Y., et al. (2013). Dysbiosis signature of fecal microbiota in colorectal cancer patients. *Microb. Ecol*., 66(2), 462-470. doi: 10.1007/s00248-013-0245-9

34. Liu, F., Ma, R., Wang, Y., and Zhang, L. (2018). The clinical importance of *Campylobacter concisus* and other human hosted Campylobacter species. *Front. Cell. Infect. Microbiol*., 8, 243. doi: 10.3389/fcimb.2018.00243

35. Kaakoush, N. O., Deshpande, N. P., Man, S. M., Burgos-Portugal, J. A., Khattak, F. A., Raftery, M. J., et al. (2015). Transcriptomic and proteomic analyses reveal key innate immune signatures in the host response to the gastrointestinal pathogen *Campylobacter concisus*. *Infect. Immun*., 83(2), 832-845. doi: 10.1128/IAI.03012-14

36. Yang, L., Lu, X., Nossa, C. W., Francois, F., Peek, R. M., and Pei, Z. (2009). Inflammation and intestinal metaplasia of the distal esophagus are associated with alterations in the microbiome. *Gastroenterology*, 137(2), 588-597. doi: 10.1053/j.gastro.2009.04.046

37. Lopetuso, L. R., Scaldaferri, F., Petito, V., and Gasbarrini, A. (2013). Commensal *Clostridia*: leading players in the maintenance of gut homeostasis. *Gut Pathog*., 5(1), 1-8. doi: 10.1186/1757-4749-5-23

38. Yamamoto, Y., Itoh, N., Sugiyama, T., and Kurai, H. (2020). Clinical features of *Clostridium* bacteremia in cancer patients: A case series review. *J. Infect. Chemother*., 26(1), 92-94. doi: 10.1016/j.jiac.2019.07.019

39. Di Bella, S., Ascenzi, P., Siarakas, S., Petrosillo, N., and Di Masi, A. (2016). *Clostridium difficile* toxins A and B: insights into pathogenic properties and extraintestinal effects. *Toxins*, 8(5), 134. doi: 10.3390/toxins8050134

40. Ahn, J., Sinha, R., Pei, Z., Dominianni, C., Goedert, J. J., Hayes, R. B., et al. (2013). Abstract 2290: Human gut microbiome and risk of colorectal cancer, a case-control study. [conference abstract]. doi: 10.1158/1538-7445.AM2013-2290

41. Liang, Q., Chiu, J., Chen, Y., Huang, Y., Higashimori, A., Fang, J., et al. Fecal bacteria act as novel biomarkers for noninvasive diagnosis of colorectal cancer. *Clin. Cancer Res*., 23(8), 2061-2070. doi: 10.1158/1078-0432.CCR-16-1599

42. Lu, Y., Zhang, B. Y., Shi, J. S., and Wu, L. Q. (2004). Expression of the bacterial gene in gallbladder carcinoma tissue and bile. *Hepatobiliary & Pancreatic Diseases International: HBPD INT*, 3(1), 133-135.

43. Ohigashi, S., Sudo, K., Kobayashi, D., Takahashi, O., Takahashi, T., Asahara, T., et al. (2013). Occurence of initiation and progression of colorectal cancer in the changed intestinal environment. *J. Clin. Oncol.,* 31(4), 416. doi: 10.1200/jco.2013.31.4_suppl.416

44. Hsieh, Y. Y., Tung, S. Y., Pan, H. Y., Yen, C. W., Xu, H. W., Lin, Y. J., et al. (2018). Increased abundance of *Clostridium* and *Fusobacterium* in gastric microbiota of patients with gastric cancer in Taiwan. *Scientific Rep*., 8(1), 1-11. doi: 10.1038/s41598-017-18596-0

45. Jahani-Sherafat, S., Azimirad, M., Alebouyeh, M., Amoli, H. A., Hosseini, P., Ghasemian-Safaei, H., et al. (2019). The rate and importance of *Clostridium difficile* in colorectal cancer patients. *Gastroenterol. Hepatol. Bed Bench*, 12(4), 358-363.

46. Liang, J. Q., Wong, S. H., Szeto, C. H., Chu, E. S., Lau, H. C., Chen, Y., et al. (2021). Fecal microbial DNA markers serve for screening colorectal neoplasm in asymptomatic subjects. *J. Gastroenterol. Hepatol*., 36(4), 1035-1043. doi: 10.1111/jgh.15171

47. Sawant, M., Baydoun, M., Creusy, C., Chabé, M., Viscogliosi, E., Certad, G., et al. (2020). *Cryptosporidium* and colon cancer: cause or consequence?. *Microorganisms*, 8(11), 1665. doi: 10.3390/microorganisms8111665

48. Kalantari, N., Gorgani-Firouzjaee, T., Ghaffari, S., Bayani, M., Ghaffari, T., and Chehrazi, M. (2020). Association between *Cryptosporidium* infection and cancer: a systematic review and meta-analysis. *Parasitol. Int*., 74, 101979. doi: 10.1016/j.parint.2019.101979

49. Abdou, A. G., Harba, N. M., Afifi, A. F., and Elnaidany, N. F. (2013). Assessment of *Cryptosporidium parvum* infection in immunocompetent and immunocompromised mice and its role in triggering intestinal dysplasia. *Int. J. Infect. Dis*., 17(8), e593-e600. doi: 10.1016/j.ijid.2012.11.023

50. El-Kersh, W. M., Bahaa, E. D. W., El-Deen, S. A. S., Ammar, A. I., and Matar, A. M. (2019). Prolonged *Cryptosporidium parvum* infection can be a risk factor for intestinal malignancy even in immunocompetent host. *Egypt. J. Med. Microbiol*., 28, 63-70.

51. Sulżyc-Bielicka, V., Kołodziejczyk, L., Jaczewska, S., Bielicki, D., Safranow, K., Bielicki, P., et al. (2018). Colorectal cancer and *Cryptosporidium* spp. infection. *PLoS One*, 13(4), e0195834. doi: 10.1371/journal.pone.0195834

52. Sulżyc-Bielicka, V., Kuźna-Grygiel, W., Kołodziejczyk, L., Bielicki, D., Kładny, J., Stępień-Korzonek, M., et al. (2007). Cryptosporidiosis in patients with colorectal cancer. *J. Parasitol*., 93(3), 722-724. doi: 10.1645/GE-1025R1.1

53. Essid, R., Menotti, J., Hanen, C., Aoun, K., and Bouratbine, A. (2018). Genetic diversity of *Cryptosporidium* isolates from human populations in an urban area of Northern Tunisia. *Infect. Genet. Evol*., 58, 237-242. doi: 10.1016/j.meegid.2018.01.004

54. Osman, M., Benamrouz, S., Guyot, K., Baydoun, M., Frealle, E., Chabe, M., et al. (2017). High association of *Cryptosporidium* spp. infection with colon adenocarcinoma in Lebanese patients. *PloS One*, 12(12), e0189422. doi: 10.1371/journal.pone.0189422

55. Sanad, M. M., Thagfan, F. A., Al Olayan, E. M., Almogren, A., Al Hammaad, A., Al-Mawash, A., et al. (2014). Opportunistic coccidian parasites among Saudi cancer patients presenting with diarrhea: prevalence and immune status. *Res. J. Parasitol*., 9(2), 55-63. doi: 10.3923/jp.2014.55.63

56. Sulżyc-Bielicka, V., Kołodziejczyk, L., Jaczewska, S., Bielicki, D., Kładny, J., and Safranow, K. (2012). Prevalence of *Cryptosporidium* sp. in patients with colorectal cancer. *Pol. J. Surg*., 2012. 84(7): p. 348-351. doi: 10.2478/v10035-012-0058-4

57. De Almeida, C. V., Taddei, A., and Amedei, A. (2018). The controversial role of *Enterococcus faecalis* in colorectal cancer. *Therap. Adv. Gastroenterol*., 11, 1756284818783606. doi: 10.1177/1756284818783606

58. De Almeida, C. V., Lulli, M., di Pilato, V., Schiavone, N., Russo, E., Nannini, G., et al. (2019). Differential responses of colorectal cancer cell lines to *Enterococcus faecalis'* strains isolated from healthy donors and colorectal cancer patients. *J. Clin. Med*., 8(3), 388. doi: 10.3390/jcm8030388

59. Strickertsson, J. A., Desler, C., Martin-Bertelsen, T., Machado, A. M. D., Wadstrøm, T., Winther, O., et al. (2013). *Enterococcus faecalis* infection causes inflammation, intracellular oxphos-independent ROS production, and DNA damage in human gastric cancer cells. *PloS One*, 8(4), e63147. doi: 10.1371/journal.pone.0063147

60. Balamurugan, R., Rajendiran, E., George, S., Samuel, G. V., and Ramakrishna, B. S. (2008). Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, *Desulfovibrio* and *Enterococcus faecalis* in the feces of patients with colorectal cancer. *J. Gastroenterol. Hepatol*., 23(8pt1), 1298-1303. doi: 10.1111/j.1440-1746.2008.05490.x

61. Corredoira, J., García-País, M. J., Coira, A., Rabuñal, R., García-Garrote, F., Pita, J., et al. (2015). Differences between endocarditis caused by *Streptococcus bovis* and *Enterococcus* spp. and their association with colorectal cancer. *Eur. J. Clin. Microbiol. Infect. Dis*., 34(8), 1657-1665. doi: 10.1007/s10096-015-2402-1

62. Geravand, M., Fallah, P., Yaghoobi, M. H., Soleimanifar, F., Farid, M., Zinatizadeh, N., et al. (2019). Investigation of Enterococcus faecalis population in patients with polyp and colorectal cancer in comparison of healthy individuals. *Arq. Gastroenterol*., 56, 141-145. doi: 10.1590/S0004-2803.201900000-28

63. D'asheesh, T. I. A., Hussen, B. M., Al-Marzoqi, A. H., and Ghasemian, A. (2021). Assessment of oncogenic role of intestinal microbiota in colorectal cancer patients. *J. Gastrointest. Cancer*, 52(3), 1016-1021. doi: 10.1007/s12029-020-00531-8

64. Pericas, J. M., Corredoira, J., Moreno, A., García-País, M. J., Falces, C., Rabuñal, R., et al. (2017). Relationship between *Enterococcus faecalis* infective endocarditis and colorectal neoplasm: preliminary results from a cohort of 154 patients. *Rev. Esp. Cardiol*. (English Edition), 70(6), 451-458. doi: 10.1016/j.rec.2016.10.013

65. Rezasoltani, S., Aghdaei, H. A., Dabiri, H., Sepahi, A. A., Modarressi, M. H., and Mojarad, E. N. (2018). The association between fecal microbiota and different types of colorectal polyp as precursors of colorectal cancer. *Microb. Pathog.*, 124, 244-249. doi: 10.1016/j.micpath.2018.08.035

66. Shoji, M., Sasaki, Y., Abe, Y., Nishise, S., Yaoita, T., Yagi, M., et al. (2021). Characteristics of the gut microbiome profile in obese patients with colorectal cancer. *JGH Open*, 5(4), 498-507. doi: 10.1002/jgh3.12529

67. Wang, T., Cai, G., Qiu, Y., Fei, N., Zhang, M., Pang, X., et al. (2012). Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *ISME J*., 6(2), 320-329. doi: 10.1038/ismej.2011.109

68. Zhou, Y., He, H., Xu, H., Li, Y., Li, Z., Du, Y., et al. (2016). Association of oncogenic bacteria with colorectal cancer in South China. *Oncotarget*, 7(49), 80794. doi: 10.18632/oncotarget.13094

69. Kaper, J. B., Nataro, J. P., and Mobley, H. L. (2004). Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol*., 2(2), 123-140. doi: 10.1038/nrmicro818

70. Dalmasso, G., Cougnoux, A., Delmas, J., Darfeuille-Michaud, A., and Bonnet, R. (2014). The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment. *Gut Microbes*, 5(5), 675-680. doi: 10.4161/19490976.2014.969989

71. Faïs, T., Delmas, J., Barnich, N., Bonnet, R., and Dalmasso, G. (2018). Colibactin: more than a new bacterial toxin. *Toxins*, 10(4), 151. doi: 10.3390/toxins10040151

72. Sheng, Q., Du, H., Cheng, X., Cheng, X., Tang, Y., Pan, L., et al. (2019). Characteristics of fecal gut microbiota in patients with colorectal cancer at different stages and different sites. *Oncol. Lett*., 18(5), 4834-4844. doi: 10.3892/ol.2019.10841

73. Tang, L., Zhou, Y. J., Zhu, S., Liang, G. D., Zhuang, H., Zhao, M. F., et al. (2020). *E. coli* diversity: low in colorectal cancer. *BMC Med. Genom*., 13(1), 1-17. doi: 10.1186/s12920-020-0704-3

74. Wang, Y., Zhang, Y., Qian, Y., Xie, Y. H., Jiang, S. S., Kang, Z. R., et al. (2021). Alterations in the oral and gut microbiome of colorectal cancer patients and association with host clinical factors. *Int. J. Cancer,* 149(4), 925-935. doi: 10.1002/ijc.33596

75. Bonnet, M., Buc, E., Sauvanet, P., Darcha, C., Dubois, D., Pereira, B., et al. (2014). Colonization of the human gut by *E. coli* and colorectal cancer risk. *Clin. Cancer Res*., 20(4), 859-867. doi: 10.1158/1078-0432.CCR-13-1343

76. Buc, E., Dubois, D., Sauvanet, P., Raisch, J., Delmas, J., Darfeuille-Michaud, A., et al. (2013). High prevalence of mucosa-associated *E. coli* producing cyclomodulin and genotoxin in colon cancer. *PloS One*, 8(2), e56964. doi: 10.1371/journal.pone.0056964

77. Iyadorai, T., Mariappan, V., Vellasamy, K. M., Wanyiri, J. W., Roslani, A. C., Lee, G. K., et al. (2020). Prevalence and association of pks+ *Escherichia coli* with colorectal cancer in patients at the University Malaya Medical Centre, Malaysia. *PloS One*, 15(1), e0228217. doi: 10.1371/journal.pone.0228217

78. Kohoutova, D., Smajs, D., Moravkova, P., Cyrany, J., Moravkova, M., Forstlova, M., et al. (2014). *Escherichia coli* strains of phylogenetic group B2 and D and bacteriocin production are associated with advanced colorectal neoplasia. *BMC Infect. Dis*., 14(1), 1-8. doi: 10.1186/s12879-014-0733-7

79. Pleguezuelos-Manzano, C., Puschhof, J., Huber, A. R., van Hoeck, A., Wood, H. M., Nomburg, J., et al. (2020). Mutational signature in colorectal cancer caused by genotoxic pks+ *E. coli*. *Nature*, 580(7802), 269-273. doi: 10.1038/s41586-020-2080-8

80. Rezasoltani, S., Asadzadeh, A. H., Dabiri, H., Akhavan, S. A., Modarressi, M. H., and Nazemalhosseini, M. E. (2020). Investigating the number of *Lactobaccilus*, *Escherichia coli* and *Prevotella* in fecal microbiota of adenomatous polyposis and colorectal cancer patients. *Tehran Univ. Med. J*., 78(3), 137-143.

81. Shimpoh, T., Hirata, Y., Ihara, S., Suzuki, N., Kinoshita, H., Hayakawa, Y., et al. (2017). Prevalence of pks-positive *Escherichia coli* in Japanese patients with or without colorectal cancer. *Gut Pathog.*, 9(1), 1-8. doi: 10.1186/s13099-017-0185-x

82. Tunsjø, H. S., Gundersen, G., Rangnes, F., Noone, J. C., Endres, A., and Bemanian, V. (2019). Detection of Fusobacterium nucleatum in stool and colonic tissues from Norwegian colorectal cancer patients. *Eur. J. Clin. Microbiol. Infect. Dis*., 38(7), 1367-1376. doi: 10.1007/s10096-019-03562-7

83. Yoshikawa, Y., Tsunematsu, Y., Matsuzaki, N., Hirayama, Y., Higashiguchi, F., Sato, M., et al. (2020). Characterization of colibactin-producing *Escherichia coli* isolated from Japanese patients with colorectal cancer. *Jpn. J. Infect. Dis*., 73(6), 437-442. doi: 10.7883/yoken.JJID.2020.066

84. Zarei, O., Arabestan, M. R., Majlesi, A., Mohammadi, Y., and Alikhani, M. Y. (2019). Determination of virulence determinants of *Escherichia coli* strains isolated from patients with colorectal cancer compared to the healthy subjects. *Gastroenterol. Hepatol. Bed Bench*, 12(1), 52.

85. Tsuchiya, Y., Loza, E., Villa-Gomez, G., Trujillo, C. C., Baez, S., Asai, T., et al. (2018). Metagenomics of microbial communities in gallbladder bile from patients with gallbladder cancer or cholelithiasis. *Asian Pac. J. Cancer Prev*., 19(4), 961. doi: 10.22034/APJCP.2018.19.4.961

86. Yang, N. Y., Zhang, Q., Li, J. L., Yang, S. H., and Shi, Q. (2014). Progression of periodontal inflammation in adolescents is associated with increased number of *Porphyromonas gingivalis*, *Prevotella intermedia*, *Tannerella forsythensis*, and *Fusobacterium nucleatum*. *Int. J. Paediatr. Dent*., 24(3), 226-233. doi: 10.1111/ipd.12065

87. Lee, S. A., Liu, F., Riordan, S. M., Lee, C. S., and Zhang, L. (2019). Global investigations of Fusobacterium nucleatum in human colorectal cancer. *Front. Oncol*., 9, 566. doi: 10.3389/fonc.2019.00566

88. Vargo, D., Moskovitz, M., and Floch, M. H. (1980). Faecal bacterial flora in cancer of the colon. *Gut*, 21(8), 701-705. doi: 10.1136/gut.21.8.701

89. Castellarin, M., Warren, R. L., Freeman, J. D., Dreolini, L., Krzywinski, M., Strauss, J., et al. (2012). *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res*., 22(2), 299-306. doi: 10.1101/gr.126516.111

90. Flanagan, L., Schmid, J., Ebert, M., Soucek, P., Kunicka, T., Liska, V., et al. (2014). *Fusobacterium nucleatum* associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. *Eur. J. Clin. Microbiol. Infect. Dis*., 33(8), 1381-1390. doi: 10.1007/s10096-014-2081-3

91. Ito, M., Kanno, S., Nosho, K., Sukawa, Y., Mitsuhashi, K., Kurihara, H., et al. (2015). Association of *Fusobacterium nucleatum* with clinical and molecular features in colorectal serrated pathway. *Int. J. Cancer*, 137(6), 1258-1268. doi: 10.1002/ijc.29488

92. Mira-Pascual, L., Cabrera-Rubio, R., Ocon, S., Costales, P., Parra, A., Suarez, A., et al. (2015). Microbial mucosal colonic shifts associated with the development of colorectal cancer reveal the presence of different bacterial and archaeal biomarkers. *J. Gastroenterol*., 50(2), 167-179. doi: 10.1007/s00535-014-0963-x

93. Amitay, E. L., Werner, S., Vital, M., Pieper, D. H., Höfler, D., Gierse, I. J., et al. (2017). *Fusobacterium* and colorectal cancer: causal factor or passenger? Results from a large colorectal cancer screening study. *Carcinogenesis*, 38(8), 781-788. doi: 10.1093/carcin/bgx053

94. Drewes, J. L., White, J. R., Dejea, C. M., Fathi, P., Iyadorai, T., Vadivelu, J., et al. (2017). High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. *NPJ Biofilms Microbiomes*, 3(1), 1-12. doi: 10.1038/s41522-017-0040-3

95. Eklöf, V., Löfgren-Burström, A., Zingmark, C., Edin, S., Larsson, P., Karling, P., et al. (2017). Cancer-associated fecal microbial markers in colorectal cancer detection. *Int. J. Cancer*, 141(12), 2528-2536. doi: 10.1002/ijc.31011

96. Wang, H. F., Li, L. F., Guo, S. H., Zeng, Q. Y., Ning, F., Liu, W. L., et al. (2016). Evaluation of antibody level against *Fusobacterium nucleatum* in the serological diagnosis of colorectal cancer. *Scientific Rep*., 6(1), 1-10. doi: 10.1038/srep33440

97. Xie, Y.H. and Fang, J.Y. (2016). Single fecal abundance of *Fusobacterium nucleatum* for early- and late-stage detection of colorectal cancer. *J. Dig. Dis*., 2016. 17: p. 3.

98. Yu, J., Chen, Y., Fu, X., Zhou, X., Peng, Y., Shi, L., et al. (2016). Invasive *Fusobacterium nucleatum* may play a role in the carcinogenesis of proximal colon cancer through the serrated neoplasia pathway. *Int. J. Cancer*, 139(6), 1318-1326. doi: 10.1002/ijc.30168

99. Scott, A., Alexander, J. L., Marchesi, J., and Kinross, J. M. (2017). Modification of the appendix microbiome in acute appendicitis and colon adenocarcinoma. *Gastroenterol.*, 152(5), S1228.

100. Suehiro, Y., Sakai, K., Nishioka, M., Hashimoto, S., Takami, T., Higaki, S., et al. (2017). Highly sensitive stool DNA testing of *Fusobacterium nucleatum* as a marker for detection of colorectal tumours in a Japanese population. *Ann. Clin. Biochem*., 54(1), 86-91. doi: 10.1177/0004563216643970

101. Yamamura, K., Baba, Y., Miyake, K., Nakamura, K., Shigaki, H., Mima, K., et al. (2017). *Fusobacterium nucleatum* in gastroenterological cancer: Evaluation of measurement methods using quantitative polymerase chain reaction and a literature review. *Oncol. Lett*., 14(6), 6373-6378. doi: 10.3892/ol.2017.7001

102. Yoon, H., Kim, N., Park, J. H., Kim, Y. S., Lee, J., Kim, H. W., et al. (2017). Comparisons of gut microbiota among healthy control, patients with conventional adenoma, sessile serrated adenoma, and colorectal cancer. *J. Cancer Prev*., 22(2), 108. doi: 10.15430/JCP.2017.22.2.108

103. Yu, J., Feng, Q., Wong, S. H., Zhang, D., Liang, Q., Qin, Y., et al. (2017). Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut*, 66(1), 70-78. doi: 10.1136/gutjnl-2015-309800

104. Butt, J., Jenab, M., Pawlita, M., Overvad, K., Tjonneland, A., Olsen, A., et al. (2019). Antibody responses to *Fusobacterium nucleatum* proteins in prediagnostic blood samples are not associated with risk of developing colorectal cancer. *Cancer Epidemiol. Prev. Biomark*., 28(9), 1552-1555. doi: 10.1158/1055-9965.EPI-19-0313

105. Repass, J. (2018). Replication Study: *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Elife*, 7, e25801. doi: 10.7554/eLife.25801

106. Russo, E., Bacci, G., Chiellini, C., Fagorzi, C., Niccolai, E., Taddei, A., et al. (2018). Preliminary comparison of oral and intestinal human microbiota in patients with colorectal cancer: a pilot study. *Front. Microbiol*., 8, 2699. doi: 10.3389/fmicb.2017.02699

107. Chen, X. H., Wang, A., Chu, A. N., Gong, Y. H., and Yuan, Y. (2019). Mucosa-associated microbiota in gastric cancer tissues compared with non-cancer tissues. *Front. Microbiol*., 10, 1261. doi: 10.3389/fmicb.2019.01261

108. De Carvalho, A. C., de Mattos Pereira, L., Datorre, J. G., Dos Santos, W., Berardinelli, G. N., Matsushita, M. D. M., et al. (2019). Microbiota profile and impact of *Fusobacterium nucleatum* in colorectal cancer patients of Barretos Cancer Hospital. *Front. Oncol*., 9, 813. doi: 10.3389/fonc.2019.00813

109. Kageyama, S., Takeshita, T., Takeuchi, K., Asakawa, M., Matsumi, R., Furuta, M., et al. (2019). Characteristics of the salivary microbiota in patients with various digestive tract cancers. *Front. Microbiol*., 10, 1780. doi: 10.3389/fmicb.2019.01780

110. Saito, K., Koido, S., Odamaki, T., Kajihara, M., Kato, K., Horiuchi, S., et al. (2019). Metagenomic analyses of the gut microbiota associated with colorectal adenoma. *PLoS One*, 14(2), e0212406. doi: 10.1371/journal.pone.0212406

111. Yachida, S., Mizutani, S., Shiroma, H., Shiba, S., Nakajima, T., Sakamoto, T., et al. (2019). Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat. Med*., 25(6), 968-976. doi: 10.1038/s41591-019-0458-7

112. Alkharaan, H., Lu, L., Gabarrini, G., Halimi, A., Ateeb, Z., Sobkowiak, M. J., et al. (2020). Circulating and salivary antibodies to *Fusobacterium nucleatum* are associated with cystic pancreatic neoplasm malignancy. *Front. Immunol*., 11, 2003. doi: 10.3389/fimmu.2020.02003

113. Boehm, E. T., Thon, C., Kupcinskas, J., Steponaitiene, R., Skieceviciene, J., Canbay, A., et al. (2020). *Fusobacterium nucleatum* is associated with worse prognosis in Lauren's diffuse type gastric cancer patients. *Sci. Rep*., 10(1), 1-12. doi: 10.1038/s41598-020-73448-8

114. Gantuya, B., El Serag, H. B., Matsumoto, T., Ajami, N. J., Uchida, T., Oyuntsetseg, K., et al. (2020). Gastric mucosal microbiota in a Mongolian population with gastric cancer and precursor conditions. *Aliment. Pharmacol. Ther*., 51(8), 770-780. doi: 10.1111/apt.15675

115. Kashani, N., Bezmin Abadi, A. T., Rahimi, F., and Forootan, M. (2020). FadA-positive *Fusobacterium nucleatum* is prevalent in biopsy specimens of Iranian patients with colorectal cancer. *New Microbes New Infect*., 2020. 34. doi: 10.1016/j.nmni.2020.100651

116. Reynolds, I. S., Thomas, V., O'Connell, E., Fichtner, M., McNamara, D. A., Kay, E. W., et al. (2020). Mucinous adenocarcinoma of the rectum: a whole genome sequencing study. *Front. Oncol*., 10, 1682. doi: 10.3389/fonc.2020.01682

117. Zhang, C., Thakkar, P. V., Sharma, P., Vennelaganti, S., Betel, D., and Shah, M. A. (2019). Abstract 2826: Understanding associations among local microbiome, immune response, and efficacy of immunotherapy in esophageal cancer. *Cancer Res.,* 79(13). doi: 10.1158/1538-7445.AM2019-2826

118. Kawasaki, M., Ikeda, Y., Ikeda, E., Takahashi, M., Tanaka, D., Nakajima, Y., et al. (2021). Oral infectious bacteria in dental plaque and saliva as risk factors in patients with esophageal cancer. *Cancer*, 127(4), 512-519. doi: 10.1002/cncr.33316

119. Kurt, M., and Yumuk, Z. (2021). Diagnostic accuracy of *Fusobacterium nucleatum* IgA and IgG ELISA test in colorectal cancer. *Sci. Rep*., 11(1), 1-6. doi: 10.1038/s41598-021-81171-1

120. Pignatelli, P., Iezzi, L., Pennese, M., Raimondi, P., Cichella, A., Bondi, D., et al. (2021). The potential of colonic tumor tissue *Fusobacterium nucleatum* to predict staging and its interplay with oral abundance in colon cancer patients. *Cancers*, 13(5), 1032. doi: 10.3390/cancers13051032

121. Eisele, Y., Mallea, P. M., Gigic, B., Stephens, W. Z., Warby, C. A., Buhrke, K., et al. (2021). *Fusobacterium nucleatum* and clinicopathologic features of colorectal cancer: results from the ColoCare study. *Clin. Colorectal Cancer*, 20(3), e165-e172. doi: 10.1016/j.clcc.2021.02.007

122. Gethings-Behncke, C., Coleman, H. G., Jordao, H. W., Longley, D. B., Crawford, N., Murray, L. J., et al. (2020). *Fusobacterium nucleatum* in the colorectum and its association with cancer risk and survival: a systematic review and meta-analysis. *Cancer Epidemiol. Prev. Biomark*., 29(3), 539-548. doi: 10.1158/1055-9965.EPI-18-1295

123. Janati, A. I., Karp, I., Laprise, C., Sabri, H., and Emami, E. (2020). Detection of *Fusobaterium nucleatum* in feces and colorectal mucosa as a risk factor for colorectal cancer: a systematic review and meta-analysis. *Syst. Rev*., 9(1), 1-15. doi: 10.1186/s13643-020-01526-z

124. Hashemi Goradel, N., Heidarzadeh, S., Jahangiri, S., Farhood, B., Mortezaee, K., Khanlarkhani, N., et al. (2019). *Fusobacterium nucleatum* and colorectal cancer: A mechanistic overview. *J. Cell. Physiol*., 234(3), 2337-2344. doi: 10.1002/jcp.27250

125. Mysak, J., Podzimek, S., Sommerova, P., Lyuya-Mi, Y., Bartova, J., Janatova, T., et al. (2014). *Porphyromonas gingivalis*: major periodontopathic pathogen overview. *J. Immunol. Res*., 2014. doi: 10.1155/2014/476068

126. Corbella, S., Veronesi, P., Galimberti, V., Weinstein, R., Del Fabbro, M., and Francetti, L. (2018). Is periodontitis a risk indicator for cancer? A meta-analysis. *PloS One*, 13(4), e0195683. doi: 10.1371/journal.pone.0195683

127. Liu, X. B., Gao, Z. Y., Sun, C. T., Wen, H., Gao, B., Li, S. B., et al. (2019). The potential role of *P. gingivalis* in gastrointestinal cancer: a mini review. *Infect. Agents Cancer*, 14(1), 1-7. doi: 10.1186/s13027-019-0239-4

128. Ahn, J., Segers, S., and Hayes, R. B. (2012). Periodontal disease, *Porphyromonas gingivalis* serum antibody levels and orodigestive cancer mortality. *Carcinogenesis*, 33(5), 1055-1058. doi: 10.1093/carcin/bgs112

129. Fan, X., Alekseyenko, A. V., Wu, J., Peters, B. A., Jacobs, E. J., Gapstur, S. M., et al. (2018). Human oral microbiome and prospective risk for pancreatic cancer: a population-based nested case-control study. *Gut*, 67(1), 120-127. doi: 10.1136/gutjnl-2016-312580

130. Gao, S., Li, S., Ma, Z., Liang, S., Shan, T., Zhang, M., et al. (2016). Presence of *Porphyromonas gingivalis* in esophagus and its association with the clinicopathological characteristics and survival in patients with esophageal cancer. *Infect. Agents Cancer*, 11(1), 1-9. doi: 10.1186/s13027-016-0049-x

131. Gao, S. G., Yang, J. Q., Ma, Z. K., Yuan, X., Zhao, C., Wang, G. C., et al. (2018). Preoperative serum immunoglobulin G and A antibodies to *Porphyromonas gingivalis* are potential serum biomarkers for the diagnosis and prognosis of esophageal squamous cell carcinoma. *BMC Cancer*, 18(1), 1-8. doi: 10.1186/s12885-017-3905-1

132. Michaud, D. S., Izard, J., Wilhelm-Benartzi, C. S., You, D. H., Grote, V. A., Tjønneland, A., et al. (2013). Plasma antibodies to oral bacteria and risk of pancreatic cancer in a large European prospective cohort study. *Gut*, 62(12), 1764-1770. doi: 10.1136/gutjnl-2012-303006

133. Peters, B. A., Wu, J., Pei, Z., Yang, L., Purdue, M. P., Freedman, N. D., et al. (2017). Oral microbiome composition reflects prospective risk for esophageal cancers. *Cancer Res*., 77(23), 6777-6787. doi: 10.1158/0008-5472.CAN-17-1296

134. Wang, X., Jia, Y., Wen, L., Mu, W., Wu, X., Liu, T., et al. (2021). *Porphyromonas gingivalis* promotes colorectal carcinoma by activating the hematopoietic NLRP3 inflammasome. *Cancer Res*., 81(10), 2745-2759. doi: 10.1158/0008-5472.CAN-20-3827

135. Sinha, R., Ahn, J., Sampson, J. N., Shi, J., Yu, G., Xiong, X., et al. (2016). Fecal microbiota, fecal metabolome, and colorectal cancer interrelations. *PloS One*, 11(3), e0152126. doi: 10.1371/journal.pone.0152126

136. Scanu, T., Spaapen, R. M., Bakker, J. M., Pratap, C. B., Wu, L. E., Hofland, I., et al. (2015). *Salmonella* manipulation of host signaling pathways provokes cellular transformation associated with gallbladder carcinoma. Cell Host Microbe, 17(6), 763-774.

137. Chang, R., Wei, J. C. C., Lin, M. C., and Hung, Y. M. (2021). The potential role of nontyphoidal salmonellosis in gastric cancer: a nationwide matched cohort study. *Gastric Cancer*, 24(2), 292-301. doi: s10120-020-01132-x

138. Duijster, J. W., Hansen, J. V., Franz, E., Neefjes, J. J., Frisch, M., Mughini-Gras, L., et al. (2021). Association between *Salmonella* infection and colon cancer: a nationwide registry-based cohort study. *Epidemiol. Infect*., 149, E56. doi: 10.1017/S0950268821000285

139. Iyer, P., Barreto, S. G., Sahoo, B., Chandrani, P., Ramadwar, M. R., Shrikhande, S. V., et al. (2016). Non-typhoidal *Salmonella* DNA traces in gallbladder cancer. *Infect. Agents Cancer*, 11(1), 1-4. doi: 10.1186/s13027-016-0057-x

140. Kato, I., Boleij, A., Kortman, G. A., Roelofs, R., Djuric, Z., Severson, R. K., et al. (2013). Partial associations of dietary iron, smoking and intestinal bacteria with colorectal cancer risk. *Nutr. Cancer*, 65(2), 169-177. doi: 10.1080/01635581.2013.748922

141. Lu, R., Bosland, M., Xia, Y., Zhang, Y. G., Kato, I., and Sun, J. (2017). Presence of *Salmonella* AvrA in colorectal tumor and its precursor lesions in mouse intestine and human specimens. *Oncotarget*, 8(33), 55104. doi: 10.18632/oncotarget.19052

142. Mughini-Gras, L., Schaapveld, M., Kramers, J., Mooij, S., Neefjes-Borst, E. A., Pelt, W. V., et al. (2018). Increased colon cancer risk after severe *Salmonella* infection. *PloS One*, 13(1), e0189721. doi: 10.1371/journal.pone.0189721

143. Risques, R. A., and Kennedy, S. R. (2018). Aging and the rise of somatic cancer-associated mutations in normal tissues. *PLoS Genet*., 14(1), e1007108. doi: 10.1371/journal.pgen.1007108

144. Di Domenico, E. G., Cavallo, I., Pontone, M., Toma, L., and Ensoli, F. (2017). Biofilm producing *Salmonella* typhi: chronic colonization and development of gallbladder cancer. *Int. J. Mol. Sci*., 18(9), 1887. doi: 10.3390/ijms18091887

145. Capoor, M. R., Nair, D., Khanna, G., Krishna, S. V., Chintamani, M. S., and Aggarwal, P. (2008). Microflora of bile aspirates in patients with acute cholecystitis with or without cholelithiasis: a tropical experience. *Braz. J. Infect. Dis.*, 12, 222-225. doi: 10.1590/S1413-86702008000300012

146. Caygill, C. P., Hill, M. J., Braddick, M., and Sharp, J. C. (1994). Cancer mortality in chronic typhoid and paratyphoid carriers. *Lancet*, 343(8889), 83-84. doi: 10.1016/S0140-6736(94)90816-8

147. Csendes, A., Becerra, M., Burdiles, P., Demian, I., Bancalari, K., and Csendes, P. (1994). Bacteriological studies of bile from the gallbladder in patients with carcinoma of the gallbladder, cholelithiasis, common bile duct stones and no gallstones disease. *Eur. J. Surg*., 160(6-7), 363-367.

148. Dutta, U., Garg, P. K., Kumar, R., and Tandon, R. K. (2000). Typhoid carriers among patients with gallstones are at increased risk for carcinoma of the gallbladder. *Am. J. Gastroenterol*., 95(3), 784-787. doi: 10.1016/S0002-9270(99)00919-3

149. Hazrah, P., Oahn, K. T. H., Tewari, M., Pandey, A. K., Kumar, K., Mohapatra, T. M., et al. (2004). The frequency of live bacteria in gallstones. HPB, 6(1), 28-32. doi: 10.1080/13651820310025192

150. Mellemgaard, A. and K. Gaarslev, Risk of hepatobiliary cancer in carriers of *Salmonella* typhi*. J. Nat. Cancer Institute*, 1988. 80(4): p. 288. doi: 10.1093/jnci/80.4.288

151. Nath, G., Singh, H., and Shukla, V. K. (1997). Chronic typhoid carriage and carcinoma of the gallbladder. *Eur. J. Cancer Prev*., 6(6), 557-559. doi: 10.1097/00008469-199712000-00011

152. Nath, G., Singh, Y. K., Kumar, K., Gulati, A. K., Shukla, V. K., Khanna, A. K., et al. (2008). Association of carcinoma of the gallbladder with typhoid carriage in a typhoid endemic area using nested PCR. *J. Infect. Dev. Ctries*., 2(04), 302-307. doi: 10.3855/jidc.226

153. Pandey, M., and Shukla, V. K. (2003). Lifestyle, parity, menstrual and reproductive factors and risk of gallbladder cancer. *Eur. J. Cancer Prev*., 12(4), 269-272.

154. Roa, I., Ibacache, G., Carvallo, J., Melo, A., Araya, J., De Aretxabala, X., et al. (1999). Microbiological study of gallbladder bile in a high risk zone for gallbladder cancer. *Rev. Med.Chile*, 127(9), 1049-1055.

155. Safaeian, M., Gao, Y. T., Sakoda, L. C., Quraishi, S. M., Rashid, A., Wang, B. S., et al. (2011). Chronic typhoid infection and the risk of biliary tract cancer and stones in Shanghai, China. *Infect. Agents Cancer*, 6(1), 1-3. doi: 10.1186/1750-9378-6-6

156. Serra, I., Yamamoto, M., Calvo, A., Cavada, G., Báez, S., Endoh, K., et al. (2002). Association of chili pepper consumption, low socioeconomic status and longstanding gallstones with gallbladder cancer in a Chilean population. *Int. J. Cancer*, 102(4), 407-411. doi: 10.1002/ijc.10716

157. Sharma, V., Chauhan, V. S., Nath, G., Kumar, A., and Shukla, V. K. (2007). Role of bile bacteria in gallbladder carcinoma. *Hepato-gastroenterology*, 54(78), 1622-1625.

158. Shukla, V. K., Singh, H., Pandey, M., Upadhyay, S. K., and Nath, G. (2000). Carcinoma of the gallbladder—is it a sequel of typhoid?. *Dig. Dis. Sci*., 45(5), 900-903. doi: 10.1023/A:1005564822630

159. Singh, H., Pandey, M., amd Shukla, V. K. (1996). *Salmonella* carrier state, chronic bacterial infection and gallbladder carcinogenesis. *Eur. J. Cancer Prev*., 5(2), 144.

160. Strom, B. L., Soloway, R. D., Rios-Dalenz, J. L., Rodriguez-Martinez, H. A., West, S. L., Kinman, J. L., et al. (1995). Risk factors for gallbladder cancer. An international collaborative case–control study. *Cancer*, 76(10), 1747-1756. doi: 10.1002/1097-0142(19951115)76:10<1747::AID-CNCR2820761011>3.0.CO;2-L

161. Tewari, M., Mishra, R. R., and Shukla, H. S. (2010). *Salmonella* typhi and gallbladder cancer: report from an endemic region. *Hepatobiliary Pancreat. Dis. Int*., 9(5), 524-530.

162. Vaishnavi, C., Kochhar, R., Singh, G., Kumar, S., Singh, S., and Singh, K. (2005). Epidemiology of typhoid carriers among blood donors and patients with biliary, gastrointestinal and other related diseases. *Microbiol. Immunol*., 49(2), 107-112. doi: 10.1111/j.1348-0421.2005.tb03709.x

163. Welton, J., Marr, J., and Friedman, S. (1979). Association between hepatobiliary cancer and typhoid carrier state. *Lancet*, 313(8120), 791-794. doi: 10.1016/S0140-6736(79)91315-1

164. Yagyu, K., Lin, Y., Obata, Y., Kikuchi, S., Ishibashi, T., Kurosawa, M., et al. (2004). Bowel movement frequency, medical history and the risk of gallbladder cancer death: a cohort study in Japan. *Cancer Sci*., 95(8), 674-678. doi: 10.1111/j.1349-7006.2004.tb03328.x

165. Nagaraja, V., and Eslick, G. D. (2014). Systematic review with meta-analysis: the relationship between chronic *Salmonella* typhi carrier status and gall-bladder cancer. *Aliment. Pharmacol. Ther*., 39(8), 745-750. doi: 10.1111/apt.12655

166. Koshiol, J., Wozniak, A., Cook, P., Adaniel, C., Acevedo, J., Azócar, L., et al. (2016). *Salmonella* enterica serovar Typhi and gallbladder cancer: a case–control study and meta-analysis. *Cancer Med*., 5(11), 3310-3235. doi: 10.1002/cam4.915

167. Elbaz, T., and Esmat, G. (2013). Hepatic and intestinal schistosomiasis. *J. Adv. Res*., 4(5), 445-452. doi: 10.1016/j.jare.2012.12.001

168. Cancer, I.A.f.R.o., Infections with Schistosomes, Liver Flukes and Helicobacter pylori. Evaluation of Carcinogenic risks to Humans. IARC: Lyon, 1994.

169. Roderfeld, M., Padem, S., Lichtenberger, J., Quack, T., Weiskirchen, R., Longerich, T., et al. (2020). *Schistosoma mansoni* Egg–Secreted Antigens Activate Hepatocellular Carcinoma–Associated Transcription Factors c-Jun and STAT3 in Hamster and Human Hepatocytes. *Hepatology*, 72(2), 626-641. doi: 10.1002/hep.30192

170. Hamid, H. K. (2019). Schistosoma japonicum–associated colorectal cancer: A review. *Am. J. Trop. Med. Hyg*., 100(3), 501. doi: 10.4269/ajtmh.18-0807

171. Abruzzi, A., Fried, B., and Alikhan, S. B. (2016). Coinfection of *Schistosoma* species with hepatitis B or hepatitis C viruses. *Adv. Parasitol*., 91, 111-231. doi: 10.1016/bs.apar.2015.12.003

172. Chen, M. G. (2014). Assessment of morbidity due to *Schistosoma japonicum* infection in China. *Inf. Dis. Poverty*, 3(1), 1-16. doi: 10.1186/2049-9957-3-6

173. Qiu, D. C., Hubbard, A. E., Zhong, B., Zhang, Y., and Spear, R. C. (2005). A matched, case–control study of the association between *Schistosoma japonicum* and liver and colon cancers, in rural China. *Ann. Trop. Med. Parasitol*., 99(1), 47-52. doi: 10.1179/136485905X19883

174. Xu, Z., and Su, D. L. (1984). Schistosoma japonicum and colorectal cancer: an epidemiological study in the People's Republic of China. *Int. J. Cancer*, 34(3), 315-318. doi: 10.1002/ijc.2910340305

175. En-Sheng, Z., Cancer of the colon and schistosomiasis. 1981, SAGE Publications Sage UK: London, England.

176. Iida, F., Iida, R., Kamijo, H., Takaso, K., Miyazaki, Y., Funabashi, W., et al. (1999). Chronic Japanese schistosomiasis and hepatocellular carcinoma: ten years of follow-up in Yamanashi Prefecture, Japan. *Bull. World Health Organ*., 77(7), 573.

177. Madbouly, K. M., Senagore, A. J., Mukerjee, A., Hussien, A. M., Shehata, M. A., Navine, P., et al. (2007). Colorectal cancer in a population with endemic *Schistosoma mansoni*: is this an at-risk population?. *Int. J. Colorectal Dis*., 22(2), 175-181. doi: 10.1007/s00384-006-0144-3

178. Toda, K. S., Kikuchi, L., Chagas, A. L., Tanigawa, R. Y., Paranaguá-Vezozzo, D. C., Pfiffer, T., et al. (2015). Hepatocellular carcinoma related to *Schistosoma mansoni* infection: case series and literature review. *J. Clin. Transl. Hepatol*., 3(4), 260. doi: 10.14218/JCTH.2015.00027

179. Palumbo, E. (2007). Association between schistosomiasis and cancer: a review. *Infect. Dis. Clin. Pract*., 15(3), 145-148. doi: 10.1097/01.idc.0000269904.90155.ce

180. Jans, C., and Boleij, A. (2018). The road to infection: host-microbe interactions defining the pathogenicity of *Streptococcus bovis/Streptococcus equinus* complex members. *Front. Microbiol*., 9, 603. doi: 10.3389/fmicb.2018.00603

181. Kumar, R., Herold, J. L., Schady, D., Davis, J., Kopetz, S., Martinez-Moczygemba, M., et al. (2017). *Streptococcus gallolyticus* subsp. *gallolyticus* promotes colorectal tumor development. *PLoS Pathog.*, 13(7), e1006440. doi: 10.1371/journal.ppat.1006440

182. Beck, M., Frodl, R., and Funke, G. (2008). Comprehensive study of strains previously designated *Streptococcus bovis* consecutively isolated from human blood cultures and emended description of *Streptococcus gallolyticus* and *Streptococcus infantarius* subsp. *coli*. *J. Clin. Microbiol*., 46(9), 2966-2972. doi: 10.1128/JCM.00078-08

183. Corredoira, J. C., Alonso, M. P., Garcia, J. F., Casariego, E., Coira, A., Rodriguez, A., et al. (2005). Clinical characteristics and significance of *Streptococcus salivarius* bacteremia and *Streptococcus bovis* bacteremia: a prospective 16-year study. *Eur. J. Clin. Microbiol. Infect. Dis*., 24(4), 250-255. doi: 10.1007/s10096-005-1314-x

184. Gonzlez-Quintela, A., Martınez-Rey, C., Castroagudın, J. F., Rajo-Iglesias, M. C., and Domınguez-Santalla, M. J. (2001). Prevalence of liver disease in patients with *Streptococcus bovis* bacteraemia. *J. Infect*., 42(2), 116-119. doi: 10.1053/jinf.2001.0799

185. Jean, S. S., Teng, L. J., Hsueh, P. R., Ho, S. W., and Luh, K. T. (2004). Bacteremic *Streptococcus bovis* infections at a university hospital, 1992-2001. *J. Formos. Med. Assoc*., 103(2), 118-123.

186. Murray, H. W., and Roberts, R. B. (1978). *Streptococcus bovis* bacteremia and underlying gastrointestinal disease. *Arch. Int. Med*., 138(7), 1097-1099. doi: 10.1001/archinte.1978.03630320037013

187. Ruoff, K. L., Miller, S. I., Garner, C. V., Ferraro, M. J., and Calderwood, S. B. (1989). Bacteremia with *Streptococcus bovis* and *Streptococcus salivarius*: clinical correlates of more accurate identification of isolates. *J. Clin. Microbiol.*, 27(2), 305-308. doi: 10.1128/jcm.27.2.305-308.1989

188. Zarkin, B. A., Lillemoe, K. D., Cameron, J. L., Effron, P. N., Magnuson, T. H., and Pitt, H. A. (1990). The triad of *Streptococcus bovis* bacteremia, colonic pathology, and liver disease. *Ann. Surg*., 211(6), 786. doi: 10.1097/00000658-199006000-00019

189. Boleij, A., Roelofs, R., Schaeps, R. M., Schülin, T., Glaser, P., Swinkels, D. W., et al. (2010). Increased exposure to bacterial antigen RpL7/L12 in early stage colorectal cancer patients. *Cancer*, 116(17), 4014-4022. doi: 10.1002/cncr.25212

190. Corredoira, J., Alonso, M. P., Coira, A., and Varela, J. (2008). Association between *Streptococcus infantarius* (formerly *S. bovis* II/1) bacteremia and noncolonic cancer. *J. Clin. Microbiol.*, 46(4), 1570-1570. doi: 10.1128/JCM.00129-08

191. Fernández-Ruiz, M., Villar-Silva, J., Llenas-García, J., Caurcel-Díaz, L., Vila-Santos, J., Sanz-Sanz, F., et al. (2010). *Streptococcus bovis* bacteraemia revisited: clinical and microbiological correlates in a contemporary series of 59 patients. *J. Infect*., 61(4), 307-313. doi: 10.1016/j.jinf.2010.07.007

192. Vaska, V. L., and Faoagali, J. L. (2009). *Streptococcus bovis* bacteraemia: identification within organism complex and association with endocarditis and colonic malignancy. *Pathology*, 41(2), 183-186. doi: 10.1080/00313020802436816

193. Boleij, A., Roelofs, R., Danne, C., Bellais, S., Dramsi, S., Kato, I., et al. (2012). Selective antibody response to *Streptococcus gallolyticus* pilus proteins in colorectal cancer patients. *Cancer Prev. Res*., 5(2), 260-265. doi: 10.1158/1940-6207.CAPR-11-0321

194. Garza-González, E., Ríos, M., Bosques-Padilla, F. J., Francois, F., Cho, I., González, G. M., et al. (2012). Immune response against *Streptococcus gallolyticus* in patients with adenomatous polyps in colon. *Int. J. Cancer*, 131(10), 2294-2299. doi: 10.1002/ijc.27511

195. Zammit, S. C., Azzopardi, N., and Ellul, P. (2014). *Streptococcus gallolyticus* bacteraemia in hepatobiliary–pancreatic and colonic pathologies. *QJM: Int. J. Med*., 107(5), 355-361. doi: 10.1093/qjmed/hct261

196. Butt, J., Romero-Hernández, B., Pérez-Gómez, B., Willhauck-Fleckenstein, M., Holzinger, D., Martin, V., et al. (2016). Association of *Streptococcus gallolyticus* subspecies *gallolyticus* with colorectal cancer: serological evidence. *Int. J. Cancer,* 138(7), 1670-1679. doi: 10.1002/ijc.29914

197. Butt, J., Werner, S., Willhauck-Fleckenstein, M., Michel, A., Waterboer, T., Zörnig, I., et al. (2017). Serology of *Streptococcus gallolyticus* subspecies *gallolyticus* and its association with colorectal cancer and precursors. *Int. J. Cancer*, 141(5), 897-904. doi: 10.1002/ijc.30765

198. Paritsky, M., Pastukh, N., Brodsky, D., Isakovich, N., and Peretz, A. (2015). Association of *Streptococcus bovis* presence in colonic content with advanced colonic lesion. *World J. Gastroenterol*., 21(18), 5663. doi: 10.3748/wjg.v21.i18.5663

199. Butt, J., Blot, W. J., Teras, L. R., Visvanathan, K., Le Marchand, L., Haiman, C. A., et al. (2018). Antibody responses to *Streptococcus Gallolyticus* subspecies *Gallolyticus* proteins in a large prospective colorectal cancer cohort consortium. *Cancer Epidemiol. Prev. Biomark*., 27(10), 1186-1194. doi: 10.1158/1055-9965.EPI-18-0249

200. Butt, J., Jenab, M., Willhauck-Fleckenstein, M., Michel, A., Pawlita, M., Kyrø, C., et al. (2018). Prospective evaluation of antibody response to *Streptococcus gallolyticus* and risk of colorectal cancer. *Int. J. Cancer*, 143(2), 245-252. doi: 10.1002/ijc.31283

201. Guven, D. C., Dizdar, O., Alp, A., Akdogan Kittana, F. N., Karakoc, D., Hamaloglu, E., et al. (2018). Analysis of *Fusobacterium nucleatum*, *Streptococcus gallolyticus* and *Porphyromonas gingivalis* in saliva in colorectal cancer patients and healthy controls. *J. Clin. Oncol.,* 36(15). doi: 10.1200/JCO.2018.36.15_suppl.e15617

202. Kale, P., Khillan, V., and Sarin, S. K. (2018). Novel association of *Streptococcus gallolyticus* subspecies *pasteurianus* and hepatocelluar carcinoma: opening new frontiers. *Scand. J. Gastroenterol*., 53(10-11), 1354-1357. doi: 10.1080/00365521.2018.1511826

203. Sheikh, A. F., Zadeh, A. R. M., Saki, M., Khani, P., Hashemi, S. J., Zadeh, S. S., et al. (2020). Detection of *Streptococcus gallolyticus* in colorectal cancer and inflammatory bowel disease patients compared to control group in southwest of Iran. *Mol. Biol. Rep*., 47(11), 8361-8365. doi: 10.1007/s11033-020-05807-7

204. Boleij, A., van Gelder, M. M., Swinkels, D. W., and Tjalsma, H. (2011). Clinical importance of *Streptococcus gallolyticus* infection among colorectal cancer patients: systematic review and meta-analysis. *Clin. Infect. Dis*., 53(9), 870-878. doi: 10.1093/cid/cir609

205. Krishnan, S., and Eslick, G. D. (2014). *Streptococcus bovis* infection and colorectal neoplasia: a meta-analysis. *Colorectal Dis*., 16(9), 672-680. doi: 10.1111/codi.12662

206. Pasquereau-Kotula, E., Martins, M., Aymeric, L., and Dramsi, S. (2018). Significance of *Streptococcus gallolyticus* subsp. *gallolyticus* association with colorectal cancer. *Front. Microbiol*., 9, 614. doi: 10.3389/fmicb.2018.00614

207. Buonfrate, D., Bisanzio, D., Giorli, G., Odermatt, P., Fürst, T., Greenaway, C., et al. (2020). The global prevalence of *Strongyloides stercoralis* infection. *Pathogens*, 9(6), 468. doi: 10.3390/pathogens9060468

208. Buonfrate, D., Requena-Mendez, A., Angheben, A., Muñoz, J., Gobbi, F., Van Den Ende, J., et al. (2013). Severe strongyloidiasis: a systematic review of case reports. *BMC Infect. Dis*., 13(1), 1-10. doi: 10.1186/1471-2334-13-78

209. Catalano, C., Aron, J., Bansal, R., and Leytin, A. (2017). Colorectal cancer associated with *Strongyloides stercoralis* colitis. *ACG Case Rep. J*., 4.

210. Tanaka, T., Hirata, T., Parrott, G., Higashiarakawa, M., Kinjo, T., Kinjo, T., et al. (2016). Relationship among *Strongyloides stercoralis* infection, human T-cell lymphotropic virus type 1 infection, and cancer: a 24-year cohort inpatient study in Okinawa, Japan. *Am. J. Trop. Med. Hyg*., 94(2), 365. doi: 10.4269/ajtmh.15-0556

211. Hirata, T., Kishimoto, K., Kinjo, N., Hokama, A., Kinjo, F., and Fujita, J. (2007). Association between *Strongyloides stercoralis* infection and biliary tract cancer. *Parasitol. Res.*, 101(5), 1345-1348. doi: 10.1007/s00436-007-0648-y

212. Seo, A. N., Goo, Y. K., Chung, D. I., Hong, Y., Kwon, O., and Bae, H. I. (2015). Comorbid gastric adenocarcinoma and gastric and duodenal *Strongyloides stercoralis* infection: a case report. *Korean J. Parasitol*., 53(1), 95. doi: 10.3347/kjp.2015.53.1.95

213. Ishikawa, S., Maeda, T., Hattori, K., Watanabe, T., Kuramoto, T., Ueno, S., et al. (2017). A case of adenocarcinoma developed in the small intestine with chronic strongyloidiasis. *Clin. J. Gastroenterol*., 10(6), 519-523. doi: 10.1007/s12328-017-0783-7

214. Sava, M., Huynh, T., Frugoli, A., Kong, L., Salehpour, M., and Barrows, B. (2020). Colorectal cancer related to chronic *Strongyloides stercoralis* infection. *Case Rep. Gastrointest. Med*., 2020. doi: 10.1155/2020/8886460

215. Tomaino, C., Catalano, C., Tiba, M., and Aron, J. (2015). Su2012 A first case report of colorectal cancer associated with chronic *Strongyloides stercoralis* colitis and the complex management decisions that follow. *Gastroenterol*., 148(4), S-575.

216. Machado, E. R., Teixeira, E. M., Goncalves-Pires, M. D. R. F., Loureiro, Z. M., Araujo, R. A., and Costa-Cruz, J. M. (2008). Parasitological and immunological diagnosis of *Strongyloides stercoralis* in patients with gastrointestinal cancer. *Scand. J. Infect. Dis*., 40(2), 154-158. doi: 10.1080/00365540701558730

217. Tjalsma, H., Boleij, A., Marchesi, J. R., and Dutilh, B. E. (2012). A bacterial driver–passenger model for colorectal cancer: beyond the usual suspects. *Nat. Rev. Microbiol*., 10(8), 575-582. doi: 10.1038/nrmicro2819

218. Kostic, A. D., Chun, E., Robertson, L., Glickman, J. N., Gallini, C. A., Michaud, M., et al. (2013). *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe*, 14(2), 207-215. doi: 10.1016/j.chom.2013.07.007

219. Sun, Z., Xiong, C., Teh, S. W., Lim, J. C. W., Kumar, S., and Thilakavathy, K. (2019). Mechanisms of oral bacterial virulence factors in pancreatic cancer. *Front. Cell. Infect. Microbiol*., 9, 412. doi: 10.3389/fcimb.2019.00412

220. Wacholder, S., Silverman, D. T., McLaughlin, J. K., and Mandel, J. S. (1992). Selection of controls in case-control studies: II. Types of controls. *Am. J. Epidemiol*., 135(9), 1029-1041. doi: 10.1093/oxfordjournals.aje.a116397

221. Paschke, S., Jafarov, S., Staib, L., Kreuser, E. D., Maulbecker-Armstrong, C., Roitman, M., et al. (2018). Are colon and rectal cancer two different tumor entities? A proposal to abandon the term colorectal cancer. *Int. J. Mol. Sci*., 19(9), 2577. doi: 10.3390/ijms19092577

222. Xi, Y., Yuefen, P., Wei, W., Quan, Q., Jing, Z., Jiamin, X., and Shuwen, H. (2019). Analysis of prognosis, genome, microbiome, and microbial metabolome in different sites of colorectal cancer. *J. Transl. Med*., 17(1), 1-22. doi: 10.1186/s12967-019-2102-1

223. Liu, X., Wang, X., Lin, S., Yuan, J., and Yu, I. T. (2014). Dietary patterns and oesophageal squamous cell carcinoma: a systematic review and meta-analysis. *Br. J. Cancer*, 110(11), 2785-2795. doi: 10.1038/bjc.2014.172

224. Hendriksen, R. S., Vieira, A. R., Karlsmose, S., Lo Fo Wong, D. M., Jensen, A. B., Wegener, H. C., et al. (2011). Global monitoring of *Salmonella* serovar distribution from the World Health Organization Global Foodborne Infections Network Country Data Bank: results of quality assured laboratories from 2001 to 2007. *Foodborne Pathog. Dis*., 8(8), 887-900. doi: 10.1089/fpd.2010.0787

225. Pijnacker, R., Friesema, I. H. M., Mughini Gras, L., Lagerweij, G. R., Van Pelt, W., and Franz, E. (2019). Disease burden of food-related pathogens in the Netherlands, 2018. doi: 10.21945/RIVM-2019-0086

226. Aya, V., Flórez, A., Perez, L., and Ramírez, J. D. (2021). Association between physical activity and changes in intestinal microbiota composition: A systematic review. *PloS One*, 16(2), e0247039. doi: 10.1371/journal.pone.0247039

227. Litwinowicz, K., Choroszy, M., and Waszczuk, E. (2020). Changes in the composition of the human intestinal microbiome in alcohol use disorder: a systematic review. *Am. J. Drug Alcohol Abuse*, 46(1), 4-12.

228. Matejcic, M., Gunter, M. J., and Ferrari, P. (2017). Alcohol metabolism and oesophageal cancer: a systematic review of the evidence. *Carcinogenesis*, 38(9), 859-872. doi: 10.1093/carcin/bgx067

229. Savin, Z., Kivity, S., Yonath, H., and Yehuda, S. (2018). Smoking and the intestinal microbiome. *Arch. Microbiol*., 200(5), 677-684. doi: 10.1007/s00203-018-1506-2

Supplementary tables and figure

Supplementary Table 1. Search terms used in literature search

Category 1† [Title]	Category 2† [Title]	Category 3† [Title/Abstract]	Category 4† [Title/Abstract]	Category 5† [Title]
digestive	cancer*	bacteria*	hazard ratio*	COVID
orodigestive	neoplasm*	bacterium	HR*	COVID-19
gastrointestinal	neoplasia*	microorganism*	relative risk*	SARS-COV2
enteric	adenocarcinoma*	microbe*	RR*	virus*
abdomen	adenoma*	microbial	standardized incidence ratio*	viral
abdominal	tumor*	parasite	SIR*	papillomavirus
esophagus	carcinoma*	parasitic	standardized mortality ratio*	HPV
oesophagus	malignancy		SMR*	Helicobacter
esophageal	malignancies		odds	pylori
oesophageal			incidence*	treatment*
stomach			Cox	chemotherap*
gastric			case-control	therapy
duodenum			cohort	therapies
duodenal				therapeutic
intestine				cholecystectomy
intestinal				hemicolectomy
gut				splenectomy
colon				colectomy
colorectal				gastrectomy
bowel				surgery
rectosigmoid				surgeries
rectum				postoperative
rectal				operative
anus				surgical
anal				rat
liver				rats
intrahepatic				mouse
hepatic				mice
bile duct				
cholangiocarcinoma				
gallbladder				
gall-bladder				
biliary				
pancreas				
pancreatic				

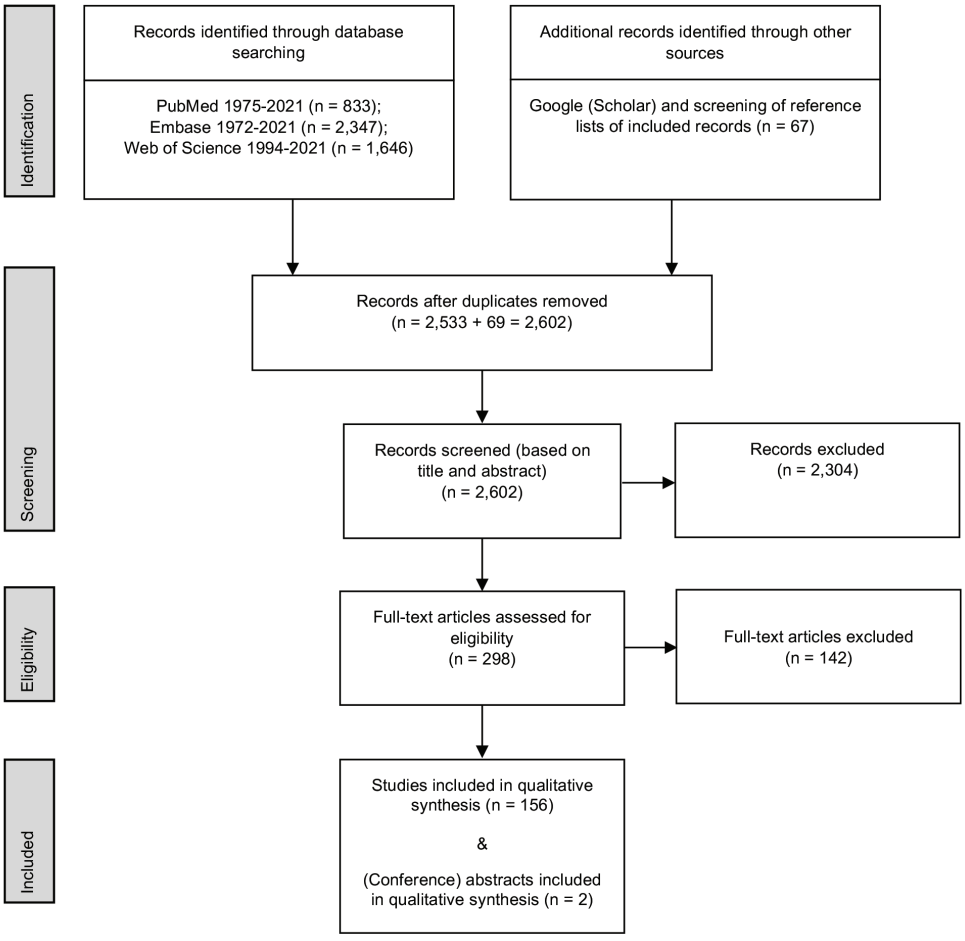
† Categories 1-4 combined with 'AND', Category 5 'AND NOT'.

Table S2. Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria

- One of the following study designs:
  - Case-control;
  - Cohort;
  - Cross-sectional.
- At least one of the following data types:
  - Surveillance or survey data (e.g. bacterial infection records, self-reported disease);
  - Serological assays;
  - Data about presence and/or abundance of microbial genetic material in human specimens (e.g. tumor tissue, blood, feces, saliva) in relation to GI cancer.
- Diagnosed malignancy in one of the following organs:
  - Esophagus;
  - Stomach;
  - Small intestine;
  - Duodenum;
  - Colon;
  - Rectum;
  - Anus;
  - Liver;
  - Intrahepatic bile ducts;
  - Biliary tract;
  - Gallbladder;
  - Pancreas.
	- Articles exclusively focusing on viruses in relation to cancer.
	- Articles exclusively focusing on *Helicobacter pylori.*
	- Articles focusing on the broad composition of the microbiome in relation to cancer, i.e. those only addressing bacterial phyla or genera rather than bacterial species*****
	- Articles addressing the association of bacteria/parasites and cancer during or after cancer treatment.
	- Articles with an experimental design (e.g. in vitro or in vivo studies).
	- Review articles.

***** Based on search of title and abstract.



Supplementary Figure 1. Prisma flow-diagram of the article selection process.

Supplementary Table 3. Characteristics and main outcomes of epidemiological studies assessing the association between ***Bacteroides fragilis*** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Toprak, 2006 [20]	Turkey	*unk*	Colorectal	Case-control	*Bf*: DNA detection in fecal samples using PCR. Cancer: *unk.*	73 CRC patients, 59 healthy controls.	*Bf* present in fecal samples of 56 (76.7%) CRC patients *vs.* 40 (67.8%) controls (p>0.05). *Bft* gene present in 38% of the CRC patients *vs.* 12% of the controls (RR 4.16; 95%CI 1.39-12.43, p=0.009).
Abdulamir, 2009 [11]	Malaysia	2006-2007	Colorectal	Case-control	*Bf*: detection of IgG antibodies in blood samples using ELISA, and detection of NF-kB and IL-8 mRNA expression in tissue (tumor, [adjacent] normal) samples using in situ hybridization assay. Cancer: diagnosis based on colonoscopy.	50 CRC patients, 14 colorectal adenoma patients, 30 apparently healthy volunteers and 30 controls without colon tumors.	Seroprevalence of *Bf* not higher in CRC patients (0.166 ± 0.013) and adenoma patients (0.178 ± 0.032) as compared with healthy volunteers (0.176 ± 0.014) and controls (0.180 ± 0.020). NF-kB and IL-8 expression significantly higher in tumor tissue of CRC patients and adenoma patients compared to adjacent normal tissue and tissue from healthy controls (p<0.05).
Rahimkhani, 2010 [19]	Iran	*unk*	Colorectal	Case-control	*Bf*: detection in fecal samples using culturing methods. Cancer: *unk.*	30 CRC patients, 30 healthy controls.	*Bf* most abundant bacterial species in fecal samples of 8 CRC patients and 2 controls.
Keenan, 2014 [16]	*unk*	*unk*	Colorectal	Case-control	*Bf*: DNA and toxin gene detection in fecal samples using PCR. Cancer: *unk.*	61 CRC patients, 125 healthy controls.	*Bf* detected in 51 (83.6%) of fecal samples from CRC patients *vs.* 114 (91.2%) of healthy controls. ETBF-positivity higher in CRC patients (19.6%) *vs.* controls (7.4%) (p=0.087).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Boleij, 2015 [12]	USA	2010-2013	Colorectal	Case-control	*Bf*: detection of *bft* gene in mucosal samples using PCR. Cancer: *unk*.	49 CRC patients, 49 healthy controls.	*Bft* gene present in mucosa of 91.7% proximal and 85.7% of distal CRC tissue or adjacent normal tissue *vs.* 55.5% (proximal) and 53.1% (distal) in controls (significance levels: proximal p=0.04; distal p=0.03). Positive association between *Bft*-positivity and CRC stage (72.7% in early CRC, 100% in late CRC, p=0.09).
Fukugaiti, 2015 [14]	Brazil	*unk*	Colorectal	Case-control	*Bf*: detection in fecal samples using PCR. Cancer: diagnosis based on colonoscopy.	7 CRC patients, 10 healthy controls.	*Bf* detected in feces of 6/7 (85.7%) CRC patients and 6/10 (60.0%) controls. No significant difference in fecal abundance of *Bf* in CRC patients *vs.* controls (4.8 ± 2.6 *vs.* 4.4 ± 2.5 $\log_{10}$ no. of copies, p=0.78).
Viljoen, 2015 [21]	South-Africa	*unk*	Colorectal	Cross-sectional	*Bf*: detection of *Bf* DNA in tissue samples using qPCR. Cancer: *unk*.	Paired tumor and normal adjacent tissue from 55 CRC patients. Additionally, 18 adenocarcinoma samples.	ETBF present in 14/54 (25.9%) of CRC patients and 15/53 (28.3%) of adjacent normal tissue. Both tumor and normal tissue infected in 71% of ETBF-positive patients. ETBF significantly more present in colon than rectum, no differences between proximal and distal colon. Positive association between ETBF-positivity and CRC stage.
Xie, 2016a* [22]	China	2011-2013	Colorectal	Case-control	*Bf*: DNA detection in tissue samples using PCR. Cancer: colon/ rectum cancer diagnosed by colonoscopy and pathology.	Paired tumor and normal adjacent tissue from 36 CRC patients. Normal tissue from 18 healthy controls.	Significant decreased number of *Bf* in CRC tissue (5.5x10^3 copies/g) *vs.* normal adjacent (4.3x10^4) and healthy control tissue (6.0x10^5). ETBF was significantly increased in CRC tissue compared to healthy normal tissue.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Purcell, 2017 [18]	New-Zealand	2003-2005	Colorectal	Prospective cohort	*Bf*: *bft* gene detection in tissue samples using PCR. Cancer: *unk*.	150 patients referred for colonoscopy, 20 of which developed CRC.	ETBF present in tissue samples of 74/150 (49.3%) patients. No significant association between ETBF-positivity and CRC was found. Significant association between ETBF-positivity and low-grade dysplasia (0.0007), tubular adenomas (p 0.027) and serrated polyps (0.007). ETBF presence more likely in lower (distal) parts of colon.
Hale, 2018 [15]	USA	*unk*	Colorectal	Cross-sectional	*Bf*: DNA detection in (tumor and adjacent normal) tissue using 16S rRNA gene sequencing and PCR to detect *bft* gene. Cancer: *unk*.	83 CRC patients (25 with deficient mismatch repair [dMMR] CRC, 58 with proficient mismatch repair [pMMR] CRC).	*Bf* significantly more abundant in tumor tissue of dMMR CRC patients as compared with normal tissue (p=0.02), not in pMMR CRC patients. No significant difference in prevalence of *bft* gene between dMMR CRC patients and pMMR CRC patients.
Kwong, 2018 [17]	China	2006-2015	Colorectal	Retrospective cohort	*Bf*: culture-confirmed bacteremia. Cancer: *unk*.	1,338 *Bf* bacteremia patients, 6,690 matched controls without history of bacteremia.	42/1,338 (3.1%) of the bacteremia patients developed CRC compared to 73/6,690 (1.1%) of the controls, aHR: 3.85 (95%CI 2.62-5.64; p<0.0001).
Bundgaard, 2019 [13]	Denmark	2002-2010	Colorectal	Prospective cohort	*Bf*: DNA detection in (tumor and adjacent normal) tissue samples using qPCR. Cancer: *unk*.	99 CRC patients, 96 adenoma patients, 104 patients with diverticular disease.	*Bf* detected in 36.4% of the CRC patients. Detection rate lowest for adenoma samples as compared to tumor tissue, adjacent normal and diverticular tissue. Detection rate significantly higher in adjacent normal tissue compared to tumor tissue (p<0.05).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Haghi, 2019 [7]	Iran	2016-2018	Colorectal	Case-control	*Bf*: detection of *neu* and *bft* genes and enterotoxin in fecal samples using PCR. Cancer: *unk*.	60 CRC patients, 60 healthy controls.	*neu* gene present in 58.3% CRC patients *vs.* 26.6% controls. *Bft* gene present in 31.6% CRC patients *vs.* 8.3% controls. Prevalence of *Bf* significantly higher in CRC patients *vs.* controls ($p<0.05$). Positive association between presence of *bft* gene and CRC stage ($p<0.05$).
Justesen, 2020* [9]	Denmark	2007-2018	Colorectal	Retrospective cohort	*Bf*: culture confirmed bacteremia (in blood samples). Cancer: *unk*.	583 patients with *Bf* bacteremia of which 11 developed CRC. Reference population: ~2 million.	11/583 (1.9%) individuals with a history of bacteremia caused by *Bf* were diagnosed with CRC, all within 1 year after bacteremia.
Zamani, 2020 [8]	Iran	2015-2017	Colorectal	Case-control	*Bf*: detection of *Bf* and *bft* gene in mucosal biopsies using rt-PCR. Cancer: histologically confirmed colon/ rectum cancer.	Mucosal biopsies from 26 CRC patients and 42 patients with precancerous conditions and from 52 healthy controls.	*Bf* detected in 15/26 (57.7%) CRC patients *vs.* 42/52 (80.8%) healthy controls. *Bft* gene detected in 11/26 (42.3%) CRC patients *vs.* 2/52 (3.8%) healthy controls (OR 18.3; 95%CI 3.7-92.0).
Ma, 2021 [10]	Europe	*unk*	Colorectal	Case-control	*Bf*: detection in fecal samples using sequencing. Cancer: *unk*.	285 CRC patients, 512 IBD patients, 290 healthy controls.	Abundance of genus-level *Bacteroides* and *Bf* significantly elevated in IBD patients compared to healthy controls and CRC patients.

*Only abstract available. *unk*: unknown. *Bf*: *Bacteroides fragilis*. (q)PCR: (quantitative) polymerase chain reaction. CRC: colorectal cancer. RR: relative risk. ETBF: enterotoxigenic *B. fragilis*. OR: odds ratio. aHR: adjusted hazard ratio. IBD: inflammatory bowel disease.

Supplementary Table 4. Characteristics and main outcomes of epidemiological studies assessing the association between ***Campylobacter* spp.** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Brauner, 2010 [28]	Sweden	1989-2007	Esophagus, Stomach, Duodenum, Colon, Rectum + anus, Liver + bile duct	Retrospective cohort	*C. jejuni*: detection in fecal samples using culturing methods. Cancer: diagnosis according to ICD-9 150-156 codes.	16,276 individuals with a registered *C. jejuni* infection, of which 3 developed esophageal cancer, 5 GC, none duodenal cancer, 21 CC, 8 rectal or anal cancer, 6 liver or bile duct cancer. Reference population: ~9 million.	Esophageal cancer: SIR 0.99 (95%CI 0.20-2.89); GC SIR: 0.75 (95%CI 0.24-1.76); Duodenal cancer: SIR: 0.00 (95%CI 0.00-2.15); CC SIR: 0.90 (95%CI 0.56-1.38); Rectal or anal cancer: SIR 0.56 (95%CI 0.24-1.10); Liver or bile duct cancer: SIR 1.00 (95%CI 0.37-2.17). Follow-up up to 19 years (median 7.6 years).
Blackett, 2013 [27]	UK	*unk*	Esophagus	Case-control	*Campylobacter* spp.: detection in esophageal tissue using culturing methods. Cancer: diagnosis based on histology.	30 esophageal cancer patients, 45 Barrett's esophagus (BE) patients, 37 gastro-esophageal reflux disease (GERD) patients and 39 healthy controls.	*Campylobacter* spp. detected in 3/34 (8.8%) of the cancer patients, 19/45 (42.2%) BE patients, 19/37 (51.4%) GERD patients, and 5/39 (12.8%) controls. Most of the isolated *Campylobacter* were *C. concisus*. Esophageal colonization by *Campylobacter* in GERD and BE patients is likely caused by refluxate (which is reduced/ minimal in cancer patients and healthy controls).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Wu, 2013 [33]	China	*unk*	Colorectal	Case-control	*Campylobacter* spp.: detection in fecal samples using pyro-sequencing of the 16S rRNA gene V3 region. Cancer: *unk*.	19 CRC patients, 20 healthy controls.	*Campylobacter* significantly more abundant in CRC patients *vs.* controls.
Allali, 2015 [26]	USA, Spain	*unk*	Colorectal	Cross-sectional	*Campylobacter* spp.: detection in (tumor) tissue using 16S rRNA gene sequencing. Cancer: *unk*.	Paired tumor and adjacent normal tissue from 90 CRC patients.	*Campylobacter* spp. significantly more abundant in tumor tissue *vs.* adjacent normal tissue of the Spanish cohort and in the adjacent normal tissue of the USA cohort *vs.* the adjacent normal tissue of the Spanish cohort.
Alexander, 2016* [25]	UK	*unk*	Colorectal	Cross-sectional	*Campylobacter* spp.: detection in (tumor) tissue using 16S rRNA gene sequencing. Cancer: *unk*.	Paired tumor and adjacent normal tissue from 46 CRC patients (19 proximal, 11 distal, 16 rectal).	*Campylobacter* spp. detected in tissue from 11/46 (23.9%) of CRC patients. Significant higher abundance of *Campylobacter* spp. in tumor tissue compared to adjacent normal tissue (p=0.007).
Mughini-Gras, 2017* [30]	The Netherlands	2000-2015	Colon	Retrospective cohort	*Campylobacter* spp: detection in feces, blood, or other samples using culturing methods. Cancer: diagnosis according to ICD-10 C180-C187 codes.	83 CC patients. Reference population: ~17 million.	Overall risk of CC not significantly higher after *Campylobacter* infection (SIR 1.06; 95%CI 0.84-1.31). Significant higher risk of CC after infection between 40-49 years of age (SIR 2.27; 95%CI 1.17-3.96).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
De Savornin-Lohman, 2020 [29]	The Netherlands	2000-2016	Biliary tract	Retrospective cohort	*Campylobacter* spp: detection in feces, blood, or other samples using culturing methods. Cancer: diagnosis according to ICD-O C239, C240, C242-C244, C248, C249 codes.	7 biliary tract cancer patients. Reference population: ~17 million.	No increased risk of biliary tract cancer after *Campylobacter* infection (SIR 0.97; 95%CI 0.39–2.00).
Wang, 2020 [31]	China	*unk*	Colorectal	Case-control	*Campylobacter* spp.: detection in (tumor) and mucosa using 16S rRNA gene sequencing. Cancer: *unk*.	Tumor, adjacent normal and off-tumor site tissue from 75 CRC patients and mucosa from 26 healthy controls.	Relative abundance of *Campylobacter* spp. significantly higher at tumor, adjacent normal and off-tumor tissue of CRC patients *vs*. healthy controls.
Wei, 2020 [32]	China	2017-2018	Pancreas	Case-control	*Campylobacter* spp: detection in saliva samples using 16S rRNA gene sequencing. Cancer: histopathologically-confirmed pancreatic adenocarcinoma.	41 PDAC patients, 69 healthy controls.	Reduced abundance of *Campylobacter* spp. in saliva of PDAC patients *vs.* controls.

*Only abstract available. ICD-9 or ICD-10: International Classification of Diseases (9th or 10th revision). GC: gastric cancer. CC: colon cancer. ICD(-O): International Classification of Diseases (for Oncology). SIR: standardized incidence ratio.. *Unk*: unknown.

Supplementary Table 5. Characteristics and main outcomes of epidemiological studies assessing the association between ***Clostridium* spp**. and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Lu, 2004* [42]	China	*unk*	Gallbladder	Cross-sectional	*Clostridium*: detection in GBC tissue using 16S rRNA PCR. Cancer: *unk.*	46 GBC patients.	36/46 (78.3%) of tissue samples contained bacterial DNA, including *C. perfringens*.
Rahimkhani, 2010 [19]	Iran	*unk*	Colorectal	Case-control	*Clostridium*: detection in fecal samples using culturing methods. Cancer: *unk.*	30 CRC patients, 30 healthy controls.	*C. perfringens* most abundant bacterial species in fecal samples of 11 CRC patients and 16 controls.
Ahn, 2013 [40]	US	1985-1989	Colorectal	Case-control	*Clostridium* spp: detection in fecal samples using 16S rRNA sequencing. Cancer: diagnosis based on histology.	47 CRC patients, 94 healthy controls.	Decreased abundance of *Clostridium* spp. in CRC patients *vs.* controls.
Ohigashi, 2013* [43]	Japan	*unk*	Colorectal	Case-control	*Clostridium*: detection in fecal samples using RT-qPCR. Cancer: *unk.*	93 CRC patients, 23 adenoma patients, 27 healthy controls.	*C. coccoides* group and *C. leptum* subgroup were significantly less abundant in CRC patients *vs.* controls.
Fukugaiti, 2015 [14]	Brazil	*unk*	Colorectal	Case-control	*Clostridium*: detection in fecal samples using PCR. Cancer: diagnosis based on colonoscopy.	7 CRC patients, 10 healthy controls.	*C. difficile* and *C. perfringens* detected 7/7 and 4/7 CRC patients respectively and 8/10 and 8/10 controls respectively. Significant higher abundance of *C. difficile* in CRC patients *vs.* controls (2.5 ± 0.6 *vs.* 1.6 ± 0.8 $\log_{10}$ no. of copies, p=0.04).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Liang, 2016 [41]	China	2009-2015	Colorectal	Case-control	*C. hathewayi*: detection in fecal samples using qPCR. Cancer: *unk.*	203 CRC patients, 236 healthy controls.	Significant higher abundance of *C. hathewayi* in feces of CRC patients *vs.* controls (p<0.0001). A combination of four bacteria including *Fusobacterium nucleatum*, *Bacteroides clarus*, *C. hathewayi* and an undefined species could serve as potential biomarker for early detection of CRC.
Hsieh, 2018 [44]	Taiwan	*unk*	Stomach	Case-control	*Clostridium* spp.: detection of *Clostridium* in gastric biopsies using 16S RNA sequencing. Cancer: *unk.*	7 GC patients, 9 gastritis patients, 7 patients with intestinal metaplasia.	*Clostridium* spp. enriched in GC patients *vs.* gastritis patients. *C. colicanis* most frequently detected species.
Kwong, 2018 [17]	China	2006-2015	Colorectal	Retrospective cohort	*Clostridium*: culture-confirmed bacteremia. Cancer: *unk.*	13 patients with *C. septicum* bacteremia *vs.* 65 matched controls without bacteremia. 522 patients with *C. perfringens* bacteremia *vs.* 2,610 matched controls.	*C. septicum*: 4/13 (30.8%) of the bacteremia patients developed CRC compared to 1/65 (1.5%) of the controls, aHR: 17.1 (95%CI 1.82-160.0; p=0.013). *C. perfringens*: 11/522 (2.1%) of the bacteremia patients developed CRC compared to 35/2,610 (1.3%) of the controls, aHR 2.29 (95%CI 1.16-4.52, p=0.017).
Jahani-Sherafat, 2019 [45]	Iran	2016-2017	Colorectal	Cross-sectional	*C. difficile*: detection in colon biopsies by culturing methods and PCR. Cancer: diagnosis based on colonoscopy and pathologic reports.	Tumor tissue and adjacent normal tissue from 30 CRC patients.	*C. difficile* detected in 18/30 (60.0%) of tumor tissues and 6/30 (20.0%) healthy adjacent tissue.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Justesen, 2020* [9]	Denmark	2007-2018	Colorectal	Retrospective cohort	*Clostridium*.: culture confirmed bacteremia (in blood samples). Cancer: *unk*.	457 patients with *Clostridium* spp. bacteremia, of which 167 caused by *C. perfringens* and 53 by *C. septicum*. Reference population: ~2 million.	22/457 (4.8%) of individuals with a history of bacteremia caused by *Clostridium* spp. were diagnosed with CRC, most of them (n=20) within 1 year. 3/167 (1.8%) *C. perfringens* and 12/53 (22.6%) *C. septicum* bacteremia patients were diagnosed with CRC, all but one *C. septicum* within one year after bacteremia.
Liang, 2021 [46]	China	2009-2014	Colorectal	Case-control	*C. hathewayi*: detection in fecal samples using PCR. Cancer: diagnosis based on colonoscopic examination, histopathological review.	210 CRC patients, 115 AA patients, 86 NAA patients, 265 healthy controls.	*Clostridium* useful as fecal diagnostic marker for CRC as compared to *Fusobacterium nucleatum* and *Lachnoclostridium* spp.

*Only abstract available. *Unk:* unknown. PCR: polymerase chain reaction. GBC: gallbladder carcinoma. CRC: colorectal cancer. GC: gastric cancer. aHR; adjusted hazard ratio. AA: advanced adenocarcinoma. NAA: non-advanced adenocarcinoma.

Supplementary Table 6. Characteristics and main outcomes of epidemiological studies assessing the association between ***Cryptosporidium* spp.** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Sulżyc-Bielicka, 2007 [52]	Poland	*unk*	Colorectal	Cross-sectional	*Cryptosporidium*: detection in fecal samples using microscopy and microplate assay. Cancer: diagnosis based on colonoscopy.	55 CRC patients.	10/55 (18.2%) of the CRC patients tested positive on *Cryptosporidium.* No statistically significant difference in positivity was observed between proximal and distal colon cancer (proximal: 1/12, 8.3%; distal: 9/43, 20.9%).
Sulżyc-Bielicka, 2012* [56]	Poland	2009-2010	Colorectal	Cross-sectional	*Cryptosporidium*: detection in fecal samples using immune-enzymatic tests. Cancer: *unk.*	87 CRC patients.	*Cryptosporidium* detected in feces of 11/87 (12.6%) CRC patients. *Cryptosporidium* infection not significantly associated with age, tumor location or cancer stage.
Sanad, 2014 [55]	Saudi	2010-2011	Colorectal	Case-control	*Cryptosporidium*: detection based on microscopy. Cancer: *unk.*	20 CRC patients, 42 healthy controls.	14/20 (70.0%) of the CRC patients *Cryptosporidium*-positive *vs.* 8/42 (19.0%) of the controls.
Osman, 2017 [54]	Lebanon	2012-2013	Stomach, Colon	Case-control	*Cryptosporidium:* detection in biopsies using 18S rRNA real-time PCR, microscopy, immunofluorescence analysis. Cancer: recently diagnosed colon or stomach intraepithelial neoplasia/ adenocarcinoma.	72 CC patients, 21 GC patients, 125 controls without cancer but with persistent digestive symptoms (non-healthy controls), 44 healthy controls.	PCR-positive CC patients *vs.* healthy controls: OR 11.32 (95%CI 1.44-89.02). PCR-positive colon cancer patients *vs.* non-healthy controls: OR 4.05 (95%CI 1.39-11.79). No PCR-positive samples of GC patients.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Essid, 2018 [53]	Tunisia	2010-2015	Colorectal	Case-control	*Cryptosporidium*: detection in fecal samples using microscopy and PCR. Cancer: *unk.*	526 fecal samples, of which 15 from CRC patients, 260 from children, 197 from immunocompromised patients, 54 from myeloma patients.	5/15 (33.3%) CRC patients tested *Cryptosporidium*-positive *vs.* 37/511 (7.2%) of the non-CRC patients and controls.
Sulżyc-Bielicka, 2018 [51]	Poland	2009-2014	Colorectal	Case-control	*Cryptosporidium*: antigen detection in fecal samples using antigen-EIA technique. Cancer: diagnosis based on histopathology.	108 CRC patients, 125 healthy controls.	*Cryptosporidium* infections in cancer patients (prior to oncological treatment) *vs.* controls OR 3.43 (95%CI 1.17-10.07).

*Only abstract available. *Unk:* unknown. CRC: colorectal cancer. CC: colon cancer. GC: gastric cancer. PCR: polymerase chain reaction. Antigen-EIA: antigen enzyme immunoassay. OR: odds ratio.

Supplementary Table 7. Characteristics and main outcomes of meta-analyses assessing the association between microorganism(s) and cancer in the gastrointestinal tract.

First author, year	Microorganism	Type of study	Study period	Malignancy/ malignancies*	Articles/ sources included	Main outcomes
Boleij, 2011 [204]	*Streptococcus bovis*	SR + M-A	up to Dec 2010	Colorectal	52 case reports, 31 case series, of which 11 used for M-A.	Significant increased risk of CRC among patients with *Sgg* infection *vs. S. bovis* biotype II-infected persons, pooled OR 7.26 (95% CI 3.94–13.36). Higher CRC occurrence among *S.* bovis IE patients than other *S. bovis* infections (OR 3.72; 95%CI 2.03-6.81). The prevalence of CRC among *Sgg* infected people (33-71%) exceeds the prevalence of CRC in general population (10-25%).
Krishnan, 2014 [205]	*Streptococcus bovis*	M-A	up to March 2014	Colorectal	9 case-control studies, 39 case series.	Strongest association between S. bovis IE and CRC (OR 14.54; 95%CI 5.66–37.35). OR for *S. bovis* septicemia and CRC: 7.48 (95%CI: 3.10–18.06). OR for *S. bovis* fecal carriage and CRC 2.52 (95%CI: 1.14–5.58). All three with moderate levels of heterogeneity.
Nagaraja, 2014 [165]	*Salmonella* Typhi	SR + M-A	up to Nov 2013	Gallbladder	15 case-control studies, 2 cohort studies.	Chronic carriage of *S.* Typhi was significantly associated with an increased risk of GBC (pooled OR 4.28; 95%CI 1.84-9.96, p<0.01). Estimated risk was higher based on culturing methods (pooled OR 4.14; 95%CI 2.48-5.00, p<0.01) compared to methods based on detection of antibody titers (pooled OR 3.52; 95%CI 2.48-5.00, p<0.01). Estimated risk was higher when using controls without as reference (pooled OR 5.86; 95%CI 3.84-8.95, p<0.01) *vs.* controls with gallstones (pooled OR 2.71; 95%CI 1.92-3.83, p<0.01).
Koshiol, 2016 [166]	*Salmonella* Typhi	M-A	up to Feb 2016	Liver, Biliary tract, Gallbladder	18 case-control studies, 4 cohort studies. 14 studies used for M-A.	Odds of detecting higher Vi-antibody titers against *S.* Typhi higher in GBC patients compared to controls (OR 4.0; 95%CI 0.9-18.3). RR of detecting *S.* Typhi in bile or fecal samples was 5.0 (95%CI 2.7-9.3).

First author, year	Microorganism	Type of study	Study period	Malignancy/ malignancies*	Articles/ sources included	Main outcomes
Gethings-Behncke, 2020 [122]	*Fusobacterium nucleatum*	SR + M-A	up to Jan 2019	Colorectal	45 relevant articles identified, of which 18 were used for M-A.	Higher odds of detecting *Fn* in tissue from CRC patients or polyp patients *vs.* healthy controls (pooled OR 10.06; 95%CI 4.48-22.58, n=6 studies, and pooled OR 1.83; 95%CI 1.07-3.16, n=5 studies respectively). Also higher OR for detecting *Fn* in tumor tissue compared to adjacent normal tissue (OR 2.42; 95%CI 1.62-3.61, n=7 studies). Higher *Fn*-positivity rates in fecal samples from CRC patients *vs.* healthy controls (OR 9.01; 95%CI 3.39-23.95, n=7 studies).
Janati, 2020 [123]	*Fusobacterium nucleatum*	SR + M-A	up to Dec 2018	Colorectal	24 relevant articles identified, of which 12 were used for M-A.	A significant overall association was found between detection of F. nucleatum in colorectal specimens (fecal, mucosal and/or tissue samples) and CRC, based on data of 12 studies (OR 8.3; 95%CI 5.2-13.0; moderate heterogeneity I^2 26.3%, p=0.018). The association appeared stronger for Asian populations (OR 12.6; 95%CI 7.2-21.9) compared to American (OR 5.6; 95%CI 2.8-11.6) and European (OR 4.6; 95%CI 2.5-8.4) populations.
Kalantari, 2020 [48]	*Cryptosporidium* spp.	SR + M-A	up to July 2018	Colorectal	9 case-control studies, 10 cross-sectional studies.	4 of the included studies reported site-specific estimates for CRC, the other studies assessed malignancies outside the GI tract or provided a single estimate for all malignancies combined. The pooled OR for CRC was 3.7 (95%CI 2.10–6.50).

*Malignancies within the gastrointestinal tract, excluding possible malignancies outside the gastrointestinal tract. SR: systematic review. M-A: meta-analysis. CRC: colorectal cancer. GI: gastrointestinal. OR: odds ratio. *Sgg: Streptococcus bovis* biotype I. IE: infective endocarditis. GBC: gallbladder carcinoma. RR: relative risk. *Fn*: *Fusobacterium nucleatum.*

Supplementary Table 8. Characteristics and main outcomes of epidemiological studies assessing the association between ***Enterococcus faecalis*** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Balamurugan, 2008 [60]	India	*unk*	Esophagus, Stomach, Colorectal	Case-control	*E. faecalis*: detection in fecal samples using PCR. Cancer: diagnosis based on colonoscopy or biopsy.	20 CRC patients, 9 patients with esophageal or gastric cancer, 17 healthy controls.	Increased abundance of *E. faecalis* in feces of CRC patients (0.0500 ± 0.0426) *vs.* healthy controls (0.0106 ± 0.0086). No difference observed between esophageal/gastric cancer patients and controls.
Wang, 2012 [67]	China	*unk*	Colorectal	Case-control	*Enterococcus*: detection in fecal samples using pyrosequencing of V3 region of 16S rRNA gene. Cancer: diagnosis based on histopathology.	46 CRC patients, 56 healthy controls.	Higher relative abundance of *Enterococcus* (genus level) in CRC patients (2.4%) *vs.* controls (0.1%) ($p<0.001$).
Corredoira, 2015 [61]	Spain	1988-2014	Colorectal	Prospective cohort	*Enterococcus* spp: detection in blood samples using culturing methods. Cancer: diagnosis based on histology.	1,061 patients with enterococcal bacteremia, of which 36 (3.4%) with endocarditis.	6/28 (21%) of patients with enterococcal endocarditis developed CRC. During the acute endocarditis episode, 9/18 (50%) patients showed colorectal neoplasm, 4 NAA, 5 AA.
Viljoen, 2015 [21]	South-Africa	*unk*	Colorectal	Cross-sectional	*E. faecalis*: detection in tissue samples using qPCR. Cancer: *unk.*	Paired tumor and normal adjacent tissue from 55 CRC patients. Additionally, 18 adenocarcinoma samples.	*E. faecalis* detected in 11/40 (27.5%) tumor tissues and 7/38 (18.4%) normal adjacent tissues. Concurrent detection in both tumor and normal tissue in 5/10 (50.0%) patients.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Zhou, 2016 [68]	China	2012-2014	Colorectal	Case-control	*E. faecalis*: detection in (tumor and/or healthy) tissue samples using PCR. Cancer: diagnosis based on pathology.	97 CRC patients, 48 healthy controls.	Higher detection rates of *E. faecalis* in tumor tissue (93/97; 95.9%) *vs.* adjacent normal tissue (91/97; 93.8%) ($p<0.05$). *E. faecalis* more abundant in CRC patients *vs.* controls.
Pericàs, 2017 [64]	Spain	1979-2015	Colorectal	Retrospective cohort	*E. faecalis*: *E. faecalis* infective endocarditis according to Duke criteria. Cancer: diagnosis based on colonoscopy.	154 EFIE patients of which 5 developed CRC, 22 NAA, 5 AA.	Prevalence of CRC over 16-fold higher in studied EFIE cohort (8.2%) compared to Spanish population (0.5%).
Rezasoltani, 2018 [65]	Iran	2015-2017	Colorectal	Case-control	*E. faecalis*: detection in fecal samples using PCR. Cancer: diagnosis based on colonoscopy.	87 patients with polyps (n=21 hyperplastic polyp [HP], n=16 sessile serrated polyp [SSA], n=29 tubular adenoma [TA], n=21 [tubu]villous polyp [VP/TVP]), 31 healthy controls.	Higher abundance of *E. faecalis* in feces from TA and VP/TVP patients compared to HP, SSA patients and healthy controls. Higher abundance in patients with proximal polyps (*vs.* distal polyps) and with a higher grade of dysplasia.
de Almeida, 2019 [58]	Italy	2016	Colorectal	Case-control	*E. faecalis*: detection in fecal samples using culturing methods. Cancer: *unk*.	9 CRC patients, 9 healthy controls.	Frequency of *E. faecalis* lower in CRC patients (22%) *vs.* healthy controls (77%).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Geravand, 2019 [62]	Iran	*unk*	Colorectal	Case-control	*E. faecalis*: detection in fecal samples using PCR. Cancer: diagnosis based on colonoscopy and positive pathology.	25 CRC patients, 28 patients with polyps, 24 healthy controls.	Higher abundance of *E. faecalis* in feces from CRC patients ($11.2x10^9$) compared to polyp patients ($9.4x10^8$) (p=0.002) and healthy controls ($9x10^8$) (p=0.001).
Hussen, 2020 [63]	Iran	2014-2019	Colorectal	Case-control	*E. faecalis*: detection in fecal samples using PCR. Cancer: *unk*.	300 patients with polyps and CRC, 300 healthy controls.	2.2-fold higher abundance of *E. faecalis* in CRC patients ($1.5x10^9$ CFU/mL) *vs.* healthy controls ($8.5x10^8$ CFU/mL) (p=0.0013).
Shoji, 2021 [66]	Japan	2018-2019	Colorectal	Case-control	*E. faecalis*: detection in fecal samples using PCR and sequencing of 16s rRNA gene. Cancer: diagnosis based on colonoscopy.	36 CRC patients, 38 healthy controls.	Lower abundance of *E. faecalis* in obese CRC patients *vs.* nonobese CRC patients (p<0.01), opposite pattern in healthy controls. No significant difference in relative abundance of *E. faecalis* across CRC stages.

Unk: unknown. PCR: polymerase chain reaction. CRC: colorectal cancer. NAA: non-advanced adenoma. AA: advanced adenoma. qPCR: (quantitative) polymerase chain reaction. EFIE: *E. faecalis* infective endocarditis. HP: hyperplastic polyp. SSA: sessile serrated polyp. TA: tubular adenoma. (T)VP: (tubu)villous polyp.

Supplementary Table 9. Characteristics and main outcomes of epidemiological studies assessing the association between ***Escherichia coli*** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Buc, 2013 [76]	France	2007-2009	Colon	Case-control	*E. coli*: detection of specific genes in tissue samples using PCR. Cancer: diagnosis based on pathology.	38 CC patients, 31 non-cancer controls (with diverticulosis).	Number of *E. coli*-negative samples significantly lower in CC patients (1/38, 2.6%) *vs.* non-cancer controls (6/31, 19.4%) (p=0.04). Prevalence of phylogroup B2 *E. coli* significantly higher in tissue from CC patients (55.3%) *vs.* non-cancer controls (19.3%) (p<0.01). Cyclomodulin-encoding genes (*pks*, *cnf1*, *cdt*) overrepresented in CC patients *vs*. non-cancer controls.
Bonnet, 2014 [75]	France	2007-2010	Colon	Case-control	*E. coli*: detection in (tumor, normal) tissue using microscopy. Cancer: *unk.*	Paired tissue samples (tumor and adjacent normal) from 50 CC patients, tissue samples from 33 non-cancer controls (with uncomplicated diverticulosis).	Significant increased mucosa-associated *E. coli* in tumor tissue (93%) *vs.* adjacent normal tissue (90%) and non-cancer control tissue (88%) (p=0.01 and p<0.001). Significant more mucosa-internalized *E. coli* in tumors (86%) *vs.* adjacent normal (54%) and non-cancer tissue (48%) (p<0.0002). More *E. coli* in normal mucosa distant from tumor in stage III/IV CC *vs.* stage I (p<0.01).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Kohoutova, 2014 [78]	Czech Republic	*unk*	Colorectal	Case-control	*E. coli*: detection in mucosal samples using PCR. Cancer: diagnosis based on colonoscopy.	30 CRC patients, 30 colorectal adenoma patients, 20 healthy controls.	Frequency of *E. coli* significantly higher in CRC mucosal samples compared to healthy controls ($p<0.001$). Statistically higher prevalence of *E. coli* phylogroup B2 in proximal CRC *vs.* distal CRC ($p=0.028$). A positive association between bacteriocinogeny and CRC stage was observed.
Fukugaiti, 2015 [14]	Brazil	*unk*	Colorectal	Case-control	*E. coli*: detection in fecal samples using PCR. Cancer: diagnosis based on colonoscopy.	7 CRC patients, 10 healthy controls.	*E. coli* detected in feces of 7/7 CRC patients and 10/10 controls. No significant difference in fecal abundance of *E. coli* in CRC patients *vs.* controls (6.7 ± 2.4 *vs.* 6.4 ± 2.5 $\log_{10}$ no. of copies, $p=0.78$).
Xie, 2016a* [22]	China	2011-2013	Colorectal	Case-control	*E. coli*: *pks+ E. coli* detection in tissue samples using PCR. Cancer: diagnosis based on colonoscopy and pathology.	Paired tumor and adjacent normal tissue from 36 CRC patients. Normal tissue from 18 healthy controls.	Expression of *pks+ E. coli* was significantly higher for tumor tissue (2.96 ± 0.28) compared to healthy control tissue (1.06 ± 0.08) ($p<0.001$).
Shimpoh, 2017 [81]	Japan	2014-2015	Colorectal	Case-control	*E. coli*: *pks+ E. coli* in colonic lavage samples using PCR. Cancer: diagnosis based on pathology.	Paired colonic lavage and biopsy samples of 13 patients. Colonic lavage samples from 35 CRC patients, 37 colorectal adenoma patients and 26 healthy controls.	*pks+ E. coli* prevalence was 43%, 51% and 46% in CRC, colorectal adenoma and healthy samples respectively. Similarly no significant differences were found in *pks+ E. coli* concentrations among the three groups.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Tsuchiya, 2018 [85]	Bolivia, Chile	2014-2016	Gallbladder	Case-control	*E. coli*: detection of microbes in bile using sequencing of V3-V4 region. Cancer: *unk.*	7 GBC patients, 30 controls with CL.	*E. coli* was detected in bile from both GBC and CL patients.
Tunsjø, 2019 [82]	Norway	2014-2017	Colorectal	Case-control	*E. coli*: detection of *pks, CNF1, tcpC, astA E. coli* genes in fecal and tissue samples using PCR. Cancer: diagnosis based on colonoscopy.	25 CRC patients, 25 patients with adenomatous polyps, 22 healthy controls.	*E. coli* toxin genes in 52% of CRC tissues, 27% of polyp tissues and 45% of healthy control tissue (no significant difference). Also no difference in fecal samples.
Zarei, 2019 [84]	Iran	2015-2017	Colorectal	Case-control	*E. coli*: gene detection in biopsy samples using PCR. Cancer: diagnosis based on colonoscopy.	40 CRC patients, 40 IBD patients, 40 healthy controls.	*E. coli* isolated from 40/48 (83.8%) CRC tissue samples, 40/51 (78.4%) IBD tissue samples and 40/43 (93.0%) healthy tissue samples. In CRC group, phylogroups B2 and A most prevalent.
Iyadorai, 2020 [77]	Malaysia	2014-2015	Colorectal	Case-control	*E. coli*: 16S rRNA and *clbB* gene detection in tissues using PCR. Cancer: *unk.*	18 CRC patients, 23 healthy controls.	*pks+ E. coli* detected in 8/48 (16.7%) of CRC patients *vs.* 1/23 (4.4%) controls. Significantly more CRC tissues *pks+ E. coli*-positive (16/96; 16.7%) compared to control tissues (1/26; 3.8%) (p 0.01).
Pleguezuelos-Manzano, 2020 [79]	The Netherlands	*unk*	Colorectal	Case-control	*E. coli*: *SBS-pks* and *ID-pks* signatures using whole-genome sequencing. Cancer: data of metastatic whole-genome cancer from the Hartwig Medical cohort.	496 CRC metastases, 2,969 metastases of other malignancies.	*SBS-pks* and *ID-pks* positive in 37/496 (7.5%) and 44/496 (8.8%) of CRC metastases *vs.* 12/2,969 (0.4%) and 134/2,969 (4.5%) of other metastases (significantly enriched, p<0.001).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Rezasoltani, 2020* [80]	Iran	2016-2017	Colorectal	Case-control	*E. coli:* detection in fecal samples using qPCR. Cancer: *unk.*	20 CRC patients, 42 patients with adenomatous polyposis, 31 healthy controls.	Higher numbers of *E. coli* in fecal samples from CRC patients and adenomatous polyposis patients *vs.* healthy controls ($p<0.015$).
Tang, 2020 [73]	China	*unk*	Colorectal	Case-control	*E. coli*: detection in fecal samples using genome sequencing and PFGE. Cancer: *unk*	15 CRC patients, 170 healthy controls (68 preschool children, 87 university students, 15 seniors).	Genomic diversity of sampled *E. coli* populations increased with age group. *E. coli* diversity level significantly lower in CRC patients compared to students and seniors.
Yoshikawa, 2020 [83]	Japan	*unk*	Colorectal	Cross-sectional	*E. coli*: detection of colibactin-producing (*clb*+) *E. coli* in tissue samples using PCR. Cancer: adenocarcinoma diagnosis based on histopathology.	34 colorectal adenoma patients. 450 *E. coli* isolates from tumor samples, 279 *E. coli* isolates from non-tumor tissues.	*clb*+ *E. coli* present in 22/34 colorectal adenoma patients. Prevalence of *clb*+ in isolates from tumor tissue higher compared to non-tumor tissue (tumor: 327/450 [72.7%], non-tumor: 123/279 [44.1%]; $p<0.05$). Prevalence of *clb*+ *E. coli* not associated with tumor location.

*Only abstract available. PCR: polymerase chain reaction. CRC: colorectal cancer. *Unk:* unknown. CC: colon cancer. IBD: inflammatory bowel disease. PFGE: pulsed field gel electrophoresis.

Supplementary Table 10. Characteristics and main outcomes of epidemiological studies assessing the association between ***Fusobacterium nucleatum*** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Castellarin, 2012 [89]	Canada	*unk*	Colorectal	Cross-sectional	*Fn*: detection in tissue samples using sequencing. Cancer: *unk.*	Paired tissue samples (tumor and normal tissue) from 11 CRC patients.	9/11 (81.8%) tissue pairs showed over 2-fold higher abundance in tumor tissue compared to normal tissue.
Flanagan, 2014 [90]	Czech Republic, Germany, Ireland	2008-2010	Colorectal	Case-control	*Fn*: detection in tissue samples and fecal samples by qPCR. Cancer: diagnosis partially based on colonoscopy.	Tumor tissue from 122 CRC patients, matched normal tissue from 105 patients. Adenoma tissue and matched normal tissue from 52 patients. Fecal samples from 7 CRC patients, 24 adenoma patients and 25 healthy controls.	Relative quantification of *Fn* in tumor tissue significantly higher compared to normal tissue (range 2^{-29} – 2^{-6} in tumor tissue *vs.* 2^{-30} – 2^{-14} in normal tissue). *Fn* more abundant in fecal samples from CRC patients compared to the control groups.
Fukugaiti, 2015 [14]	Brazil	*unk*	Colorectal	Case-control	*Fn*: detection in fecal samples using PCR. Cancer: diagnosis based on colonoscopy.	7 CRC patients, 10 healthy controls.	*Fn* detected in 7/7 CRC patients and 9/10 controls. Significant higher abundance of *Fn* in CRC patients *vs.* controls (6.2 ± 1.5 *vs.* 4.0 ± 1.5 $\log_{10}$ no. of copies, p=0.01).
Mira-Pascual, 2015 [92]	Spain	*unk*	Colorectal	Case-control	*Fn*: detection in fecal samples using pyro-sequencing of 16S rRNA gene and PCR. Cancer: diagnosis based on colonoscopy.	7 CRC patients, 11 patients with tubular adenomas, 10 healthy controls.	Higher abundance of *Fn* in CRC patients *vs.* controls (9/15, 60% *vs.* 2/9, 22%; p<0.07).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Ito, 2015 [91]	Japan	2001-2013	Colorectal	Case-control	*Fn*: detection in tumor tissue samples using PCR. Cancer: *unk.*	544 CRC patients, 343 patients with serrated lesions, 122 patients with non-serrated adenomas.	511/544 (93.9%) of tissue samples from CRC patients *Fn*-positive. *Fn*-positivity significantly higher in CRC patients *vs.* patients with precancerous lesions ($p<0.0001$). Gradual increase in *Fn* abundance from sigmoid to proximal colon in sessile serrated adenoma patients.
Liang, 2016 [41]	China	2009-2015	Colorectal	Case-control	*Fn*: detection in fecal samples using qPCR. Cancer: *unk.*	203 CRC patients, 236 healthy controls.	Relative abundance of *Fn* significantly higher in CRC patients *vs.* healthy controls ($p<0.0001$).
Wang, 2016 [96]	China	June-Dec 2013	Colorectal	Case-control	*Fn*: IgA and IgG antibody detection in serum samples using indirect whole-cell ELISA. Cancer: diagnosis based on histology.	258 CRC patients, 150 benign colon disease patients, 200 healthy controls.	IgA and IgG titers significantly higher ($p<0.001$) in CRC patients *vs.* the control groups. Anti-*Fn*-IgA: average absorbance ± SD in CRC patients 0.390 ± 0.215; benign colon disease group 0.268 ± 0.158; healthy controls 0.246 ± 0.132. anti-*Fn*-IgG: average absorbance ± SD in CRC patients 0.362 ± 0.194; benign colon disease group 0.270 ± 0.162; healthy controls 0.262 ± 0.152.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Xie, 2016b* [97]	China	2010-2016	Colorectal	Case-control	*Fn*: detection in fecal samples using qPCR. Cancer: diagnosis based on endoscopy and pathological findings.	62 CRC patients (38 early stage, 24 late stage), 38 adenoma patients and 54 healthy controls.	Significant increasing trend in *Fn* abundance from healthy controls, early stage CRC to late stage CRC ($p<0.01$).
Yu, 2016 [98]	China	2014-2015	Colorectal	Case-control	*Fn*: detection in tissue samples using 16S rRNA FISH. Cancer: diagnosis based on colonoscopy.	48 proximal CRC patients, 45 distal CRC patients, 79 patients with traditional adenomas, 40 patients with hyperplastic polyps, 68 patients with serrated polyps, 20 mucosal tissues from healthy controls.	Invasive *Fn* more frequently detected in proximal CRC vs distal CRC (89.6% *vs.* 42.2%, $p<0.05$). Presence of bacterial biofilms slightly higher in proximal *vs.* distal CRC (52.1% *vs.* 48.9%).
Amitay, 2017 [93]	Germany	2005-2013	Colorectal	Case-control	*Fn*: detection in fecal samples using 16S rRNA gene analysis, PCR. Cancer: diagnosis based on colonoscopy.	46 CRC patients, 113 AA patients, 110 NAA patients, 231 healthy controls.	*Fn* more frequently detected in CRC patients *vs.* other three groups (54.3% *vs.* 23.6-25.1%, $p<0.001$). Relative abundance higher with increasing CRC stage ($p=0.049$).
Drewes, 2017 [94]	Malaysia	*unk*	Colon	Case-control	*Fn*: detection in tissue samples using 16S rRNA gene amplicon sequencing and FISH (for biofilm quantification). Cancer: *unk.*	23 paired tissue samples (tumor and adjacent normal tissue) of proximal CC patients.	Higher abundance of invasive biofilms on proximal CC tumors compared to normal tissue. *Fn* more frequently present in tumors with biofilms (16/17; 94.1%) compared to normal tissue.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Eklöf, 2017 [95]	Sweden	2008-2013	Colorectal	Case-control	*Fn*: detection in fecal samples using PCR. Cancer: diagnosis based on colonoscopy.	39 CRC patients, 135 patients with low- and high-grade dysplasia, 66 healthy controls.	*Fn* present in fecal samples of all study participants, although at significantly higher levels in CRC patients compared to dysplasia patients and controls ($p<0.001$).
Scott, 2017* [99]	UK	*unk*	Colon	Case-control	*Fn.*: detection in (tumor) tissue using 16S rRNA gene sequencing. Cancer: diagnosis based on histology.	17 CC patients, 37 patients with suspected appendicitis.	Higher abundance of *Fn* in appendixes of CC patients compared to adjacent normal tissue ($p=0.059$).
Suehiro, 2017 [100]	Japan	*unk*	Colorectal	Case-control	*Fn*: detection in fecal samples using PCR. Cancer: diagnosis based on colonoscopy.	158 CRC patients, 19 patients with AA or carcinoma *in situ,* 11 NAA patients and 60 healthy controls.	Significant higher abundance of *Fn* in CRC patients (median copy number: 317 *vs.* control group 17.5; $p<0.0001$).
Yamamura, 2017 [101]	Japan	*unk*	Esophagus, Stomach, Colorectal, Liver, Pancreas	Cross-sectional	*Fn:* detection in tissue samples using qPCR. Cancer: histopathological cancer confirmation.	20 paired tissue samples (tumor and adjacent normal tissue) for each cancer included.	*Fn* DNA detection: Esophagus: 4/20 tumor tissue; 1/20 normal tissue. Stomach: 2/20 tumor tissue; 0/20 normal tissue. Colorectal: 9/20 tumor tissue; 8/20 tumor tissue. Liver and pancreatic tissues: no DNA detected in tumor nor normal tissues.
Yoon, 2017 [102]	Korea	*unk*	Colorectal	Case-control	*Fn*: detection in mucosal tissue samples using 16S rRNA gene pyrosequencing. Cancer: *unk.*	6 CRC patients, 6 patients with conventional adenoma, 6 patients with sessile serrated adenoma and 6 healthy controls.	*Fn* not detected in samples from CRC patients. Diversity of mucosal communities lowest in CRC patients as compared to other three groups.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Yu, 2017 [103]	China, Denmark	*unk*	Colorectal	Case-control	*Fn*: detection in fecal samples using metagenomic sequencing and qPCR. Cancer: diagnosis based on colonoscopy.	Chinese cohort1: 74 CRC patients, 54 controls. Chinese cohort2: 47 CRC patients, 109 controls. Denmark: 16 CRC patients, 24 controls.	*Fn* enriched in fecal samples of CRC patients. *Fn* gene present in 4/109 (3.7%) control samples, potential suitable marker for CRC.
Hale, 2018 [15]	USA	*unk*	Colorectal	Cross-sectional	*Fn*: detection in (tumor and adjacent normal) tissue using 16S rRNA gene sequencing Cancer: *unk.*	83 CRC patients (25 with deficient mismatch repair [dMMR] CRC, 58 with proficient mismatch repair [pMMR] CRC).	*Fn* significantly more abundant in tumor tissue of dMMR CRC patients as compared with normal tissue (p=0.03), not in pMMR CRC patients. Abundance of *Fusobacterium* spp. more frequently observed in proximal *vs.* distal CC.
Hsieh, 2018 [44]	Taiwan	*unk*	Stomach	Case-control	*Fn*: detection in tissue samples using 16S rRNA gene sequencing. Cancer: *unk.*	11 GC patients, 9 gastritis patients, 7 intestinal metaplasia patients.	*Fn* significantly enriched in gastric tumor tissue compared to gastritis tissue (p<0.01).
Kwong, 2018 [17]	China	2006-2015	Colorectal	Retrospective cohort	*Fn*: culture-confirmed bacteremia. Cancer: *unk.*	79 *Fn* bacteremia patients, 395 matched controls without history of bacteremia.	4/79 (5.1%) of the bacteremia patients developed CRC compared to 4/395 (1.0%) of the controls, aHR: 6.89 (95%CI 1.70-27.9; p<0.007).
Repass, 2018 [105]	USA	*unk*	Colorectal	Cross-sectional	*Fn*: detection in tissue samples using sequencing and PCR. Cancer: *unk.*	16 CRC tissue samples, 10 adjacent normal tissue samples.	*Fn* present in 25% of CRC samples *vs.* 15% and 0% of adjacent normal and matched normal tissue respectively.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Russo, 2018 [106]	Italy	2015-2016	Colorectal	Case-control	*Fn*: detection in saliva, feces and tumor tissue samples using 16S rRNA gene sequencing and qPCR. Cancer: diagnosis confirmed by histological analysis.	10 CRC patients, 10 healthy controls.	No significant difference in observed *Fn* abundance in fecal samples of CRC patients *vs.* controls. Higher abundance in saliva compared to feces in both CRC patients ($p<0.01$) and controls ($p<0.002$).
Tsuchiya, 2018 [85]	Bolivia, Chile	2014-2016	Gallbladder	Case-control	*Fn*: detection of microbes in bile using sequencing of V3-V4 region. Cancer: *unk.*	7 GBC patients, 30 controls with CL.	*Fn* detected in bile of GBC patients.
Bundgaard, 2019 [13]	Denmark	2002-2010	Colorectal	Prospective cohort	*Fn*: detection in (tumor and adjacent normal) tissue samples using qPCR. Cancer: *unk.*	99 CRC patients, 96 adenoma patients, 104 patients with diverticular disease.	*Fn* detected in 29.3% of the tumor samples. Detection rate lowest for adenoma samples as compared to tumor tissue, adjacent normal and diverticular tissue.
Butt, 2019 [104]	10 European countries	1992-2000	Colorectal	Case-control	*Fn*: detection of antibodies against 11 *Fn* proteins in serum using multiplex serology. Cancer: diagnosis according to ICD-10 C18-C20 codes.	485 CRC patients (samples drawn 0.4-8.5 years before CRC diagnosis), 485 healthy controls.	Positivity to: ≥1 protein in 47% of patients *vs.* 53% of controls (OR 0.81; 95%CI 0.62-1.06). ≥2 proteins: 17% patients *vs.* 21% controls. ≥3 proteins: 9% patients and 9% controls. Positivity to a either single or multiple *Fn* proteins not associated with higher cancer risk.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Chen, 2019 [107]	China	2012-2014	Stomach	Cross-sectional	*Fn*: detection in tissue samples using 16S rRNA gene sequencing. Cancer: *unk.*	62 paired GC tissue samples (tumor tissue and adjacent normal tissue).	Relative abundance of *Fn* in tumor tissue significantly higher compared to adjacent normal tissue (0.257% *vs.* 0.041%) (p<0.001).
de Carvalho, 2019 [108]	Brazil	2008-2015	Colorectal	Prospective cohort	*Fn*: detection in (tumor and adjacent normal) tissue using 16S rRNA gene sequencing. Cancer: *unk.*	152 CRC patients.	*Fn* present in 35/152 (23.0%) of the tumor tissue samples and 6/57 (10.5%) of the adjacent normal tissue samples. Higher levels of *Fn* were found in tumor compared to adjacent normal (p=0.0033). Patients with *Fn* in tumor tissue were more likely to have proximal tumors (p=0.001), more advanced stage of cancer (p=0.033), higher level of invasion (p=0.014) and lower differentiation grade (p=0.011).
Kageyama, 2019 [109]	Japan	2015-2017	Esophagus, Stomach, Colorectal	Case-control	*Fn*: detection in saliva samples using 16S rRNA gene sequencing. Cancer: *unk.*	59 patients with cancer in the gastrointestinal tract, 118 healthy controls.	*Fn* significantly more abundant in saliva of esophageal cancer patients as compared to controls (p<0.05).
Saito, 2019 [110]	Japan	*unk*	Colorectal	Case-control	*Fn*: detection in colonoscopy aspirates using 16S rRNA gene sequencing. Cancer: diagnosis based on colonoscopy.	24 intramucosal CRC patients (stage 0 cancer), 47 colorectal adenoma patients, 10 healthy controls.	60% of the *Fusobacterium* spp. were *Fn* in healthy controls. Overabundance of *Fusobacterium* spp. in intramucosal CRC patients *vs.* healthy controls.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Tunsjø, 2019 [82]	Norway	2014-2017	Colorectal	Case-control	*Fn*: detection in fecal and tissue samples using qPCR. Cancer: diagnosis based on colonoscopy.	25 CRC patients, 25 patients with adenomatous polyps, 22 healthy controls.	*Fn* significantly more abundant in CRC patients compared to the group with polyps ($p<0.003$) and the healthy controls ($p<0.008$). *Fn* detected in tumor tissue from 13 patients, of which 9 (69.2%) also had high *Fn* levels in feces.
Yachida, 2019 [111]	Japan	*unk*	Colorectal	Case-control	*Fn*: detection in fecal samples using whole-genome shotgun sequencing. Cancer: diagnosis based on colonoscopy and histological findings.	225 CRC patients, 140 adenoma patients, 251 healthy controls.	Significant increased abundance of *Fn* in pre-malignant cancer stages as well as the different CRC stages ($p<0.005$).
Zhang, 2019* [117]	USA	*unk*	Esophagus	Case-control	*Fn*: detection in tissue samples using WGS. Cancer: *unk.*	6 patients with esophageal adenocarcinoma, 8 patients with non-dysplastic Barrett's esophagus and 9 healthy controls.	High abundance of *Fn* in samples from esophageal adenocarcinoma patients.
Alkharaan, 2020 [112]	Sweden	2017-2019	Pancreas	Case-control	*Fn*: DNA detection using qPCR and antibody detection using ELISA in plasma samples and saliva samples. Cancer: dysplasia or cancer diagnosis based on histopathology.	46 cancer patients, 45 low-grade dysplasia IPMN patients, 18 non-IPMN controls.	IgG binding reactivities in plasma significantly higher in cancer patients *vs.* non-IPMN control group ($p<0.0006$). Salivary IgA reactivity to *Fn* higher in the cancer group compared to controls (based on 65 samples) ($p<0.007$).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Boehm, 2020 [113]	Lithuania, Germany	*unk*	Stomach, Colorectal	Case-control	*Fn*: detection in tissue samples using qPCR. Cancer: diagnosis based on histology.	81 paired GC samples and 27 paired CRC samples (tumor and adjacent normal tissue), 17 gastritis patients, 9 atrophic gastritis or intestinal metaplasia patients.	*Fn* present in 16/27 (59.3%) CRC tumor tissue *vs.* 13/26 (50.0%) normal colon tissue samples. *Fn* load significantly correlated between CRC and GC tumor and normal tissues. *Fn* present in 23/80 (28.8%) GC tumor tissue *vs.* 18/78 (23.1%) normal gastric tissue samples.
Gantuya, 2020 [114]	Mongolia	2014-2016	Stomach	Cross-sectional	*Fn*: detection in tissue samples using 16S rRNA gene sequencing. Cancer: diagnosis based on histopathology.	48 GC patients, 120 non-cancer patients (20 healthy controls, 20 gastritis patients, 40 atrophy patients, 40 intestinal metaplasia patients).	*Fn* significantly enriched in GC tissue compared to tissue from intestinal metaplasia patients.
Kashani, 2020 [115]	Iran	2017-2019	Colorectal	Case-control	*Fn*: detection in tissue samples using PCR. Cancer: diagnosis based on colonoscopy and pathological findings.	35 CRC patients, 45 controls with colorectal disorders other than cancer.	*Fn* present in 24/35 (68.6%) of the CRC patients, of which 11 had the *fad*A gene-positive *Fn*. *Fn* present in 11/45 (24.4%) of the controls, none of which carried the *fad*A gene.
Reynolds, 2020 [116]	Ireland	2008-2017	Rectum	Cross-sectional	*Fn*: detection in tumor and adjacent normal tissue samples using WGS. Cancer: diagnosis based on biopsy, meeting the WHO mucinous rectal cancer diagnostic criteria.	10 patients with mucinous rectal cancer.	*Fn* present in 10/10 tumor tissue samples and 9/10 adjacent normal tissue samples.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Eisele, 2021 [121]	Germany, USA	2010-2018	Colorectal	Prospective cohort	*Fn*: detection in fecal samples using qPCR. Cancer: *unk.*	105 CRC patients.	22/105 (21.0%) of the CRC patients were classified as having high levels of *Fn* in feces (remaining 83 had none/low levels of *Fn*). OR for getting diagnosed with rectal cancer *vs.* colon cancer was 3.01 (95%CI 1.06-8.57; p=0.04) for those with high levels of *Fn.* In the same group, the OR for rectal cancer *vs.* proximal CC was 5.32 (95%CI 1.23-22.98; p=0.03).
Kawasaki, 2021 [118]	Japan	2018-2020	Esophagus	Case-control	*Fn*: detection in dental plaque samples, saliva and mucus samples from cancer tissue, using PCR. Cancer: *unk.*	61 esophageal cancer patients, 62 healthy controls.	*Fn* detected in mucus of 35/37 (94.6%) esophageal cancer patients and in the saliva samples of all cancer patients and healthy controls.
Kurt, 2021 [119]	Turkey	2018-2019	Colorectal	Case-control	*Fn*: detection in tissue samples by qPCR, serum antibody detection using ELISA. Cancer: diagnosis based on colonoscopy.	22 CRC patients, 35 precancerous-benign colon disease patients, 21 healthy controls.	*Fn* DNA present in tissue samples of 13/15 (86.7%) CRC patients *vs.* 19/26 (73.1%) P-BCD patients. Anti-*Fn*-IgA and -IgG were significantly higher in CRC patients *vs.* healthy controls.
Liang, 2021 [46]	China	2009-2014	Colorectal	Case-control	*Fn*: detection in fecal samples using qPCR. Cancer: diagnosis based on colonoscopic examination and histopathological review	210 CRC patients, 115 AA patients, 86 NAA patients, 265 healthy controls.	*Fn* significantly more abundant in CRC patients compared to controls (p<0.0001). *Fn* showed to be a potential marker for detection of CRC (sensitivity 81.8%; specificity 52.8%; positive predictive value 61.4%).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Pignatelli, 2021 [120]	Italy	2018-2019	Colon	Cross-sectional	*Fn*: detection in oral and (tumor and adjacent normal) tissue samples using qPCR. Cancer: *unk.*	36 CC patients.	Abundance of *Fn* was significantly higher in oral samples (median 108.7 CFU/mL) compared to tumor tissue (median 4.8 CFU/mL) or adjacent normal tissue (2.19 CFU/mL). Difference in abundance between tumor tissue and adjacent normal tissue not significant. Abundance was not significantly different between cancer subsites (proximal, distal).

*Only abstract available. *Fn: Fusobacterium nucleatum.* RFLP: restriction fragment length polymorphism. GC: gastric cancer. CRC: colorectal cancer. (q)PCR: (quantitative) polymerase chain reaction. SD: standard deviation. FISH: fluorescence in situ hybridization. NAA: non-advanced adenoma. AA: advanced adenoma. CC: colon cancer. ESCC: esophageal squamous cell carcinoma. EAC: esophageal adenocarcinoma. ICD-10: International Classification of Diseases (10th revision). aHR: adjusted hazard ratio. OR: odds ratio. ELISA: enzyme-linked immunosorbent assay. WGS: whole-genome sequencing. WHO: World Health Organization. P-BCD: precancerous-benign colon disease. IPMN: intraductal papillary mucinous neoplasm.

Supplementary Table 11. Characteristics and main outcomes of epidemiological studies assessing the association between ***Porphyromonas gingivalis*** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Ahn, 2012 [128]	USA	1988-2006	GI cancers combined (esophagus, stomach, colon, rectum, anus, liver, pancreas)	Prospective cohort	*Pg*: detection of serum IgG titers using ELISA. Cancer: diagnosis according to ICD-10 C15, C16, C18-C22, C25 codes.	817 orodigestive cancer deaths, 7,765 healthy controls.	Orodigestive cancer mortality in individuals with IgG titers against *Pg* ≥69.1 *vs.* <69.1: RR 2.25 (95%CI 1.23-4.14).
Michaud, 2013 [132]	Denmark, France Germany, Greece, Italy, The Netherlands, Norway, Spain, Sweden, United Kingdom	1992-2000	Pancreas	Prospective cohort	*Pg*: detection of antibody titers in serum samples using an immunoblot array. Cancer: data obtained from cancer registries, pathology based diagnosis.	405 pancreatic cancer patients, 416 healthy controls.	Individuals with antibody titers >200 ng/ml had an over 2-fold higher risk of pancreatic cancer compared to individuals with an antibody titer of ≤200 ng/ml (OR 2.14; 95%CI 1.05-4.36). Mean follow-up time was 5 years.
Gao, 2016 [130]	China	2010-2014	Esophagus	Case-control	*Pg*: immuno-histochemial detection of *Pg* and gingipain Kgp (protease) and PCR for presence of rDNA in tumors *vs.* healthy tissues. Cancer: *unk.*	100 ESCC tissue samples, 130 samples from healthy controls.	*Pg* present in: 61/100 cancer tissues, 12/100 adjacent normal tissues, 0/30 normal esophageal mucosa. *Pg* positively correlated to poor differentiation, severe lymph node metastasis and cancer stage.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Sinha, 2016 [135]	USA	1985-1987	Colorectal	Case-control	*Pg*: detection in fecal samples using 16S rRNA gene sequencing. Cancer: newly diagnosed with colorectal adenocarcinoma.	42 CRC patients, 89 healthy controls.	Significant higher abundance of *Pg* in CRC patients *vs.* controls (OR 3.83; 95%CI 1.03-14.22).
Peters, 2017 [133]	USA	1993-2010	Esophagus	Case–control study nested within a prospective cohort	*Pg*: detection in pre-diagnostic oral wash samples using 16S rRNA gene sequencing. Cancer: diagnosis according to ICD codes. EAC: 8140, 8144, 8480, 8481, 8560. ESCC: 8070, 8071, 8072, 8074, 8052.	81 EAC patients, 160 controls. 25 ESCC patients, 50 controls.	*Pg* present in higher number of cancer patients *vs.* controls: EAC: OR 1.06 (95%CI 0.93–1.20). ESCC: OR 1.30 (95%CI 0.96-1.77). Follow-up time between oral sample and cancer diagnosis <1-9 years.
Gao, 2018 [131]	China	*unk*	Esophagus	Case-control	*Pg*: serum IgA and IgG titers measured by ELISA. Cancer: *unk.*	96 ESCC patients, 130 healthy controls.	Median serum titers in ESCC patients *vs.* controls: IgG: 150.7 EU *vs.* 109.1 EU ($p<0.001$). IgA: 33.2 EU *vs.* 19.1 EU ($p<0.01$). Diagnostic potential: sensitivity: IgG 29.2%; IgA 52.1%. specificity: IgG 96.9%; IgA 70.8%.
Fan, 2018 [129]	USA	2002-2008, 1993-2010	Pancreas	Case–control study nested within a prospective cohort	*Pg*: detection in pre-diagnostic oral wash samples using 16S rRNA gene sequencing. Cancer: diagnosis according to ICD–O–2 codes C25.0–C25.3, C25.7–25.9.	361 cancer patients, 371 healthy controls.	*Pg* present in higher number of cancer patients *vs.* controls, OR 1.60 (95%CI 1.15-2.22). Oral samples collected up to 10 years before cancer diagnosis.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Kageyama, 2019 [109]	Japan	2015-2017	Esophagus, Stomach, Colorectal	Case-control	*Pg*: detection in saliva samples using 16S rRNA gene sequencing. Cancer: *unk.*	59 patients with cancer in the gastrointestinal tract, 118 healthy controls.	*Pg* significantly more abundant in saliva of EAC, GC and CRC patients as compared to controls ($p<0.05$).
Wei, 2020 [32]	China	2017-2018	Pancreas	Case-control	*Pg*: detection in saliva samples using 16S rRNA gene sequencing. Cancer: histopathologically-confirmed pancreatic adenocarcinoma.	41 PDAC patients, 69 healthy controls.	Significant reduced abundance of *Pg* in saliva of PDAC patients *vs.* controls.
Kawasaki, 2021 [118]	Japan	2018-2020	Esophagus	Case-control	*Pg*: detection in dental plaque samples, saliva and mucus samples from cancer tissue, using PCR. Cancer: *unk.*	61 patients with esophageal cancer, 62 healthy controls.	*Pg* detected in mucus of 13/37 (35.1%) esophageal cancer patients and in the saliva samples of 48/61 (78.7%) of the cancer patients and 49/62 (79.0%) of the healthy controls.
Pignatelli, 2021 [120]	Italy	2018-2019	Colon	Cross-sectional	*Pg*: DNA detection in oral and (tumor and adjacent normal) tissue samples using qPCR. Cancer: *unk.*	36 patients with CC.	*Pg* present in oral samples but none of the colon samples (neither tumor tissue nor adjacent normal tissue).
Wang, 2021 [134]	China	2012-2018	Colorectal	Case-control	*Pg*: detection in fecal and tissue samples using 16S rDNA gene sequencing and IHC staining of tissue samples. Cancer: *unk.*	23 CRC patients, 32 patients with CRC adenoma, 22 healthy controls.	High levels of *Pg* expression in 40.9% of CRC patients *vs.* 15.0% in adenoma group and 10.0% in controls. *Pg* significantly enriched in feces and CRC tissue compared to controls.

GI: gastrointestinal. *Pg: Porphyromonas gingivalis.* ELISA: enzyme-linked immunosorbent assay. ICD: International Classification of Diseases. RR: relative risk. OR: odds ratio. PCR: polymerase chain reaction. ESCC: esophageal squamous cell carcinoma. CRC: colorectal cancer. EAC: esophageal adenocarcinoma. EU: ELISA unit. PDAC: pancreatic adenocarcinoma.

Supplementary Table 12. Characteristics and main outcomes of epidemiological studies assessing the association between **nontyphoidal *Salmonella*** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Kato, 2013 [140]	USA, The Netherlands	2003-2005	Colorectal	Case-control	NTS: detection of *Salmonella* antiflagellin (FliC) IgG in blood samples using ELISA. Cancer: diagnosis based on histology.	70 CRC patients, 23 controls with colorectal polyps, 74 healthy controls (from two independent cohorts).	Significantly higher Flic antibody titers in CRC patients and patients with colorectal polyps *vs.* healthy controls (Dutch cohort: 3.93 *vs.* 2.23; p=0.014; US cohort: 6.65 *vs.* 4.37; p<0.001).
Iyer, 2016 [139]	India	*unk*	Gallbladder	Cross-sectional	NTS: DNA detection in tissue using PCR and whole exosome sequencing. Cancer: diagnosis based on histopathology.	17 tumor tissues from GBC patients, 9 matched adjacent normal tissue.	None of gallbladder samples tested NTS-positive. Traces of typhoidal *Salmonella* found in 12/26 (46.2%) gallbladder isolates. *S.* Typhimurium in 10/26 (38.5%) samples, *S.* Choleraesuis in 5/26 samples (19.2%).
Lu, 2017 [141]	USA	*unk*	Colorectal	Case-control	NTS: AvrA protein detection in fecal samples using rt-PCR, AvrA staining and IHC of human TMA Cancer: diagnosis based on histopathology.	TMAs from 48 CRC tumors, 13 adjacent normal mucosa, 14 metastasized lymph nodes, 61 benign lesions, 19 healthy colorectal mucosa. 24 fecal samples from healthy controls.	Significantly lower mean normalized AvrA staining score in CRC tumors (1.37) *vs.* healthy mucosa (1.96) (p=0.013). Significantly higher staining score in adjacent normal mucosa (2.72) *vs.* healthy mucosa (1.96) (p=0.018). IHC-results: AvrA present in cancer tissue, absent in normal mucosa. Fecal samples exhibited amplification for AvrA.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Mughini-Gras, 2018 [142]	The Netherlands	2000-2015	Colon	Retrospective cohort	NTS: culture confirmed infection (feces, blood, other). Cancer: diagnosis according to ICD-10 C180-C187 codes.	96 CC patients. Reference population: ~17 million.	The SIRs of developing CC after NTS infection expressed as compared to the general population were: - Overall CC: 1.17 (95%CI 0.95-1.43). - Proximal CC: 1.48 (95%CI 1.14-1.88; p<0.01) - Distal CC: 0.86 (95%CI 0.57-1.24). In patients aged <60 years when infected a higher risk was observed: SIR 1.54 (95%CI 1.09-2.10).
De Savornin-Lohman, 2020 [29]	The Netherlands	2000-2016	Biliary tract	Retrospective cohort	NTS: culture confirmed infection (feces, blood, other). Cancer: diagnosis according to ICD-O C239, C240, C242-C244, C248, C249 codes.	9 patients with biliary tract cancer. Reference population: ~17 million.	Individuals with a history of NTS infection had a 1.5-fold increased risk of developing biliary tract cancer as compared to the general population (SIR 1.53; 95%CI 0.70-2.91). The risk was slightly higher for people infected before 60 years of age: SIR 1.74 (95%CI 0.36–5.04).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Chang, 2021 [137]	Taiwan	2000-2013	Stomach, Colorectal, Liver, Biliary tract, Pancreas	Retrospective cohort	NTS: culture confirmed infection (feces, blood) with hospitalization. Cancer: diagnosis according to ICD-9 151, 153-157 codes.	9097 NTS-patients, 9097 non-NTS controls. 192 cancer patients (stomach n=37, colorectal n=55, biliary tract n=11, liver n=70, pancreas n=7).	The HRs of developing cancer after NTS infection as compared to non-NTS controls were: - Stomach 2.02 (95%CI 1.18-3.45); - Colorectal: 1.04 (95%CI 0.72-1.50); - Liver 1.03 (95%CI 0.72-1.47); - Biliary tract: 1.79 (95%CI 0.65-4.97); - Pancreas: 0.81 (95%CI 0.29-2.30). All combined 1.24 (95%CI 1.01-1.53).
Duijster, 2021 [138]	Denmark	1994-2016	Colon	Retrospective cohort	NTS: culture confirmed infection (feces, blood, other). Cancer: diagnosis according to ICD-10 C180-C187 codes.	245 CC patients. Reference population: 7.6 million.	The HRs of developing CC after NTS infection as compared to non-NTS controls were: - Overall CC 0.99 (95%CI 0.88-1.13); - Proximal CC: 1.09 (95%CI 0.93-1.29); - Distal CC: 0.87 (95%CI 0.71-1.05).

NTS: nontyphoidal *Salmonella.* ELISA: enzyme-linked immunosorbent assay. CRC: colorectal cancer. IHC: immunohistochemistry. TMA: tissue microarray. ICD: International Classification of Diseases. CC: colon cancer. SIR: standardized incidence ratio. HR: hazard ratio.

Supplementary Table 13. Characteristics and main outcomes of epidemiological studies assessing the association between ***Salmonella* Typhi** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Welton, 1979 [163]	USA	1922-1975	Liver, Biliary tract, Gallbladder	Case–control	ST: detection in fecal samples using culturing methods. Cancer: death due to hepatobiliary tract cancer.	471 *S. Typhi* carriers, 942 controls.	Risk of death due to hepatobiliary cancer 6 times higher in chronic typhoid carriers compared to controls ($p<0.001$).
Mellemgaard, 1988 [150]	Denmark	1943-1982	Liver, Biliary tract, Gallbladder	Retrospective cohort	ST: registered chronic typhoid carriers. Cancer: *unk.*	219 typhoid carriers of which 3 with hepatobiliary tract cancer.	Significant increased risk of hepatobiliary tract cancer among chronic typhoid carriers: SIR 3.85 (90%CI 1.05-9.94) (3 observed cancers, 0.78 expected cancers).
Csendes, 1994* [147]	Chile	*unk*	Gallbladder	Cross-sectional	ST: detection in bile by culturing methods. Cancer: *unk.*	58 GBC patients, 67 common bile duct stone patients, 165 CL patients, 46 acute cholecystitis patients, 36 healthy controls.	Pathogenic bacteria present in 47/58 (81.0%) of the GBC patients *vs.* 52/165 (31.5%) of the CL patients.
Caygill, 1994 [146]	United Kingdom	1964-*unk*	Colorectal, Gallbladder, Pancreas	Retrospective cohort	ST: *unk.* Cancer: *unk.*	83 chronic carriers of *S.* Typhi or Paratyphi, 386 individuals with history of acute *S.* Typhi (without chronic carriage).	SIR GBC 167.0 (95%CI 54.1-389). Also 8-fold and 3-fold increased risk of pancreatic cancer and CRC respectively.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Strom, 1995 [160]	Bolivia, Mexico	1984-1988	Gallbladder	Case-control	ST: detection of antibodies using ELISA. Cancer: diagnosis based on histology.	84 GBC/biliary tract cancer patients, 264 CL or choledocholithiasis patients and 126 controls without biliary stones. (not all included in serological analysis).	Seropositivity in 7/15 (46.7%) cancer patients, 5/10 (50.0%) CL or choledocholithiasis patients and 4/8 (50.0%) controls. Significant elevated risk of self-reported history of physician-diagnosed typhoid fever (OR 12.7; 95%CI 1.5-598).
Singh, 1996 [159]	India	*unk*	Gallbladder	Case-control	ST: detection in bile samples using culturing methods. Cancer: *unk.*	38 GBC patients, 67 CL patients.	*S*T present in bile of 4/38 (10.5%) GBC patients and 2/67 (3.0%) CL patients. No mixed infections with other pathogens in GBC patients (i.e. based on bile samples).
Nath, 1997* [151]	India	*unk*	Gallbladder	Case-control	ST: detection in bile samples using culturing methods. Cancer: *unk.*	28 GBC patients, 56 CL patients and 17 healthy controls.	Significant higher prevalence of *S*T in GBC patients *vs.* CL patients and controls ($p<0.05$).
Roa, 1999* [154]	Chile	*unk*	Gallbladder	Cross-sectional	ST: detection in bile samples using culturing methods. Cancer: *unk*	24 GBC patients, 468 chronic cholecystitis\\s patients, 140 acute cholecystitis patients, 5 gallbladder dysplasia patients.	13/29 (44.8%) of patients with gallbladder cancer or dysplasia had positive bacterial cultures in bile, not including *S*T.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Dutta, 2000 [148]	India	1994-1995	Gallbladder	Case-control	ST: detection of antibodies against *S*T Vi antigen using ELISA. Cancer: gallbladder mass on ultrasound or computed tomography.	37 GBC patients with concurrent CL, 80 cholelithiasis patients with without cancer.	Antibody positivity in 6/37 (16.2%) of GBC patients *vs.* 2/80 (2.5%) of controls without cancer. 14-fold increased risk of GBC in chronic typhoid carriers with CL.
Shukla, 2000 [158]	India	*unk*	Gallbladder	Case-control	ST: detection of antibodies against *S*T Vi antigen (ViAb) in serum using indirect haemagglutination assay. Cancer: *unk*.	51 GBC patients, 56 patients with cholelithiasis, 40 healthy controls.	Significantly higher ViAb positivity in GBC patients (29%) *vs.* healthy controls (5%) (OR 7.92; 95%CI 1.69-37.09, p<0.01). No significant difference between CL patients and controls (OR 3.47; 95%CI 0.44-11.93, p>0.05).
Serra, 2002 [156]	Chile	1992-1995	Gallbladder	Case-control	ST: self-reported history of *S. Typhi* infection. Cancer: diagnosis based on histology.	114 GBC patients, 114 hospitalized controls with CL.	9.7% of GBC patients reported a history of *S*T infection *vs.* 7.5% of the controls.
Pandey, 2003 [153]	India	*unk*	Gallbladder	Case-control	ST: self-reported history of *S. Typhi* infection. Cancer: *unk*.	64 GBC patients, 101 controls with CL.	14/64 (21.9%) of GBC patients reported a history of typhoid fever *vs.* 13/101 (12.9%) of the controls.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Hazrah, 2004 [149]	India	1997-1999	Gallbladder	Cross-sectional	ST: culturing bacteria from the core of the gallstones. Cancer: *unk.*	14 GBC patients, 83 patients with cholecystitis or choledocholithiasis, 2 patients with empyema, 1 patient with periampullary carcinoma. All 100 patients had CL as well.	*Salmonella* spp. present in 1.5% of patients and controls combined.
Yagyu, 2004 [164]	Japan	1988-1990	Gallbladder	Prospective cohort	ST: self-reported history of *ST* infection. Cancer: GBC diagnosis using ICD-9 classification.	113,394 individuals (healthy at start of study).	2/116 (1.7%) individuals who died from GBC during a mean of 9.7 years of follow-up had a self-reported history of typhoid fever.
Vaishnavi, 2005 [162]	India	2000-2002	Gallbladder	Case-control	ST: detection of antibodies against *ST* Vi antigen in serum using ELISA. Cancer: *unk.*	446 patients with gastrointestinal, biliary or other related diseases including 27 GBC patients, 705 healthy controls.	Significantly higher antibody positivity in GBC patients (2/27, 7.4%) *vs.* healthy controls (13/705, 1.8%) ($p<0.05$).
Sharma, 2007* [157]	India	*unk*	Gallbladder	Case-control	ST: detection of antibodies against ViAb in serum and detection of bacteria in bile. Cancer: *unk.*	65 GBC patients, 125 CL patients, 200 healthy controls.	Significantly higher ViAb positivity in GBC patients (20/65, 30.8%) *vs.* controls (22/200, 11.0%) (OR: 3.60; $p<0.05$).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Capoor, 2008 [145]	India	2005-2007	Gallbladder	Case-control	ST: detection of *ST* in bile and feces samples using culturing methods. Cancer: diagnosis based on histopathological examination.	6 GBC patients, 53 patients with acute cholecystitis with concurrent CL, 45 patients with acute cholecystitis with gastrointestinal ailments requiring biliary drainage.	*ST* isolated from bile of 1/6 (16.7%) GBC patients and 2/53 (3.8%) patients with acute cholecystitis with concurrent CL.
Nath, 2008 [152]	India	2004-2007	Gallbladder	Case-control	ST: ViAb detection in serum using indirect hemagglutination assay. *ST* flagellin gene detection in bile using PCR. Cancer: diagnosis based on ultrasonography, aspiration cytology and/ or histopathology.	52 GBC patients, 223 with benign gallbladder diseases, 508 healthy controls, 424 corpses (without visible gallbladder pathology).	Significantly higher ViAb-positivity in GBC patients (20/52, 38.4%) *vs.* controls (47/508, 9.2%) (OR 6.13; p<0.001). Significantly higher prevalence of St flagellin positivity in GBC patients (35/52, 67.3%) *vs.* benign disease group (95/223, 42.6%) (p<0.01). No significant difference in *ST* isolation from gallbladder tissue of GBC patients (2/52, 3.8%) *vs.* benign disease group (2/223, 0.9%).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Tewari, 2010 [161]	India	2007-2009	Gallbladder	Case-control	ST: detection in tissue/ bile using PCR and culturing methods, and antibody detection in serum using the Widal test and IHA. Cancer: *unk.*	54 GBC patients, 54 controls with CL.	24/54 (44.4%) and 12/54 (22.2%) of GBC patients positivity on Widal and IHA test respectively *vs.* 13/54 (24.1%) and 5/54 (9.3%) of the controls. Positive *ST* PCR in 18/54 (33.3%) tissue samples and 2/54 (3.7%) bile samples of GBC patients *vs.* none of the controls.
Safaeian, 2011 [155]	China	1997-2001	Biliary tract, Gallbladder	Case-control	ST: detection of Vi antibody against *ST*. Cancer: diagnosis according to ICD-9 156 codes.	627 GBC/biliary tract cancers, 774 CL patients, 263 bile-duct stone patients, 959 healthy controls.	Low prevalence of chronic typhoid carriers. No significantly elevated seropositivity in cancer patients as compared to controls. Seropositivity in 1/457 (0.22%) cancer patients, 4/977 (0.41%) patients with stones and 1/859 (0.12%) healthy controls.
Scanu, 2015 [136]	India	2009-2013	Gallbladder	Case-control	ST: detection of antibodies against *ST* Vi antigen in serum. Cancer: *unk*	23 GBC patients, 60 controls with benign gallbladder disease.	Significant higher ViAb-positivity in GBC patients (12/23, 52.2%) *vs.* benign disease group (7/60, 11.7%) ($p<0.001$).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Iyer, 2016 [139]	India	*unk*	Gallbladder	Cross-sectional	ST: DNA detection in (tumor and normal) tissue using PCR and whole exosome sequencing. Cancer: histopathologically confirmed gallbladder cancer.	17 tumor tissues, 9 matched adjacent normal tissue.	None of gallbladder samples *S*T -positive based on PCR. Traces of *S*T found in 11/26 gallbladder isolates.
Tsuchiya, 2018 [85]	Bolivia, Chile	2014-2016	Gallbladder	Case-control	ST: detection of microbes in bile using sequencing of V3-V4 region. Cancer: *unk.*	7 GBC patients, 30 controls with CL.	*S*T t not detected in bile of GBC patients.

*Only abstract available. ST: *Salmonella* Typhi. SIR: standardized incidence ratio. *Unk:* unknown. GBC: gallbladder carcinoma. CL: cholelithiasis. CRC: colorectal cancer. ELISA: enzyme-linked immunosorbent assay. OR: odds ratio. ViAb: antibodies against *Salmonella* Typhi Vi antigen. ICD: International Classification of Diseases. PCR: polymerase chain reaction. IHA: indirect hemagglutination assay.

Supplementary Table 14. Characteristics and main outcomes of epidemiological studies assessing the association between ***Schistosoma*** **spp.** and cancer in the gastrointestinal tract

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Zhao, 1981 [175]	China	1970s	Colon	Cross-sectional	Schistosoma: *unk.* Cancer: *unk.*	*unk.*	337 CC patients with concurrent schistosomiasis.
Xu, 1984 [174]	China	1973-1979	Colorectal	Case-control	*S. japonicum*: *unk.* Cancer: diagnosis based on histopathology, radiology, endoscopy, surgery.	98 CC patients, 154 rectal cancer patients, 252 controls with non-GI cancer, 252 healthy controls.	Strong correlation between infection with *S. japonicum* and mortality from CRC. RRs of rectal cancer in patients with history of schistosomiasis 8.3 (95%CI 3.1-22.6) and 4.5 (95%CI 1.7-12.1) as compared to non-GI cancers and healthy controls respectively. RR for colon cancer not significantly increased (1.20; 95%CI 0.48-3.18).
Iida, 1999 [176]	Japan	1985-1996	Liver	Retrospective cohort	*S. japonicum*: history of anti-schistosomiasis treatment, positive skin test or ultrasonographic/ computer tomographic confirmation. Cancer: *unk.*	26 HCC patients, 484 chronic schistosomiasis patients (follow-up time of up to 10 years).	Development of HCC in 26/484 (5.4%) of schistosomiasis patients. Hepatitis C virus seropositivity in 39.5% of HCC patients with schistosomiasis. Out of 571 autopsies, 21/54 (39%) of patients with HCC had concurrent schistosomiasis, also 21/144 (14.6%) of patients with chronic schistosomiasis had HCC *vs.* 33/427 (7.7%) individuals without schistosomiasis (p=0.015).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Qiu, 2005 [173]	China	1995-2002	Colon, Liver	Case-control	*S. japonicum*: history of infection retrieved from medical records, surveillance data or by questionnaire. Cancer: *unk*	142 CC patients, 285 controls matched to CC patients, 127 liver cancer patients, 127 controls matched to liver cancer patients.	Significant increased risk of CC (OR 3.3; 95%CI 1.8-6.1) and liver cancer (OR 3.7; 95%CI 1.0-1.3) after schistosomiasis. Fraction attributable to schistosomiasis of 24% for CC and 27% for liver cancer.
Madbouly, 2007 [177]	Egypt	1999-2001	Colorectal	Case-control	*S. mansoni*: diagnosed by histopathologic examination of the colonic mucosa. Cancer: *unk.*	60 CC patients with schistosomal colitis (SCC), 40 CC patients without schistosomal colitis (NDCC).	More mucinous adenocarcinomas in SCC group compared to NDCC group (35% *vs.* 10%, p= 0.02) and more late stage tumors. Patients in SCC group significantly younger than NDCC controls (34.5 *vs.* 50.7 years; p=0.02)
Toda, 2015 [178]	Brazil	2002-2015	Liver	Case series	*S. mansoni*: epidemiological evidence, history of schistosomiasis treatment, fecal samples, imaging patterns and/or histological findings. Cancer: noninvasive diagnostic criteria of American Association for Study of Liver Diseases.	7 patients with history of *S. mansoni* infection and HCC.	All 7 HCC patients tested negative on serological test for hepatitis C virus-antibodies, 4 were positive on hepatitis B virus core antibodies .

Unk: unknown. CC: colon cancer. Non-GI: outside the gastrointestinal tract. CRC: colorectal cancer. RR: relative risk. HCC: hepatocellular carcinoma. OR: odds ratio. SCC: schistosomal colits. NDCC: without schistosomal colitis.

Supplementary Table 15. Characteristics and main outcomes of epidemiological studies assessing the association between ***Streptococcus*** **spp.** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Murray, 1978 [186]	USA	1970-1976	Colon	Cross-sectional	*S. bovis*: *unk.* Cancer: *unk.*	36 patients with *S. bovis* bacteremia, of which 10 with *S. bovis* IE.	2/36 (5.6%) bacteremia patients had a concurrent CC diagnosis and another 2/36 (5.6%) had potentially malignant villous adenomas.
Ruoff, 1989 [187]	USA	1982-1987	Colon	Cross-sectional	*S. bovis*: detection in blood samples using culturing methods. Cancer: *unk.*	38 patients with *S. bovis* bacteremia.	15/38 (39.5%) of *S. bovis* bacteremia patients were diagnosed with CC.
Zarkin, 1990 [188]	USA	1979-1988	Colon	Cross-sectional	*S. bovis*: detection in blood samples using culturing methods. Cancer: *unk.*	92 patients with *S. bovis* bacteremia, of which 26 with *S. bovis* IE.	43/92 (46.7%) bacteremia patients underwent colonic evaluation (e.g. colonoscopy, pathologic examination). Colonic lesions found in 11/24 (26.2%) *S. bovis* bacteremia patients and 11/19 (57.9%) *S. bovis* IE patients. CC diagnosis 6 *S. bovis* IE patients.
Gonzalez-Quintela, 2001 [184]	Spain	1993-2000	Colon	Cross-sectional	*S. bovis*: detection in blood samples using culturing methods. Cancer: *unk.*	20 patients with *S. bovis* bacteremia.	Colon evaluation data available for 13 *S. bovis* bacteremia patients, of which 3/13 (15.0%) were diagnosed with CC.
Jean, 2004* [185]	Taiwan	1992-2001	Colon	Cross-sectional	*S. bovis*: detection in blood samples using the API20 strep method. Cancer: diagnosis based on colonoscopy.	62 patients with *S. bovis* bacteremia.	19/62 (30.6%) of bacteremia patients underwent colonoscopy, of which 9/19 (47.3%) had colonic lesions (carcinoma or tubular adenoma).

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Corredoira, 2005 [183]	Spain	1987-2003	Colon	Prospective cohort	*Streptococcus*: detection in blood samples using culturing methods. Cancer: *unk.*	617 patients with *Streptococcus* bacteremia, of which 41 with *Sgg.*	More frequent detection of (pre)malignant CC in patients with a history of *Sgg* bacteremia (24/42, 57.1%) *vs. S. salivarius* bacteremia (0/52) (p=0.005).
Beck, 2008 [182]	Germany	*unk*	Colon	Retrospective cohort	*Streptococcus*: detection in blood samples using 16S rRNA gene sequencing. Cancer: *unk.*	58 patients with *Streptococcus* bacteremia of which 46 with available patient record.	Less than 10% of *Streptococcus* bacteremia patients were diagnosed with CC.
Corredoira, 2008* [190]	Spain	1988-2007	Colon, Other GI cancers	Prospective cohort	*Streptococcus:* detection in blood samples using API20 strep method or GP card of the VITEK2 system. Cancer: *unk.*	133 bacteremias (90 *Sgg*, 15 *S. gallolyticus* subsp. *pasteurianus*, 28 S. *infantarius*).	51/105 (48.5%) and 3/28 (10.7%) of *S. gallolyticus* and *S. infantarious* infections respectively developed/had CC. 6/105 (5.7%) and 16/28 (57.1%) of *S. gallolyticus* and *S. infantarious* infections respectively had noncolonic cancer, of which 12 were GI cancers, mainly pancreatic and biliary tract.
Abdulamir, 2009 [11]	Malaysia	2006-2007	Colorectal	Case-control	*S. gallolyticus*: detection of IgG antibodies in blood samples using ELISA. Cancer: diagnosis based on colonoscopy.	50 CRC patients, 14 colorectal adenoma patients, 30 apparently healthy volunteers and 30 controls without colon tumors.	IgG antibody titers significantly higher in CRC patients (0.158 ± 0.032) and adenoma patients (0.173 ± 0.024) *vs.* healthy volunteers (0.064 ± 0.011) and controls (0.046 ± 0.024) (p<0.05). No difference in antibody titers between CRC patients and adenoma patients.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Vaska, 2009 [192]	Australia	1999-2006	Colon	Retrospective cohort	*S. bovis*: detection in blood samples using culturing methods. Cancer: diagnosis based on colonoscopy and histopathology.	20 patients with *S. bovis* bacteremia, of which 10 caused by *Sgg* and 10 by *S. bovis* biotype II.	9/10 patients with a history of *Sgg* bacteremia underwent colonoscopy, of which 5/9 (55.6%) were diagnosed with CC and 4 had polyps or adenomas. 0/5 patients with a history *S. bovis* biotype II bacteremia who underwent colonoscopy developed CC.
Boleij, 2010 [189]	USA, The Netherlands	*unk*	Colorectal	Case-control	*S. bovis*: detection in blood samples using culturing methods, detection of antibodies against *S. bovis* antigen RpL7/L12 using ELISA. Cancer: *unk.*	Dutch population: 82 CRC patients, 10 patients with a systemic bacterial infection, 127 healthy controls. USA population: 64 CRC or polyp patients, 48 healthy controls.	Higher immune response against the RpL7/L12 antigen in stage I/II CRC patients and polyp patients compared to healthy controls (p=0.013) or late stage CRC patients (p=0.025).
Fernández-Ruiz, 2010 [191]	Spain	1997-2008	Colon	Cross-sectional	*S. bovis*: detection in blood samples using culturing methods. Cancer: diagnosis based on colonoscopy or pathologic examination of endoscopic/surgical samples.	59 patients with *S. bovis* bacteremia.	4/59 (6.8%) of the bacteremia patients was diagnosed with a GI malignancy before the bacteremia (1 of which with CC). 33/59 (55.9%) of the bacteremia patients underwent colonic evaluation, most of them ≤3 months after the bacteremia. 6/33 (18.2%) patients had CC.
Rahimkhani, 2010 [19]	Iran	*unk*	Colorectal	Case-control	*S. bovis*: detection in fecal samples using culturing methods. Cancer: *unk.*	30 CRC patients, 30 healthy controls.	*S. bovis* most abundant bacterial species in fecal samples of 9 CRC patients and 6 controls.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Boleij, 2012 [193]	USA, The Netherlands	*unk*	Colorectal	Case-control	*S. bovis*: detection in blood samples using culturing methods, detection of IgG antibodies against 4 pilus proteins using ELISA. Cancer: *unk.*	Dutch population: 37 CRC patients, 12 patients with polyps, 15 patients with clinical bacterial infection (*E. coli, Klebsiella pneumonia* or *Sgg*), 27 healthy controls. USA population: 33 CRC patients, 11 patients with polyps and 47 healthy controls.	The immune response to the 4 pilus proteins (Gallo1569, -2039, -2178, -2179) was specific for *Sgg* infection (compared to the other bacterial infections), however also high interindividual variation observed.
Garza-González, 2012 [194]	USA	*unk*	Colon	Case-control	*S. bovis*: detection of serological markers in blood samples using Western blot. Cancer: diagnosis based on colonoscopy.	133 patients with colonic adenomatous polyps, 53 healthy controls (with normal colonoscopy).	22 immunogenic proteins detected by the Western blot, of which two (of ~22 kDa and ~30 kDa) were most prominent. Presence of the 22 kDa protein associated with a higher risk of adenomatous polyps (OR 7.98; 95%CI 3.54-17.93, p<0.001). Higher OR when both proteins were present (OR 22.37; 95%CI 3.77-131.64, p<0.001).
Zammit, 2014 [195]	Malta	2007-2012	Colon, Liver, Biliary tract, Pancreas	Retrospective cohort	*S. bovis*: bacteremia confirmed by blood culture, catalase test or other test. Cancer: *unk.*	42 patients with a history of *S. bovis* bacteremia, 2 CRC patients, 6 colorectal adenoma patients, 11 patients with hepatobiliary–pancreatic pathologies.	19% of patients with history of *S. bovis* bacteremia had colonic adenomas or colonic adenocarcinomas. Choledocholithiasis or cholelithiasis with cholangitis or cholecystitis observed in 4/42 patients. 2/42 patients had pancreatic cancer.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Corredoira, 2015 [61]	Spain	1988-2014	Colorectal	Prospective cohort	*S. bovis*: detection using conventional phenotypic test and/or molecular test on blood samples. Cancer: colonoscopy.	109 patients with *S. bovis* IE.	Colorectal neoplasia found in 70/93 (75.3%) of patients who underwent colonoscopy during acute *S. bovis* IE (25 NAA, 37 AA, 8 CRC). 43 patients with previous *S. bovis* IE developed colorectal neoplasm (20 NAA, 18 AA, 5 CRC) during follow-up after a mean of 60.6 months (range 12-181 months).
Paritsky, 2015 [198]	Israel	2012-2013	Colorectal	Cross-sectional	*S.bovis*: detection in tissue using a VITEK2 system. Cancer: diagnosis based on colonoscopy.	203 patients who underwent colonoscopy (for different reasons).	49/203 (24.1%) of the individuals who underwent colonoscopy were tested positive on *S. bovis*. Of these 49 *S. bovis*-positive individuals, 17 (34.7%) had a malignant tumor, 22 (44.9%) had polyps, 4 (8.2%) had colitis and 6 (12.2%) showed no colonic abnormalities. No malignant tumors were found among the 154 individuals who tested *S. bovis*-negative.
Butt, 2016 [196]	Spain	2008-2013	Colorectal	Case-control	*Streptococcus*: multiplex serology against 4 *sgg* antigens (Gallo1569, 2039, 2178, 2179). Cancer: histologically confirmed colorectal cancer.	576 CRC patients, 576 healthy controls.	(Significant) increased risk of CRC associated with: - positivity to Gallo2039: OR 1.58 (95%CI 1.09-2.28); - positivity to Gallo2178: OR 1.58 (95%CI 1.09-2.30); - positivity to Gallo2179: OR 1.45 (95%CI 1.00-2.11); - Double positivity for Gallo2178 and Gallo 2179: OR 3.54 (95%CI 1.49-8.44); - Two or more positivity of Gallo1569, Gallo2178, Gallo2179: OR 1.93 (95%CI 1.04-3.56). Stronger association in people aged <65 years.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Butt, 2017 [197]	Germany	2005-2013, 2003-2007	Colorectal	Case-control	*Streptococcus*: detection of 11 *sgg* antigens in serum using multiplex serology. Cancer: newly diagnosed colon or rectum cancer patients (*nfs*).	50 newly diagnosed CRC patients (*BliTz study*), 30 NAA patients, 100 AA patients, 228 healthy controls. 318 additional CRC patients from *DACHSplus* study included.	(Significant) increased risk of CRC associated with: - positivity to Gallo2178: OR 4.13 (95%CI 2.11-8.08) (*DACHSplus)*; - Gallo2178: OR 3.19 (95%CI 1.11–9.21) (*BliTz*); - positivity to ≥2 of the 6 new SGG markers: OR 1.81 (95%CI 1.07–3.06) (*DACHSplus)*; - positivity to any of the 11 SGG antigens: OR 1.27 (95%CI 0.64-2.51) (*BliTz*); - positivity to ≥2 of the 6 new SGG markers: OR 1.50 (95%CI 0.61-3.72) (*BliTz study*).
Butt, 2018a [199]	Denmark, France, Greece, Germany, Italy, Netherlands, Norway, Spain, Sweden, United Kingdom	1992-2003	Colorectal	Case–control study nested within a prospective cohort	*Streptococcus*: detection of 11 *sgg* antigens in serum using multiplex serology. Cancer: ICD-10 C18-C20 diagnosis.	485 CRC patients (samples drawn 0.4-8.5 years before CRC diagnosis), 485 healthy controls who did not develop cancer.	Significant increased risk of CRC associated with: - positivity to any of the 11 SGG antigens: OR 1.36 (95%CI 1.04-1.77); - positivity to ≥2 of the 6 new SGG markers: OR 2.17 (95%CI 1.44-3.27).
Butt, 2018b [200]	US	*unk*	Colorectal	Prospective cohort	*Streptococcus*: detection of 9 *sgg* antigens in serum using multiplex serology. Cancer: ICD-O-3 C180-C189, C199, C209	Participants selected from 10 prospective cohorts. 4,063 CRC patients (samples drawn up to >10 years before CRC diagnosis), 4,063 healthy controls.	Increased risk of CRC with positivity to Gallo2178: OR 1.23 (95%CI 0.99-1.52), particularly when diagnosed <10 years after blood draw (OR 1.40; 95%CI 1.09-1.79). No association between Gallo2178-positivity and CRC risk in the subgroup diagnosed ≥10 years after blood draw.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Guven, 2018* [201]	Turkey	*unk*	Colorectal	Case-control	*Streptococcus*: detection in saliva samples using PCR. Cancer: *unk.*	71 CRC patients, 77 healthy controls (aged >50 years).	*Sgg* positivity similar in CRC patients (31%) *vs.* controls (27%). Significantly higher amount of *Sgg* in saliva of CRC patients *vs.* controls (4.12 ±0.99 and 3.15 ±0.58 log^{10} copies/ml respectively, p< 0.001). No difference in prevalence or quantity between different tumor locations (proximal *vs.* distal) or tumor stage (stage 1-3 *vs*. stage 4).
Kale, 2018* [202]	India	2013-2017	Rectum, Liver, Pancreas	Retrospective cohort study	*Streptococcus*: detection in blood, ascitic fluid, bile or pleural fluid. Cancer: *unk.*	68 patients with a *S. gallolyticus subsp pasteurianus* infection, of which 7 with a malignancy.	5/68 patients with *S. gallolyticus subsp pasteurianus* had concomitant HCC, 1/68 rectal adenocarcinoma, 1 pancreatic carcinoma.
Kwong, 2018 [17]	China	2006-2015	Colorectal	Retrospective cohort	*S. bovis*: culture-confirmed bacteremia. Cancer: *unk.*	203 *S. bovis* bacteremia patients, 1,0115 matched controls without history of bacteremia.	8/203 (3.9%) of the bacteremia patients developed CRC compared to 10/1,015 (1.0%) of the controls, aHR: 5.73 (95%CI 2.18-15.1; p=0.0004).
Bundgaard, 2019 [13]	Denmark	2002-2010	Colorectal	Prospective cohort	*S. bovis*: DNA detection in (tumor and adjacent normal) tissue samples using qPCR. Cancer: *unk.*	99 CRC patients, 96 adenoma patients, 104 patients with diverticular disease.	*S. bovis* not found in any of the samples of all study participants.
Justesen, 2020* [9]	Denmark	2007-2018	Colorectal	Retrospective cohort	*S. bovis*: culture confirmed bacteremia (in blood samples). Cancer: *unk.*	117 patients with bacteremia caused by *S. bovis*. Reference population: ~2 million.	6/117 (5.1%) of the individuals with a history of bacteremia caused by *S. bovis* were diagnosed with CRC, 5 of them within 1 year after bacteremia.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Sheikh, 2020 [203]	Iran	2017-2018	Colorectal	Case-control	*Streptococcus*: detection in fecal samples using culturing and PCR. Cancer: clinical, radiological, histological criteria and endoscopy and colonoscopy.	22 CRC patients, 44 IBD patients, 40 healthy controls.	Significantly higher *Sgg* positivity in CRC and IBD patients *vs.* controls based on culturing (p=0.013) and PCR (p=0.001). Culturing: *Sgg* positivity in 2/22 (9.1%) CRC patients and 7/44 (15.9%) IBD patients *vs.* 0/40 of the controls. PCR: *Sgg* positivity in 9/22 (40.9%) CRC patients, 15/44 (34.1%) IBD patients *vs.* 3/40 (7.5%) healthy controls.
Wang, 2020 [31]	China		Colorectal	Case-control	*Streptococcus* spp.: detection in (tumor) and mucosa using 16S rRNA gene sequencing. Cancer: *unk.*	Tumor, adjacent normal and off-tumor site tissue from 75 CRC patients and mucosa from 26 healthy controls.	Relative abundance of *Streptococcus* spp. significantly higher in tumor tissue *vs.* mucosa from healthy controls.

*Only abstract available. *Unk*: unknown. IE: infective endocarditis. CC: colon cancer. *Sgg*: *Streptococcus gallolyticus* subsp. *gallolyticus.* GI: gastrointestinal. NAA: non-advanced adenoma. AA: advanced adenoma. CRC: colorectal cancer. OR: odds ratio. *Nfs:* not further specified. ICD(-O): International Classification of Diseases (for Oncology). PCR: polymerase chain reaction. HCC: hepatocellular carcinoma. aHR: adjusted hazard ratio. IBD: inflammatory bowel disease.

Supplementary Table 16. Characteristics and main outcomes of epidemiological studies assessing the association between ***Strongyloides stercoralis*** and cancer in the gastrointestinal tract.

First author, year	Country	Study period	Malignancy/ malignancies	Study type	Diagnostic method(s)	Population size or no. cases & controls	Main outcomes
Hirata, 2007 [211]	Japan	1991-2005	Liver, Biliary tract, Pancreas	Case-control	*Strongyloides*: detection in fecal samples using culturing methods. Cancer: diagnosis based on histology, cytology, radiological findings.	196 (liver, pancreatic, biliary tract) cancer patients, 1,458 controls without cancer.	Prevalence of *S. stercoralis* significantly higher in biliary tract cancer patients (18.4%) *vs.* controls (7.5%). OR for developing cancer in pancreas 2.2 (95%CI 0.7-6.6), liver 0.9 (95%CI 0.5-1.8), biliary tract 2.7 (95%CI 1.1-6.3).
Tanaka, 2016 [210]	Japan	1991-2014	Esophagus, Stomach, Colorectal, Liver, Biliary tract, Pancreas,	Retrospective cohort	*Strongyloides*: detection in fecal samples using culturing methods. Cancer: diagnosis based on histology, cytology, radiological findings.	1,352 cancer patients (different malignancies including 7 esophagus, 24 stomach, 15 colorectal, 2 pancreas, 9 liver, 10 biliary tract), 2,596 controls without cancer.	OR for developing cancer in esophagus 0.65 (95%CI 0.29-1.45), stomach 1.22 (95%CI 0.76-1.97), colorectal 0.94 (95%CI 0.53-1.66), pancreas 0.83 (95%CI 0.19-3.55), liver 0.72 (95%CI 0.35-1.47), biliary tract 1.90 (95%CI 0.93-3.87).

OR: odds ratio.

References

7. Haghi, F., Goli, E., Mirzaei, B., and Zeighami, H. (2019). The association between fecal enterotoxigenic *B. fragilis* with colorectal cancer. *BMC Cancer*, 19(1), 1-4.
8. Zamani, S., Taslimi, R., Sarabi, A., Jasemi, S., Sechi, L. A., Feizabadi, M. M. (2020). Enterotoxigenic *Bacteroides fragilis*: a possible etiological candidate for bacterially-induced colorectal precancerous and cancerous lesions. *Front. Cell. Inf. Microbiol.* 9, 449.
9. Justesen, U. (2020). Bacteraemia with anaerobic bacteria and association with colorectal cancer, in ECCMID-2020. [conference abstract].
10. Ma, Y., Zhang, Y., Xiang, J., Xiang, S., Zhao, Y., Xiao, M., et al. (2021). Metagenome analysis of intestinal bacteria in healthy people, patients with inflammatory bowel disease and colorectal cancer. *Front. Cell. Inf. Microbiol*., 11.
11. Abdulamir, A.S., Hafidh, R. R., Mahdi, L. K., Al-jeboori, T., and Abubaker, F. (2009). Investigation into the controversial association of *Streptococcus gallolyticus* with colorectal cancer and adenoma. *BMC Cancer*, 9(1), 1-12.
12. Boleij, A., Hechenbleikner, E. M., Goodwin, A. C., Badani, R., Stein, E. M., Lazarev, M. G., et al. (2015). The *Bacteroides fragilis* toxin gene is prevalent in the colon mucosa of colorectal cancer patients. *Clin. Infect. Dis.,* 60(2), 208-215.
13. Bundgaard-Nielsen, C., Baandrup, U. T., Nielsen, L. P., and Sørensen, S. (2019). The presence of bacteria varies between colorectal adenocarcinomas, precursor lesions and non-malignant tissue. *BMC Cancer*, 19(1), 1-13.
14. Fukugaiti, M. H., Ignacio, A., Fernandes, M. R., Ribeiro Júnior, U., Nakano, V., and Avila-Campos, M. J. (2015). High occurrence of *Fusobacterium nucleatum* and *Clostridium difficile* in the intestinal microbiota of colorectal carcinoma patients. *Braz. J. Microbiol*., 46(4), 1135-1140.
15. Hale, V. L., Jeraldo, P., Chen, J., Mundy, M., Yao, J., Priya, S., et al. (2018). Distinct microbes, metabolites, and ecologies define the microbiome in deficient and proficient mismatch repair colorectal cancers. *Genome Med*., 10(1), 1-13.
16. Keenan, J., Aitchison, A., Greenlees, R., and Frizelle, F. (2014). Are toxin-producing strains of *Bacteroides fragilis* involved in colorectal cancer?:CR27. *Anz. J. Surg*. 84.
17. Kwong, T. N., Wang, X., Nakatsu, G., Chow, T. C., Tipoe, T., Dai, R. Z., et al. (2018). Association between bacteremia from specific microbes and subsequent diagnosis of colorectal cancer. *Gastroenterology*, 155(2), 383-390.
18. Purcell, R. V., Pearson, J., Aitchison, A., Dixon, L., Frizelle, F. A., and Keenan, J. I. (2017). Colonization with enterotoxigenic *Bacteroides fragilis* is associated with early-stage colorectal neoplasia. *PloS One*, 12(2), e0171602.
19. Rahimkhani, M., Mohagheghi, M. A., and Yaraei, K. (2010). Fecal microbial flora in colorectal cancer. *Afr. J. Microbiol. Res*., 4(23), 2622-2625.
20. Ulger Toprak, N., Yagci, A. Y. Ş. E. G. Ü. L., Gulluoglu, B. M., Akin, M. L., Demirkalem, P., Celenk, T., et al. (2006). A possible role of *Bacteroides fragilis* enterotoxin in the aetiology of colorectal cancer. *Clin. Microbiol. Infect*., 12(8), 782-786.
21. Viljoen, K. S., Dakshinamurthy, A., Goldberg, P., and Blackburn, J. M. (2015). Quantitative profiling of colorectal cancer-associated bacteria reveals associations between fusobacterium spp., enterotoxigenic *Bacteroides fragilis* (ETBF) and clinicopathological features of colorectal cancer. *PloS One*, 10(3), e0119462.

22. Xie, L. L., Wu, N., Zhu, Y. M., Qiu, X. Y., Chen, G. D., Zhang, L. M., et al. (2016). Expression of enterotoxigenic *Bacteroides fragilis* and polyketide synthase gene-expressing *Escherichia coli* in colorectal adenoma patients. *Zhonghua yi xue za zhi*, 96(12), 954-959.

25. Alexander, J. L., Scott, A., Mroz, A., Perdones-Montero, A., Mckenzie, J., Rees, D. N., et al. (2016). 91 Mass spectrometry imaging (MSI) of microbiome-metabolome interactions in colorectal cancer. *Gastroenterology*, 150(4), S23.

26. Allali, I., Delgado, S., Marron, P. I., Astudillo, A., Yeh, J. J., Ghazal, H., et al. (2015). Gut microbiome compositional and functional differences between tumor and non-tumor adjacent tissues from cohorts from the US and Spain. *Gut Microbes*, 6(3), 161-172.

27. Blackett, K. L., Siddhi, S. S., Cleary, S., Steed, H., Miller, M. H., Macfarlane, S., et al. (2013). Oesophageal bacterial biofilm changes in gastro-oesophageal reflux disease, Barrett's and oesophageal carcinoma: association or causality?. *Aliment. Pharmacol. Ther*., 37(11), 1084-1092.

28. Brauner, A., Brandt, L., Frisan, T., Thelestam, M., and Ekbom, A. (2010). Is there a risk of cancer development after *Campylobacter* infection?. *Scand. J. Gastroenterol*., 45(7-8), 893-897.

29. De Savornin Lohman, E., Duijster, J., Groot Koerkamp, B., van der Post, R., Franz, E., Mughini Gras, L., et al. (2020). Severe *Salmonella* spp. Or *Campylobacter* spp. infection and the risk of biliary tract cancer: A population-based study. *Cancers*, 12(11), 3348.

30. Mughini Gras, L. (2017). Colon cancer risk in patients with a history of *Campylobacter* infection, in 19th International workshop CHRO 2017 on *Campylobacter*, *Helicobacter* and Related Organisms. 2017: Nantes, France. p. 250. [conference abstract].Alexander, J.L., et al., 91 Mass spectrometry imaging (MSI) of microbiome-metabolome interactions in colorectal cancer. Gastroenterology, 2016. 150(4): p. S23.

31. Wang, Y., Zhang, C., Hou, S., Wu, X., Liu, J., and Wan, X. (2020). Analyses of potential driver and passenger bacteria in human colorectal cancer. *Cancer Manag. Res*., 12, 11553.

32. Wei, A. L., Li, M., Li, G. Q., Wang, X., Hu, W. M., Li, Z. L., et al. (2020). Oral microbiome and pancreatic cancer. *World J. Gastroenterol*., 26(48), 7679.

33. Wu, N., Yang, X., Zhang, R., Li, J., Xiao, X., Hu, Y., et al. (2013). Dysbiosis signature of fecal microbiota in colorectal cancer patients. *Microb. Ecol*., 66(2), 462-470.

40. Ahn, J., Sinha, R., Pei, Z., Dominianni, C., Wu, J., Shi, J., ... & Yang, L. (2013). Human gut microbiome and risk for colorectal cancer. *J. Nat. Cancer Inst*., 105(24), 1907-1911.

41. Liang, Q., Chiu, J., Chen, Y., Huang, Y., Higashimori, A., Fang, J., et al. Fecal bacteria act as novel biomarkers for noninvasive diagnosis of colorectal cancer. *Clin. Cancer Res*., 23(8), 2061-2070.

42. Lu, Y., Zhang, B. Y., Shi, J. S., and Wu, L. Q. (2004). Expression of the bacterial gene in gallbladder carcinoma tissue and bile. *Hepatobiliary & Pancreatic Diseases International: HBPD INT*, 3(1), 133-135.

43. Ohigashi, S., Sudo, K., Kobayashi, D., Takahashi, O., Takahashi, T., Asahara, T., et al. (2013). Occurence of initiation and progression of colorectal cancer in the changed intestinal environment. *J. Clin. Oncol.,* 31(4), 416.

44. Hsieh, Y. Y., Tung, S. Y., Pan, H. Y., Yen, C. W., Xu, H. W., Lin, Y. J., et al. (2018). Increased abundance of *Clostridium* and *Fusobacterium* in gastric microbiota of patients with gastric cancer in Taiwan. *Scientific Rep*., 8(1), 1-11.

45. ahani-Sherafat, S., Azimirad, M., Alebouyeh, M., Amoli, H. A., Hosseini, P., Ghasemian-Safaei, H., et al. (2019). The rate and importance of *Clostridium difficile* in colorectal cancer patients. *Gastroenterol. Hepatol. Bed Bench*, 12(4), 358-363.

46. Liang, J. Q., Wong, S. H., Szeto, C. H., Chu, E. S., Lau, H. C., Chen, Y., et al. (2021). Fecal microbial DNA markers serve for screening colorectal neoplasm in asymptomatic subjects. *J. Gastroenterol. Hepatol*., 36(4), 1035-1043.

48. Kalantari, N., Gorgani-Firouzjaee, T., Ghaffari, S., Bayani, M., Ghaffari, T., and Chehrazi, M. (2020). Association between *Cryptosporidium* infection and cancer: a systematic review and meta-analysis. *Parasitol. Int*., 74, 101979.

51. Sulżyc-Bielicka, V., Kołodziejczyk, L., Jaczewska, S., Bielicki, D., Safranow, K., Bielicki, P., et al. (2018). Colorectal cancer and *Cryptosporidium* spp. infection. *PLoS One*, 13(4), e0195834.

52. Sulżyc-Bielicka, V., Kuźna-Grygiel, W., Kołodziejczyk, L., Bielicki, D., Kładny, J., Stępień-Korzonek, M., et al. (2007). Cryptosporidiosis in patients with colorectal cancer. *J. Parasitol*., 93(3), 722-724.

53. Essid, R., Menotti, J., Hanen, C., Aoun, K., and Bouratbine, A. (2018). Genetic diversity of *Cryptosporidium* isolates from human populations in an urban area of Northern Tunisia. *Infect. Genet. Evol*., 58, 237-242.

54. Osman, M., Benamrouz, S., Guyot, K., Baydoun, M., Frealle, E., Chabe, M., et al. (2017). High association of *Cryptosporidium* spp. infection with colon adenocarcinoma in Lebanese patients. *PloS One*, 12(12), e0189422.

55. Sanad, M. M., Thagfan, F. A., Al Olayan, E. M., Almogren, A., Al Hammaad, A., Al-Mawash, A., et al. (2014). Opportunistic coccidian parasites among Saudi cancer patients presenting with diarrhea: prevalence and immune status. *Res. J. Parasitol*., 9(2), 55-63.

56. Sulżyc-Bielicka, V., Kołodziejczyk, L., Jaczewska, S., Bielicki, D., Kładny, J., and Safranow, K. (2012). Prevalence of *Cryptosporidium* sp. in patients with colorectal cancer. *Pol. J. Surg*., 2012. 84(7): p. 348-351.

58. De Almeida, C. V., Lulli, M., di Pilato, V., Schiavone, N., Russo, E., Nannini, G., et al. (2019). Differential responses of colorectal cancer cell lines to *Enterococcus faecalis*' strains isolated from healthy donors and colorectal cancer patients. *J. Clin. Med*., 8(3), 388.

60. Balamurugan, R., Rajendiran, E., George, S., Samuel, G. V., and Ramakrishna, B. S. (2008). Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, *Desulfovibrio* and *Enterococcus faecalis* in the feces of patients with colorectal cancer. *J. Gastroenterol. Hepatol*., 23(8pt1), 1298-1303.

61. Corredoira, J., García-País, M. J., Coira, A., Rabuñal, R., García-Garrote, F., Pita, J., et al. (2015). Differences between endocarditis caused by *Streptococcus bovis* and *Enterococcus* spp. and their association with colorectal cancer. *Eur. J. Clin. Microbiol. Infect. Dis*., 34(8), 1657-1665.

62. Geravand, M., Fallah, P., Yaghoobi, M. H., Soleimanifar, F., Farid, M., Zinatizadeh, N., et al. (2019). Investigation of Enterococcus faecalis population in patients with polyp and colorectal cancer in comparison of healthy individuals. *Arq. Gastroenterol*., 56, 141-145.

63. D'asheesh, T. I. A., Hussen, B. M., Al-Marzoqi, A. H., and Ghasemian, A. (2021). Assessment of oncogenic role of intestinal microbiota in colorectal cancer patients. *J. Gastrointest. Cancer*, 52(3), 1016-1021.

64. Pericas, J. M., Corredoira, J., Moreno, A., García-País, M. J., Falces, C., Rabuñal, R., et al. (2017). Relationship between *Enterococcus faecalis* infective endocarditis and colorectal neoplasm: preliminary results from a cohort of 154 patients. *Rev. Esp. Cardiol*. (English Edition), 70(6), 451-458.

65. Rezasoltani, S., Aghdaei, H. A., Dabiri, H., Sepahi, A. A., Modarressi, M. H., and Mojarad, E. N. (2018). The association between fecal microbiota and different types of colorectal polyp as precursors of colorectal cancer. *Microb. Pathog.*, 124, 244-249.

66. Shoji, M., Sasaki, Y., Abe, Y., Nishise, S., Yaoita, T., Yagi, M., et al. (2021). Characteristics of the gut microbiome profile in obese patients with colorectal cancer. *JGH Open*, 5(4), 498-507.

67. Wang, T., Cai, G., Qiu, Y., Fei, N., Zhang, M., Pang, X., et al. (2012). Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *ISME J*., 6(2), 320-329.

68. Zhou, Y., He, H., Xu, H., Li, Y., Li, Z., Du, Y., et al. (2016). Association of oncogenic bacteria with colorectal cancer in South China. *Oncotarget*, 7(49), 80794.

73. Tang, L., Zhou, Y. J., Zhu, S., Liang, G. D., Zhuang, H., Zhao, M. F., et al. (2020). *E. coli* diversity: low in colorectal cancer. *BMC Med. Genom*., 13(1), 1-17.

75. Bonnet, M., Buc, E., Sauvanet, P., Darcha, C., Dubois, D., Pereira, B., et al. (2014). Colonization of the human gut by *E. coli* and colorectal cancer risk. *Clin. Cancer Res*., 20(4), 859-867.

76. Buc, E., Dubois, D., Sauvanet, P., Raisch, J., Delmas, J., Darfeuille-Michaud, A., et al. (2013). High prevalence of mucosa-associated *E. coli* producing cyclomodulin and genotoxin in colon cancer. *PloS One*, 8(2), e56964.

77. Iyadorai, T., Mariappan, V., Vellasamy, K. M., Wanyiri, J. W., Roslani, A. C., Lee, G. K., et al. (2020). Prevalence and association of pks+ *Escherichia coli* with colorectal cancer in patients at the University Malaya Medical Centre, Malaysia. *PloS One*, 15(1), e0228217.

78. Kohoutova, D., Smajs, D., Moravkova, P., Cyrany, J., Moravkova, M., Forstlova, M., et al. (2014). *Escherichia coli* strains of phylogenetic group B2 and D and bacteriocin production are associated with advanced colorectal neoplasia. *BMC Infect. Dis*., 14(1), 1-8.

79. Pleguezuelos-Manzano, C., Puschhof, J., Huber, A. R., van Hoeck, A., Wood, H. M., Nomburg, J., et al. (2020). Mutational signature in colorectal cancer caused by genotoxic pks+ *E. coli*. *Nature*, 580(7802), 269-273.

80. Rezasoltani, S., Asadzadeh, A. H., Dabiri, H., Akhavan, S. A., Modarressi, M. H., and Nazemalhosseini, M. E. (2020). Investigating the number of *Lactobaccilus*, *Escherichia coli* and *Prevotella* in fecal microbiota of adenomatous polyposis and colorectal cancer patients. *Tehran Univ. Med. J*., 78(3), 137-143.

81. Shimpoh, T., Hirata, Y., Ihara, S., Suzuki, N., Kinoshita, H., Hayakawa, Y., et al. (2017). Prevalence of pks-positive *Escherichia coli* in Japanese patients with or without colorectal cancer. *Gut Pathog.*, 9(1), 1-8.

82. Tunsjø, H. S., Gundersen, G., Rangnes, F., Noone, J. C., Endres, A., and Bemanian, V. (2019). Detection of Fusobacterium nucleatum in stool and colonic tissues from Norwegian colorectal cancer patients. *Eur. J. Clin. Microbiol. Infect. Dis*., 38(7), 1367-1376.

83. Yoshikawa, Y., Tsunematsu, Y., Matsuzaki, N., Hirayama, Y., Higashiguchi, F., Sato, M., et al. (2020). Characterization of colibactin-producing *Escherichia coli* isolated from Japanese patients with colorectal cancer. *Jpn. J. Infect. Dis*., 73(6), 437-442.

84. Zarei, O., Arabestan, M. R., Majlesi, A., Mohammadi, Y., and Alikhani, M. Y. (2019). Determination of virulence determinants of *Escherichia coli* strains isolated from patients with colorectal cancer compared to the healthy subjects. *Gastroenterol. Hepatol. Bed Bench*, 12(1), 52.

85. Tsuchiya, Y., Loza, E., Villa-Gomez, G., Trujillo, C. C., Baez, S., Asai, T., et al. (2018). Metagenomics of microbial communities in gallbladder bile from patients with gallbladder cancer or cholelithiasis. *Asian Pac. J. Cancer Prev*., 19(4), 961.

89. Castellarin, M., Warren, R. L., Freeman, J. D., Dreolini, L., Krzywinski, M., Strauss, J., et al. (2012). *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res*., 22(2), 299-306.

90. Flanagan, L., Schmid, J., Ebert, M., Soucek, P., Kunicka, T., Liska, V., et al. (2014). *Fusobacterium nucleatum* associates with stages of colorectal neoplasia development, colorectal cancer and disease outcome. *Eur. J. Clin. Microbiol. Infect. Dis*., 33(8), 1381-1390.

91. Ito, M., Kanno, S., Nosho, K., Sukawa, Y., Mitsuhashi, K., Kurihara, H., et al. (2015). Association of *Fusobacterium nucleatum* with clinical and molecular features in colorectal serrated pathway. *Int. J. Cancer*, 137(6), 1258-1268.

92. Mira-Pascual, L., Cabrera-Rubio, R., Ocon, S., Costales, P., Parra, A., Suarez, A., et al. (2015). Microbial mucosal colonic shifts associated with the development of colorectal cancer reveal the presence of different bacterial and archaeal biomarkers. *J. Gastroenterol*., 50(2), 167-179.

93. Amitay, E. L., Werner, S., Vital, M., Pieper, D. H., Höfler, D., Gierse, I. J., et al. (2017). *Fusobacterium* and colorectal cancer: causal factor or passenger? Results from a large colorectal cancer screening study. *Carcinogenesis*, 38(8), 781-788

94. Drewes, J. L., White, J. R., Dejea, C. M., Fathi, P., Iyadorai, T., Vadivelu, J., et al. (2017). High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. *NPJ Biofilms Microbiomes*, 3(1), 1-12.

95. Eklöf, V., Löfgren-Burström, A., Zingmark, C., Edin, S., Larsson, P., Karling, P., et al. (2017). Cancer-associated fecal microbial markers in colorectal cancer detection. *Int. J. Cancer*, 141(12), 2528-2536.

96. Wang, H. F., Li, L. F., Guo, S. H., Zeng, Q. Y., Ning, F., Liu, W. L., et al. (2016). Evaluation of antibody level against *Fusobacterium nucleatum* in the serological diagnosis of colorectal cancer. *Scientific Rep*., 6(1), 1-10.

97. Xie, Y.H. and Fang, J.Y. (2016). Single fecal abundance of *Fusobacterium nucleatum* for early- and late-stage detection of colorectal cancer. *J. Dig. Dis*., 2016. 17: p. 3.

98. Yu, J., Chen, Y., Fu, X., Zhou, X., Peng, Y., Shi, L., et al. (2016). Invasive *Fusobacterium nucleatum* may play a role in the carcinogenesis of proximal colon cancer through the serrated neoplasia pathway. *Int. J. Cancer*, 139(6), 1318-1326.

99. Scott, A., Alexander, J. L., Marchesi, J., and Kinross, J. M. (2017). Modification of the appendix microbiome in acute appendicitis and colon adenocarcinoma. *Gastroenterol.*, 152(5), S1228.

100. Suehiro, Y., Sakai, K., Nishioka, M., Hashimoto, S., Takami, T., Higaki, S., et al. (2017). Highly sensitive stool DNA testing of *Fusobacterium nucleatum* as a marker for detection of colorectal tumours in a Japanese population. *Ann. Clin. Biochem*., 54(1), 86-91.

101. Yamamura, K., Baba, Y., Miyake, K., Nakamura, K., Shigaki, H., Mima, K., et al. (2017). *Fusobacterium nucleatum* in gastroenterological cancer: Evaluation of measurement methods using quantitative polymerase chain reaction and a literature review. *Oncol. Lett*., 14(6), 6373-6378.

102. Yoon, H., Kim, N., Park, J. H., Kim, Y. S., Lee, J., Kim, H. W., et al. (2017). Comparisons of gut microbiota among healthy control, patients with conventional adenoma, sessile serrated adenoma, and colorectal cancer. *J. Cancer Prev*., 22(2), 108.

103. Yu, J., Feng, Q., Wong, S. H., Zhang, D., Liang, Q., Qin, Y., et al. (2017). Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut*, 66(1), 70-78.

104. Butt, J., Jenab, M., Pawlita, M., Overvad, K., Tjonneland, A., Olsen, A., et al. (2019). Antibody responses to *Fusobacterium nucleatum* proteins in prediagnostic blood samples are not associated with risk of developing colorectal cancer. *Cancer Epidemiol. Prev. Biomark*., 28(9), 1552-1555.

105. Repass, J. (2018). Replication Study: *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Elife*, 7, e25801.

106. Russo, E., Bacci, G., Chiellini, C., Fagorzi, C., Niccolai, E., Taddei, A., et al. (2018). Preliminary comparison of oral and intestinal human microbiota in patients with colorectal cancer: a pilot study. *Front. Microbiol*., 8, 2699.

107. Chen, X. H., Wang, A., Chu, A. N., Gong, Y. H., and Yuan, Y. (2019). Mucosa-associated microbiota in gastric cancer tissues compared with non-cancer tissues. *Front. Microbiol*., 10, 1261.

108. De Carvalho, A. C., de Mattos Pereira, L., Datorre, J. G., Dos Santos, W., Berardinelli, G. N., Matsushita, M. D. M., et al. (2019). Microbiota profile and impact of *Fusobacterium nucleatum* in colorectal cancer patients of Barretos Cancer Hospital. *Front. Oncol*., 9, 813.

109. Kageyama, S., Takeshita, T., Takeuchi, K., Asakawa, M., Matsumi, R., Furuta, M., et al. (2019). Characteristics of the salivary microbiota in patients with various digestive tract cancers. *Front. Microbiol*., 10, 1780.

110. Saito, K., Koido, S., Odamaki, T., Kajihara, M., Kato, K., Horiuchi, S., et al. (2019). Metagenomic analyses of the gut microbiota associated with colorectal adenoma. *PLoS One*, 14(2), e0212406.

111. Yachida, S., Mizutani, S., Shiroma, H., Shiba, S., Nakajima, T., Sakamoto, T., et al. (2019). Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat. Med*., 25(6), 968-976.

112. Alkharaan, H., Lu, L., Gabarrini, G., Halimi, A., Ateeb, Z., Sobkowiak, M. J., et al. (2020). Circulating and salivary antibodies to *Fusobacterium nucleatum* are associated with cystic pancreatic neoplasm malignancy. *Front. Immunol*., 11, 2003.

113. Boehm, E. T., Thon, C., Kupcinskas, J., Steponaitiene, R., Skieceviciene, J., Canbay, A., et al. (2020). *Fusobacterium nucleatum* is associated with worse prognosis in Lauren's diffuse type gastric cancer patients. *Sci. Rep*., 10(1), 1-12.

114. Gantuya, B., El Serag, H. B., Matsumoto, T., Ajami, N. J., Uchida, T., Oyuntsetseg, K., et al. (2020). Gastric mucosal microbiota in a Mongolian population with gastric cancer and precursor conditions. *Aliment. Pharmacol. Ther*., 51(8), 770-780.

115. Kashani, N., Bezmin Abadi, A. T., Rahimi, F., and Forootan, M. (2020). FadA-positive *Fusobacterium nucleatum* is prevalent in biopsy specimens of Iranian patients with colorectal cancer. *New Microbes New Infect*., 2020. 34.

116. Reynolds, I. S., Thomas, V., O'Connell, E., Fichtner, M., McNamara, D. A., Kay, E. W., et al. (2020). Mucinous adenocarcinoma of the rectum: a whole genome sequencing study. *Front. Oncol*., 10, 1682.

117. Zhang, C., Thakkar, P. V., Sharma, P., Vennelaganti, S., Betel, D., and Shah, M. A. (2019). Abstract 2826: Understanding associations among local microbiome, immune response, and efficacy of immunotherapy in esophageal cancer. *Cancer Res.,* 79(13).

118. Kawasaki, M., Ikeda, Y., Ikeda, E., Takahashi, M., Tanaka, D., Nakajima, Y., et al. (2021). Oral infectious bacteria in dental plaque and saliva as risk factors in patients with esophageal cancer. *Cancer*, 127(4), 512-519.

119. Kurt, M., and Yumuk, Z. (2021). Diagnostic accuracy of *Fusobacterium nucleatum* IgA and IgG ELISA test in colorectal cancer. *Sci. Rep*., 11(1), 1-6.

120. Pignatelli, P., Iezzi, L., Pennese, M., Raimondi, P., Cichella, A., Bondi, D., et al. (2021). The potential of colonic tumor tissue *Fusobacterium nucleatum* to predict staging and its interplay with oral abundance in colon cancer patients. *Cancers*, 13(5), 1032.

121. Eisele, Y., Mallea, P. M., Gigic, B., Stephens, W. Z., Warby, C. A., Buhrke, K., et al. (2021). *Fusobacterium nucleatum* and clinicopathologic features of colorectal cancer: results from the ColoCare study. *Clin. Colorectal Cancer*, 20(3), e165-e172.

122. Gethings-Behncke, C., Coleman, H. G., Jordao, H. W., Longley, D. B., Crawford, N., Murray, L. J., et al. (2020). *Fusobacterium nucleatum* in the colorectum and its association with cancer risk and survival: a systematic review and meta-analysis. *Cancer Epidemiol. Prev. Biomark*., 29(3), 539-548.

123. Janati, A. I., Karp, I., Laprise, C., Sabri, H., and Emami, E. (2020). Detection of *Fusobaterium nucleatum* in feces and colorectal mucosa as a risk factor for colorectal cancer: a systematic review and meta-analysis. *Syst. Rev*., 9(1), 1-15.

128. Ahn, J., Segers, S., and Hayes, R. B. (2012). Periodontal disease, *Porphyromonas gingivalis* serum antibody levels and orodigestive cancer mortality. *Carcinogenesis*, 33(5), 1055-1058.

129. Fan, X., Alekseyenko, A. V., Wu, J., Peters, B. A., Jacobs, E. J., Gapstur, S. M., et al. (2018). Human oral microbiome and prospective risk for pancreatic cancer: a population-based nested case-control study. *Gut*, 67(1), 120-127.

130. Gao, S., Li, S., Ma, Z., Liang, S., Shan, T., Zhang, M., et al. (2016). Presence of *Porphyromonas gingivalis* in esophagus and its association with the clinicopathological characteristics and survival in patients with esophageal cancer. *Infect. Agents Cancer*, 11(1), 1-9.

131. Gao, S. G., Yang, J. Q., Ma, Z. K., Yuan, X., Zhao, C., Wang, G. C., et al. (2018). Preoperative serum immunoglobulin G and A antibodies to *Porphyromonas gingivalis* are potential serum biomarkers for the diagnosis and prognosis of esophageal squamous cell carcinoma. *BMC Cancer*, 18(1), 1-8.

132. Michaud, D. S., Izard, J., Wilhelm-Benartzi, C. S., You, D. H., Grote, V. A., Tjønneland, A., et al. (2013). Plasma antibodies to oral bacteria and risk of pancreatic cancer in a large European prospective cohort study. *Gut*, 62(12), 1764-1770.

133. Peters, B. A., Wu, J., Pei, Z., Yang, L., Purdue, M. P., Freedman, N. D., et al. (2017). Oral microbiome composition reflects prospective risk for esophageal cancers. *Cancer Res*., 77(23), 6777-6787.

134. Wang, X., Jia, Y., Wen, L., Mu, W., Wu, X., Liu, T., et al. (2021). *Porphyromonas gingivalis* promotes colorectal carcinoma by activating the hematopoietic NLRP3 inflammasome. *Cancer Res*., 81(10), 2745-2759.

135. Sinha, R., Ahn, J., Sampson, J. N., Shi, J., Yu, G., Xiong, X., et al. (2016). Fecal microbiota, fecal metabolome, and colorectal cancer interrelations. *PloS One*, 11(3), e0152126.

136. Scanu, T., Spaapen, R. M., Bakker, J. M., Pratap, C. B., Wu, L. E., Hofland, I., et al. (2015). *Salmonella* manipulation of host signaling pathways provokes cellular transformation associated with gallbladder carcinoma. Cell Host Microbe, 17(6), 763-774.

137. Chang, R., Wei, J. C. C., Lin, M. C., and Hung, Y. M. (2021). The potential role of nontyphoidal salmonellosis in gastric cancer: a nationwide matched cohort study. *Gastric Cancer*, 24(2), 292-301.

138. Duijster, J. W., Hansen, J. V., Franz, E., Neefjes, J. J., Frisch, M., Mughini-Gras, L., et al. (2021). Association between *Salmonella* infection and colon cancer: a nationwide registry-based cohort study. *Epidemiol. Infect*., 149, E56.

139. Iyer, P., Barreto, S. G., Sahoo, B., Chandrani, P., Ramadwar, M. R., Shrikhande, S. V., et al. (2016). Non-typhoidal *Salmonella* DNA traces in gallbladder cancer. *Infect. Agents Cancer*, 11(1), 1-4.

140. Kato, I., Boleij, A., Kortman, G. A., Roelofs, R., Djuric, Z., Severson, R. K., et al. (2013). Partial associations of dietary iron, smoking and intestinal bacteria with colorectal cancer risk. *Nutr. Cancer*, 65(2), 169-177.

141. Lu, R., Bosland, M., Xia, Y., Zhang, Y. G., Kato, I., and Sun, J. (2017). Presence of *Salmonella* AvrA in colorectal tumor and its precursor lesions in mouse intestine and human specimens. *Oncotarget*, 8(33), 55104.

142. Mughini-Gras, L., Schaapveld, M., Kramers, J., Mooij, S., Neefjes-Borst, E. A., Pelt, W. V., et al. (2018). Increased colon cancer risk after severe *Salmonella* infection. *PloS One*, 13(1), e0189721.

145. Capoor, M. R., Nair, D., Khanna, G., Krishna, S. V., Chintamani, M. S., and Aggarwal, P. (2008). Microflora of bile aspirates in patients with acute cholecystitis with or without cholelithiasis: a tropical experience. *Braz. J. Infect. Dis.*, 12, 222-225.

146. Caygill, C. P., Hill, M. J., Braddick, M., and Sharp, J. C. (1994). Cancer mortality in chronic typhoid and paratyphoid carriers. *Lancet*, 343(8889), 83-84.

147. Csendes, A., Becerra, M., Burdiles, P., Demian, I., Bancalari, K., and Csendes, P. (1994). Bacteriological studies of bile from the gallbladder in patients with carcinoma of the gallbladder, cholelithiasis, common bile duct stones and no gallstones disease. *Eur. J. Surg*., 160(6-7), 363-367.

148. Dutta, U., Garg, P. K., Kumar, R., and Tandon, R. K. (2000). Typhoid carriers among patients with gallstones are at increased risk for carcinoma of the gallbladder. *Am. J. Gastroenterol*., 95(3), 784-787.

149. Hazrah, P., Oahn, K. T. H., Tewari, M., Pandey, A. K., Kumar, K., Mohapatra, T. M., et al. (2004). The frequency of live bacteria in gallstones. HPB, 6(1), 28-32.

150. Mellemgaard, A. and K. Gaarslev, Risk of hepatobiliary cancer in carriers of *Salmonella* typhi*. J. Nat. Cancer Institute*, 1988. 80(4): p. 288.

151. Nath, G., Singh, H., and Shukla, V. K. (1997). Chronic typhoid carriage and carcinoma of the gallbladder. *Eur. J. Cancer Prev*., 6(6), 557-559.

152. Nath, G., Singh, Y. K., Kumar, K., Gulati, A. K., Shukla, V. K., Khanna, A. K., et al. (2008). Association of carcinoma of the gallbladder with typhoid carriage in a typhoid endemic area using nested PCR. *J. Infect. Dev. Ctries*., 2(04), 302-307.

153. Pandey, M., and Shukla, V. K. (2003). Lifestyle, parity, menstrual and reproductive factors and risk of gallbladder cancer. *Eur. J. Cancer Prev*., 12(4), 269-272.

154. Roa, I., Ibacache, G., Carvallo, J., Melo, A., Araya, J., De Aretxabala, X., et al. (1999). Microbiological study of gallbladder bile in a high risk zone for gallbladder cancer. *Rev. Med.Chile*, 127(9), 1049-1055.

155. Safaeian, M., Gao, Y. T., Sakoda, L. C., Quraishi, S. M., Rashid, A., Wang, B. S., et al. (2011). Chronic typhoid infection and the risk of biliary tract cancer and stones in Shanghai, China. *Infect. Agents Cancer*, 6(1), 1-3.

156. Serra, I., Yamamoto, M., Calvo, A., Cavada, G., Báez, S., Endoh, K., et al. (2002). Association of chili pepper consumption, low socioeconomic status and longstanding gallstones with gallbladder cancer in a Chilean population. *Int. J. Cancer*, 102(4), 407-411.

157. Sharma, V., Chauhan, V. S., Nath, G., Kumar, A., and Shukla, V. K. (2007). Role of bile bacteria in gallbladder carcinoma. *Hepato-gastroenterology*, 54(78), 1622-1625. 158. Shukla, V. K., Singh, H., Pandey, M., Upadhyay, S. K., and Nath, G. (2000). Carcinoma of the gallbladder—is it a sequel of typhoid?. *Dig. Dis. Sci*., 45(5), 900-903.

159. Singh, H., Pandey, M., amd Shukla, V. K. (1996). *Salmonella* carrier state, chronic bacterial infection and gallbladder carcinogenesis. *Eur. J. Cancer Prev*., 5(2), 144.

160. Strom, B. L., Soloway, R. D., Rios-Dalenz, J. L., Rodriguez-Martinez, H. A., West, S. L., Kinman, J. L., et al. (1995). Risk factors for gallbladder cancer. An international collaborative case–control study. *Cancer*, 76(10), 1747-1756.

161. Tewari, M., Mishra, R. R., and Shukla, H. S. (2010). *Salmonella* typhi and gallbladder cancer: report from an endemic region. *Hepatobiliary Pancreat. Dis. Int*., 9(5), 524-530.

162. Vaishnavi, C., Kochhar, R., Singh, G., Kumar, S., Singh, S., and Singh, K. (2005). Epidemiology of typhoid carriers among blood donors and patients with biliary, gastrointestinal and other related diseases. *Microbiol. Immunol*., 49(2), 107-112.

163. Welton, J., Marr, J., and Friedman, S. (1979). Association between hepatobiliary cancer and typhoid carrier state. *Lancet*, 313(8120), 791-794.

164. Yagyu, K., Lin, Y., Obata, Y., Kikuchi, S., Ishibashi, T., Kurosawa, M., et al. (2004). Bowel movement frequency, medical history and the risk of gallbladder cancer death: a cohort study in Japan. *Cancer Sci*., 95(8), 674-678.

165. Nagaraja, V., and Eslick, G. D. (2014). Systematic review with meta-analysis: the relationship between chronic *Salmonella* typhi carrier status and gall-bladder cancer. *Aliment. Pharmacol. Ther*., 39(8), 745-750.

166. Koshiol, J., Wozniak, A., Cook, P., Adaniel, C., Acevedo, J., Azócar, L., et al. (2016). *Salmonella* enterica serovar Typhi and gallbladder cancer: a case–control study and meta-analysis. *Cancer Med*., 5(11), 3310-3235.

173. Qiu, D. C., Hubbard, A. E., Zhong, B., Zhang, Y., and Spear, R. C. (2005). A matched, case–control study of the association between *Schistosoma japonicum* and liver and colon cancers, in rural China. *Ann. Trop. Med. Parasitol*., 99(1), 47-52.

174. Xu, Z., and Su, D. L. (1984). Schistosoma japonicum and colorectal cancer: an epidemiological study in the People's Republic of China. *Int. J. Cancer*, 34(3), 315-318.

175. En-Sheng, Z., Cancer of the colon and schistosomiasis. 1981, SAGE Publications Sage UK: London, England.

176. Iida, F., Iida, R., Kamijo, H., Takaso, K., Miyazaki, Y., Funabashi, W., et al. (1999). Chronic Japanese schistosomiasis and hepatocellular carcinoma: ten years of follow-up in Yamanashi Prefecture, Japan. *Bull. World Health Organ*., 77(7), 573.

177. Madbouly, K. M., Senagore, A. J., Mukerjee, A., Hussien, A. M., Shehata, M. A., Navine, P., et al. (2007). Colorectal cancer in a population with endemic *Schistosoma mansoni*: is this an at-risk population?. *Int. J. Colorectal Dis*., 22(2), 175-181.

178. Toda, K. S., Kikuchi, L., Chagas, A. L., Tanigawa, R. Y., Paranaguá-Vezozzo, D. C., Pfiffer, T., et al. (2015). Hepatocellular carcinoma related to *Schistosoma mansoni* infection: case series and literature review. *J. Clin. Transl. Hepatol*., 3(4), 260.

182. Beck, M., Frodl, R., and Funke, G. (2008). Comprehensive study of strains previously designated *Streptococcus bovis* consecutively isolated from human blood cultures and emended description of *Streptococcus gallolyticus* and *Streptococcus infantarius* subsp. *coli*. *J. Clin. Microbiol*., 46(9), 2966-2972.

183. Corredoira, J. C., Alonso, M. P., Garcia, J. F., Casariego, E., Coira, A., Rodriguez, A., et al. (2005). Clinical characteristics and significance of *Streptococcus salivarius* bacteremia and *Streptococcus bovis* bacteremia: a prospective 16-year study. *Eur. J. Clin. Microbiol. Infect. Dis*., 24(4), 250-255.

184. Gonzlez-Quintela, A., Martınez-Rey, C., Castroagudın, J. F., Rajo-Iglesias, M. C., and Domınguez-Santalla, M. J. (2001). Prevalence of liver disease in patients with *Streptococcus bovis* bacteraemia. *J. Infect*., 42(2), 116-119.

185. Jean, S. S., Teng, L. J., Hsueh, P. R., Ho, S. W., and Luh, K. T. (2004). Bacteremic *Streptococcus bovis* infections at a university hospital, 1992-2001. *J. Formos. Med. Assoc*., 103(2), 118-123.

186. Murray, H. W., and Roberts, R. B. (1978). *Streptococcus bovis* bacteremia and underlying gastrointestinal disease. *Arch. Int. Med*., 138(7), 1097-1099.

187. Ruoff, K. L., Miller, S. I., Garner, C. V., Ferraro, M. J., and Calderwood, S. B. (1989). Bacteremia with *Streptococcus bovis* and *Streptococcus salivarius*: clinical correlates of more accurate identification of isolates. *J. Clin. Microbiol.*, 27(2), 305-308.

188. Zarkin, B. A., Lillemoe, K. D., Cameron, J. L., Effron, P. N., Magnuson, T. H., and Pitt, H. A. (1990). The triad of *Streptococcus bovis* bacteremia, colonic pathology, and liver disease. *Ann. Surg*., 211(6), 786.

189. Boleij, A., Roelofs, R., Schaeps, R. M., Schülin, T., Glaser, P., Swinkels, D. W., et al. (2010). Increased exposure to bacterial antigen RpL7/L12 in early stage colorectal cancer patients. *Cancer*, 116(17), 4014-4022.

190. Corredoira, J., Alonso, M. P., Coira, A., and Varela, J. (2008). Association between *Streptococcus infantarius* (formerly *S. bovis* II/1) bacteremia and noncolonic cancer. *J. Clin. Microbiol.*, 46(4), 1570-1570.

191. Fernández-Ruiz, M., Villar-Silva, J., Llenas-García, J., Caurcel-Díaz, L., Vila-Santos, J., Sanz-Sanz, F., et al. (2010). *Streptococcus bovis* bacteraemia revisited: clinical and microbiological correlates in a contemporary series of 59 patients. *J. Infect*., 61(4), 307-313.

192. Vaska, V. L., and Faoagali, J. L. (2009). *Streptococcus bovis* bacteraemia: identification within organism complex and association with endocarditis and colonic malignancy. *Pathology*, 41(2), 183-186.

193. Boleij, A., Roelofs, R., Danne, C., Bellais, S., Dramsi, S., Kato, I., et al. (2012). Selective antibody response to *Streptococcus gallolyticus* pilus proteins in colorectal cancer patients. *Cancer Prev. Res*., 5(2), 260-265.

194. Garza-González, E., Ríos, M., Bosques-Padilla, F. J., Francois, F., Cho, I., González, G. M., et al. (2012). Immune response against *Streptococcus gallolyticus* in patients with adenomatous polyps in colon. *Int. J. Cancer*, 131(10), 2294-2299.

195. Zammit, S. C., Azzopardi, N., and Ellul, P. (2014). *Streptococcus gallolyticus* bacteraemia in hepatobiliary–pancreatic and colonic pathologies. *QJM: Int. J. Med*., 107(5), 355-361.

196. Butt, J., Romero-Hernández, B., Pérez-Gómez, B., Willhauck-Fleckenstein, M., Holzinger, D., Martin, V., et al. (2016). Association of *Streptococcus gallolyticus* subspecies *gallolyticus* with colorectal cancer: serological evidence. *Int. J. Cancer,* 138(7), 1670-1679.

197. Butt, J., Werner, S., Willhauck-Fleckenstein, M., Michel, A., Waterboer, T., Zörnig, I., et al. (2017). Serology of *Streptococcus gallolyticus* subspecies *gallolyticus* and its association with colorectal cancer and precursors. *Int. J. Cancer*, 141(5), 897-904.

198. Paritsky, M., Pastukh, N., Brodsky, D., Isakovich, N., and Peretz, A. (2015). Association of *Streptococcus bovis* presence in colonic content with advanced colonic lesion. *World J. Gastroenterol*., 21(18), 5663.

199. Butt, J., Blot, W. J., Teras, L. R., Visvanathan, K., Le Marchand, L., Haiman, C. A., et al. (2018). Antibody responses to *Streptococcus Gallolyticus* subspecies *Gallolyticus* proteins in a large prospective colorectal cancer cohort consortium. *Cancer Epidemiol. Prev. Biomark*., 27(10), 1186-1194.

200. Butt, J., Jenab, M., Willhauck-Fleckenstein, M., Michel, A., Pawlita, M., Kyrø, C., et al. (2018). Prospective evaluation of antibody response to *Streptococcus gallolyticus* and risk of colorectal cancer. *Int. J. Cancer*, 143(2), 245-252.

201. Guven, D. C., Dizdar, O., Alp, A., Akdogan Kittana, F. N., Karakoc, D., Hamaloglu, E., et al. (2018). Analysis of *Fusobacterium nucleatum*, *Streptococcus gallolyticus* and *Porphyromonas gingivalis* in saliva in colorectal cancer patients and healthy controls. *J. Clin. Oncol.,* 36(15).

202. Kale, P., Khillan, V., and Sarin, S. K. (2018). Novel association of *Streptococcus gallolyticus* subspecies *pasteurianus* and hepatocelluar carcinoma: opening new frontiers. *Scand. J. Gastroenterol*., 53(10-11), 1354-1357.

203. Sheikh, A. F., Zadeh, A. R. M., Saki, M., Khani, P., Hashemi, S. J., Zadeh, S. S., et al. (2020). Detection of *Streptococcus gallolyticus* in colorectal cancer and inflammatory bowel disease patients compared to control group in southwest of Iran. *Mol. Biol. Rep*., 47(11), 8361-8365.

204. Boleij, A., van Gelder, M. M., Swinkels, D. W., and Tjalsma, H. (2011). Clinical importance of *Streptococcus gallolyticus* infection among colorectal cancer patients: systematic review and meta-analysis. *Clin. Infect. Dis*., 53(9), 870-878.

205. Krishnan, S., and Eslick, G. D. (2014). *Streptococcus bovis* infection and colorectal neoplasia: a meta-analysis. *Colorectal Dis*., 16(9), 672-680.

210. Tanaka, T., Hirata, T., Parrott, G., Higashiarakawa, M., Kinjo, T., Kinjo, T., et al. (2016). Relationship among *Strongyloides stercoralis* infection, human T-cell lymphotropic virus type 1 infection, and cancer: a 24-year cohort inpatient study in Okinawa, Japan. *Am. J. Trop. Med. Hyg*., 94(2), 365.

211. Hirata, T., Kishimoto, K., Kinjo, N., Hokama, A., Kinjo, F., and Fujita, J. (2007). Association between *Strongyloides stercoralis* infection and biliary tract cancer. *Parasitol. Res.*, 101(5), 1345-1348.

Chapter 8

FROM BEDSIDE TO BENCH: A MOLECULAR EPIDEMIOLOGICAL INVESTIGATION INTO THE ASSOCIATION BETWEEN NONTYPHOIDAL SALMONELLA INFECTION AND COLON CANCER

J.W. Duijster
V. Stévenin
C.E. Coipan
D.M. van Elsland A.H.A.M van Hoek
E.A. Neefjes-Borst
E. Franz
L. Mughini-Gras
J.J.C. Neefjes

In preparation

Abstract

Introduction: Epidemiological and experimental research suggests an association between nontyphoidal *Salmonella* (NTS) and colon cancer development. Yet, the drivers and potential mechanisms involved in this putative oncogenic role have not been deciphered. This study aims to unravel a causal link between NTS and colon cancer from the bacterial perspective.

Methods: We performed a matched case-control study based on NTS isolates obtained from 30 individuals who were diagnosed with colon cancer later in life (i.e., case-isolates) and 30 people without colon cancer diagnosis (i.e., control-isolates). All 60 NTS isolates were subjected to the following experiments/analyses: a) *in vitro* infection and host cell transformation assay; b) whole-genome sequencing; c) passage through an *in vitro* model system resembling the human gastrointestinal tract; d) *in vitro* quantification of different carbon (C-), nitrogen (N-), phosphorus (P-) and sulfur (S-) source utilization. The outcomes of the different experiments and analyses were used to assess whether case-isolates were different from control-isolates in terms of genotype or phenotype and whether this was associated with transformation efficiency.

Results: Substantial variation was present in the isolates' capacity to induce infection and cellular transformation *in vitro*, with a tendency towards higher transformation efficiency among the case-isolates. This could, however, not be explained by the genotype, neither were significant genotypic differences observed between case- and control-isolates. However, higher transformation efficiency was correlated with increased metabolic utilization capacity of multiple N-, P- and S-sources.

Conclusion: The outcomes of this study indicate a phenotypic rather than a genotypic driver for transformation efficiency. Yet, RNA sequencing of the isolates can reveal whether expression of genes differs between isolates with a high *versus* a low transformation efficiency.

Introduction

In the last decades, the role of bacteria in the onset and progression of cancers is being gradually acknowledged. Numerous mechanisms have been identified by which bacteria manipulate the host during infection, for instance by alteration of the host signaling pathways, the induction of chromosomal instability or prevention of apoptosis of damaged cells [1]. While *S*. Typhi and *Helicobacter pylori* as causative agents of respectively gallbladder and gastric cancer are well established, more species are added to the list of bacteria (potentially) contributing to cancer formation. An example hereof includes nontyphoidal *Salmonella* (NTS), which provokes cellular transformation in predisposed gallbladder organoids, as well as the development of colon tumors in mice after oral infection with the NTS serovar Typhimurium [2]. This was corroborated by an epidemiological study in the Netherlands where a significant positive association was found between infection with NTS and the risk of colon cancer (CC) [3]. In this registry-based study, routinely collected surveillance data of confirmed human NTS infections were linked at the person-level to nationwide CC diagnosis data. The overall risk of CC among people with a history of NTS infection between 20-60 years of age was 1.5-fold higher as compared to the general Dutch population. The association concerned particularly the proximal part of the colon, with over a two-fold increased risk, whereas no excess risk of distal CC was observed after NTS infection. The proximal colon is the most exposed part to NTS bacteria leaving the ileum (i.e., were NTS mainly resides). Moreover, the CC risk appeared to be higher after infection with the serovar Enteritidis as compared to Typhimurium or other serovars [3]. Whether the observed differences in risk estimates between the serovars can be explained by a difference in oncogenic capacity between NTS serovars or strains and which mechanisms and/or pathways are involved, is not yet known. Moreover, much about how transformation is maintained after *Salmonella* has been cleared from the host and which virulence factor(s) might be involved in the oncogenicity of NTS remains to be elucidated. To address these questions, we conducted a multi-faceted explorative study aiming to investigate the causal link between NTS infection and CC from bedside to bench, at a molecular epidemiology level. We first analyzed whether tumors obtained from people with a reported NTS infection are different in nature compared to tumors from people without such history. Second, we conducted a matched case-control study using a multi-angle approach, including *in vitro* and *in vivo* experiments to assess the infection, invasion, cellular transformation and metabolic capacities of NTS isolates as well as a genomic analysis of NTS isolates.

Materials and methods

Part A - Pathology

Formalin-fixed paraffin-embedded (FFPE) tissue blocks of 24 patients with proximal CC with a registered NTS infection in the past and 67 tissue blocks from age- and gender-matched controls (with proximal CC) without such history were obtained from the Dutch nationwide network and registry of histo- and cytopathology (PALGA).The tissue blocks were sectioned and the sections stained and counter-stained according to standard protocols. The sections were analyzed with immunohistochemistry for p53, c-MYC and MAPK/ERK. In addition, the tumor grading was determined for each of the 91 CC patients by an experienced CC pathologist.

Part B - Case-control study

We used data from the Dutch national surveillance program for *Salmonella* where public health laboratories send NTS isolates from human salmonellosis patients to the National Institute for Public Health and the Environment (RIVM) for serotyping [4]. This data was previously used in the epidemiological study in which 65 people with a reported NTS infection between January 1999-December 2015 were identified who developed proximal CC ≥1 year after the reported salmonellosis [3]. Of these 65 people, the NTS isolates from 30 people were still available for further analyses. We focused on cancers in the proximal part of the colon (ICD-10 codes C180-C185), as the association between NTS and CC was strongest for the proximal colon. Hence, we defined these 30 NTS isolates as cases. In addition, we selected NTS isolates (in the surveillance database) obtained from people with salmonellosis who did not develop CC during the period January 1999-December 2015 as controls. The 30 case isolates were matched on serovar, type of infection (enteric, septicemic, other [urinary, wound etc.]), year of infection, age at infection and gender to 30 controls (1:1 ratio), totaling 60 isolates.

Gastrointestinal tract model system

Prior to assessing the cellular transformation capacities of the 60 NTS isolates, we studied their host invasion and host infection potential in an *in vitro* model. The NTS isolates were cultured overnight at 37°C and subsequently exposed to conditions resembling the human digestive tract in a gastrointestinal tract (GIT) model system consisting of two parts: the simulated gastrointestinal passage and the attachment and invasion assay (Supplementary Figure S1) [5-7]. First, an overnight culture (ON) of each NTS isolate was sequentially exposed to simulated gastric fluid

(SGF) and simulated intestinal fluid (SIF) at 37°C for 30 minutes and 2 hours, respectively. After that, differentiated Caco-2 cells mimicking the small intestinal epithelium were inoculated with the SGF/SIF/bacterial-mixture at 37 °C for 1 hour on a 12-well plate to test the bacterial attachment (ATT) and invasion (INV) capacities. Between each step of the GIT model (ON, SGF, SIF, ATT and INV), serial 10-fold dilutions of samples were made and NTS bacteria present were enumerated. For quantification of attachment, 6 out of 12 wells containing the Caco-2 cells were lysed (to release attached and invaded bacteria), whereas for enumeration of invaded bacteria only, the other 6/12 wells were treated with gentamicin to kill attached bacteria before lysing cells to release invaded NTS. Details about the cell cultures, the dilutions steps, the compositions of SGF and SIF and the experimental procedures are described elsewhere [5-7]. The GIT model was applied for each of the 60 NTS isolates separately. For analysis of the change in bacterial count between each of the steps in the GIT model, we used a hierarchical Bayesian framework by applying Markov chain Monte Carlo sampling assuming Poisson distributed bacterial counts and lognormal distributed concentrations [6, 7]. Following the methodology of Wijnands et al. (2017), we calculated the *in vitro* infectivity (expressed as log P[inf]) as the sum of all log changes in NTS concentrations throughout the GIT model from the overnight culture until the concentration of invaded bacteria. The bacterial count data for all GIT system stages of the 60 NTS isolates were subjected to principal component analysis (PCA) to assess whether isolates obtained from cases differ from those obtained from controls in terms of their behavior/survival in the GIT model system [6]. Statistical analyses were performed in RStudio version 1.4 1103.

Bacterial strains and cell lines for NTS infection and transformation assays

S. Typhimurium strain SL1344 was a courtesy of S. Méresse [8]. This strain was used as reference strain in the *in vitro* infection and transformation assays. Mouse Embryonic Fibroblasts (MEFs) were derived from *Arf*-deficient C57BL/6 mice. MEFs overexpressing c-MYC were generated by retroviral transduction using a pLZRS-GFP(ires)-HA backbone. MEFs were cultured at 37°C, 5% CO2 in Dulbecco's Modified Eagle's Medium (DMEM) (Invitrogen) [2].

In vitro NTS infection, CFU and transformation assays

NTS infection of MEFs was performed as described previously [9]. In brief, NTS were grown overnight at 37°C in LB medium. The next day, the bacteria were sub-cultured at a dilution of 1:33 in fresh LB medium and incubated for 2 hours at 37°C while shaking. Cells were infected with NTS at multiplicity of infection (MOI) 20 in DMEM medium without antibiotics for 20 minutes at 37°C, 5% CO_2 in a tissue culture chamber and then incubated in the presence

of 100 µg/mL gentamicin (GIBCO) for 1 hour to eliminate extracellular bacteria. In the CFU assays, cells were lysed with lysis buffer (ddH2O + 1%NP-40), and serial dilutions of the lysate were plated on LB plates. In the anchorage-independent growth assays, MEFs were cultured for another 2 hours in the presence of 10 µg/mL gentamicin. The infected MEFs were subsequently collected and resuspended in DMEM medium supplemented with 10 µg/mL gentamicin and 0.35% low melting point agarose (UltraPure™, Invitrogen) and were poured on a soft agar bottom layer consisting of 0.7% low melting point agarose in DMEM with 10 µg/mL gentamicin. Anchorage-independent cell growth and number of soft agar colonies were assessed after 1-3 weeks of incubation at 37°C, 5% CO_2 using GelCountTM (Oxford Optronix, UK). For microscopy analysis, samples were fixed with 4% paraformaldehyde for 10 min at room temperature, and stained with rabbit polyclonal anti-*S*. Typhimurium LPS (Difco, Detroit, MI) and DAPI (Life Technologies). Images were acquired using a Leica TCS SP8 (Leica Microsystems, Wetzlar, Germany) at 40x or 63x magnification. Every experiment was performed in triplicate.

Genomic analysis

DNA extraction, whole genome sequencing (WGS) and assembly

As a next step in unraveling the role of NTS in cancer formation, we analyzed whether the degree of transformation capacity of the NTS isolates could be explained by differences in presence/absence of virulence genes. To this end, all 60 isolates were submitted to whole-genome sequencing (WGS). DNA isolation, 2×125 bp paired-end library preparation and WGS analysis on a HiSeq 2500 platform (Illumina) was performed by BaseClear (Leiden, the Netherlands). All resulting fastq files were subjected to quality control with CheckM v1.0.7 [10], and *de novo* assembled using an in-house developed pipeline (https://github.com/RIVM-bioinformatics/Juno_pipeline). The assembled genomes were analyzed with the SISTR application to confirm the *Salmonella* serovar [11]. Genome annotation of the assembled genomes was performed with Prokka v1.14.6 [12]. Next, the annotation data was used as input for Roary v3.13.0 to construct the core- and accessory genome of the 60 isolates, with a blastp identity cut-off of 95%, indicating a presence in at least 57/60 NTS isolates for a gene to be defined as part of the core-genome [13]. The core genome alignment from Roary was used to build a phylogenetic tree in RAxML v8.2.12 [14]. Single nucleotide polymorphisms (SNPs) were extracted using parsnp v1.2 [15]. The large number of serovars in the dataset, with consequently high genetic variability among the sequences, restrained us from calling SNPs on the full dataset. Instead, we created subsets comprising only *S*. Enteritidis, only *S*. Typhimurium, and both these serovars taken together. High density nucleotide

polymorphisms were filtered out with Gubbins v2.3.4 [16]. Protein function annotation was performed with Pannzer2 [17].

Genome-wide association analysis (GWAS)

Several genome-wide association tests were conducted using the R-package TreeWAS, which accounts for recombination and population structure [18]. Reconstruction of ancestral states was done with the parsimony method, and phylogenetic trees were constructed using maximum likelihood (ML) phylogeny. For all association tests, Bonferroni correction was applied to account for multiple testing. The outcome variable used in the association tests was the ranked transformation capacity of the bacterial strains, as inferred from the *in vitro* tests (see above). As differences in properties between bacterial strains may be the result of a modified structure of proteins or mutations in regulatory regions, which cannot be assessed from presence/absence of genes, we also performed the analysis at SNP (single nucleotide polymorphism) level. Statistical analyses were performed with R v3.6.2.

Phenotype Microarray analysis

In addition to the genomic analyses, we investigated phenotypic traits of the 60 NTS isolates by means of analyzing the utilization of carbon (C), nitrogen (N), phosphorus (P) and sulfur (S) sources. For this, we used the BioLog® Phenotype MicroArray (plates PM1, PM3 and PM4), which allows for high-throughput metabolic quantification of bacterial respiration and growth on a range of different substrates [19, 20]. Briefly, the quantification is based on redox technology, in which cell respiration is measured by the degree of irreversible reduction of a tetrazolium dye. Following Biolog instructions, a cell suspension of each individual cultured NTS isolate and a defined medium (including a dye) were added to 96-well plates containing a single C-, N-, P-, or S-source in each well. Plates were incubated at 37°C for 24h and color formation was measured every 15 minutes using an ELx808 Microplate Reader and Gen5 software (BioTek). The analyses were performed twice for each of the 60 NTS isolates. The ratio of the integrals of each C-, N-, P-, and S-source and a negative control (i.e. the PM1, PM3 and PM4 plates contain a negative control well without substrate for each source type), was used as outcome for further analysis. An hierarchical cluster analysis (HCA) and principal component analysis (PCA) was performed on the scaled data for each of the three plates to compare the metabolic phenotypes of NTS isolates obtained from cases *versus* controls taking into account the transformation capacity of the strains (as defined in the transformation assays). The analyses were performed in RStudio version 1.4 1103.

Results

Part A- Pathology

The objective of the pathology examination of the tumors was to determine whether the patients with a history of (severe) salmonellosis have different types of tumors with regard to the aforementioned markers as compared to the patients without such exposure. None of the markers (p53, c-MYC, MAPK/ERK) was significantly associated with a history of reported *Salmonella* infection. The tumors from patients with a history of *Salmonella* infection showed a tendency to be less likely undifferentiated than those of the patients without a reported *Salmonella* infection (odds ratio 0.21, 95% CI 0.04-1.06; p 0.059) (Figure 1, Supplementary Table S1 and Figure S2).

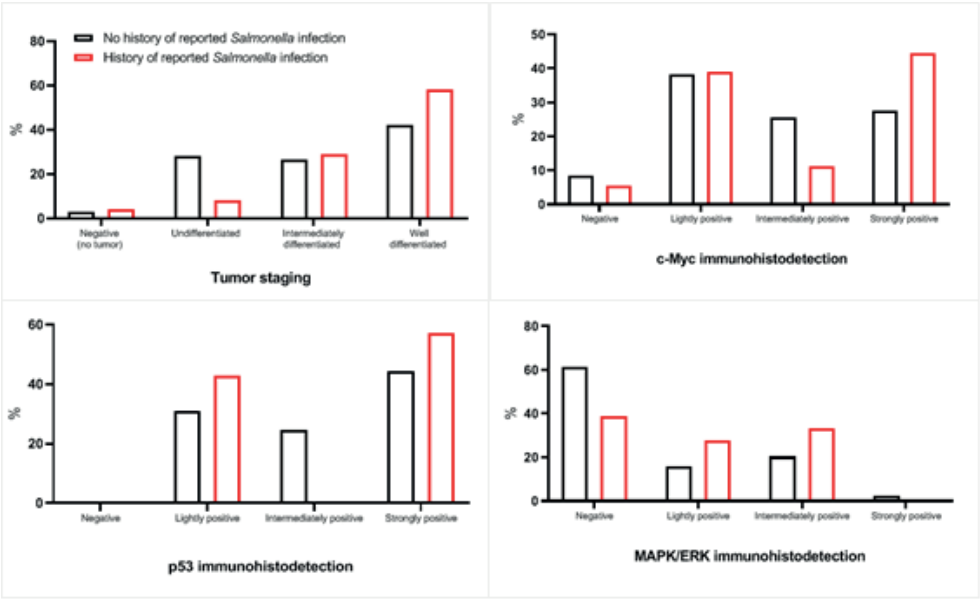


Figure 1. Immunohistochemistry and tumor staging results of the colon tumor blocks from patients with and without history of severe salmonellosis.

Part B – Case-control study

Description of study population

Supplementary table S2 shows the characteristics of the patients from whom the 30 case and 30 matched control NTS isolates were obtained. Two-third of the salmonellosis patients were male. The median age at infection was 63 years (interquartile range [IQR] 51-72) for controls and 67 years (IQR 55-76) for cases. Eleven different NTS serovars were included, mostly Enteritidis (n=22) and Typhimurium (including its monophasic variant) (n=18) (Figure 2, Supplementary table S2). Serovars other than Typhimurium and Enteritidis are hereafter referred to as 'other'. The vast majority (87%) of isolates were obtained from feces (i.e. enteric infections). One pair of Typhimurium isolates were obtained from blood or other normally sterile sites (i.e. invasive infections) and three pairs of isolates, belonging to the serovars Enteritidis, Typhimurium and Hadar, were obtained from urine or wound infections (Supplementary table S2).

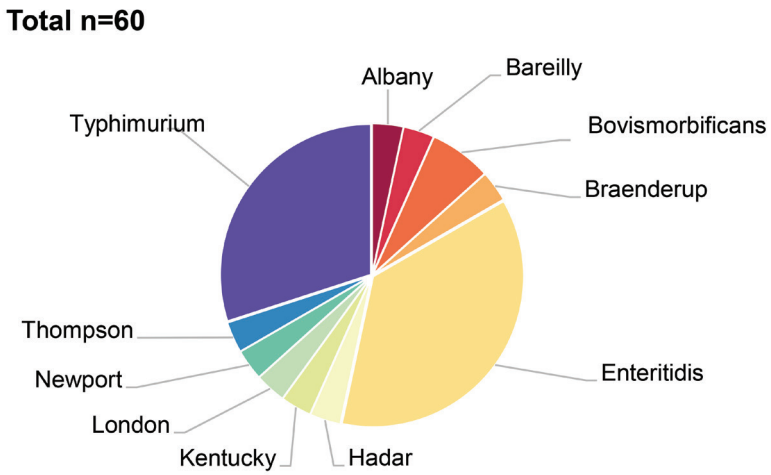


Figure 2. Serovar distribution of the 60 nontyphoidal *Salmonella* isolates.

Gastrointestinal tract model system

The mean *in vitro* infectivity (P[inf]) tends to be higher in NTS isolates obtained from cases (-1.74 ± 0.69; range -5.60 / -0.08) as compared to isolates obtained from controls (-1.39 ± 0.20; range -3.30 / -0.22), though (just) not statistically significant (paired t-test: t(29): 1.85, p 0.07). Figure 3 shows the log fractions of surviving bacteria in each of the four transitions

(SGF/ON, SIF/SGF, ATT/SIF, INV/ATT) of the GIT model for isolates obtained from cases *versus* controls as well as the mean *in vitro* infectivity. No statistically significant difference was observed between isolates obtained from cases *versus* controls for all four transitions (based on conditional logistic regression analysis). These results were confirmed by the outcomes of the PCA in which no clusters could be observed (Supplementary Figure 3). In fact, the data showed a negative correlation between the level of attachment (ATT_SIF) and invasion (INV_ATT), as well as between the amount of bacteria surviving the gastric fluid (SGF_ON) and those surviving intestinal fluid (SIF_SGF), as indicated by the arrows in Supplementary Figure 3.

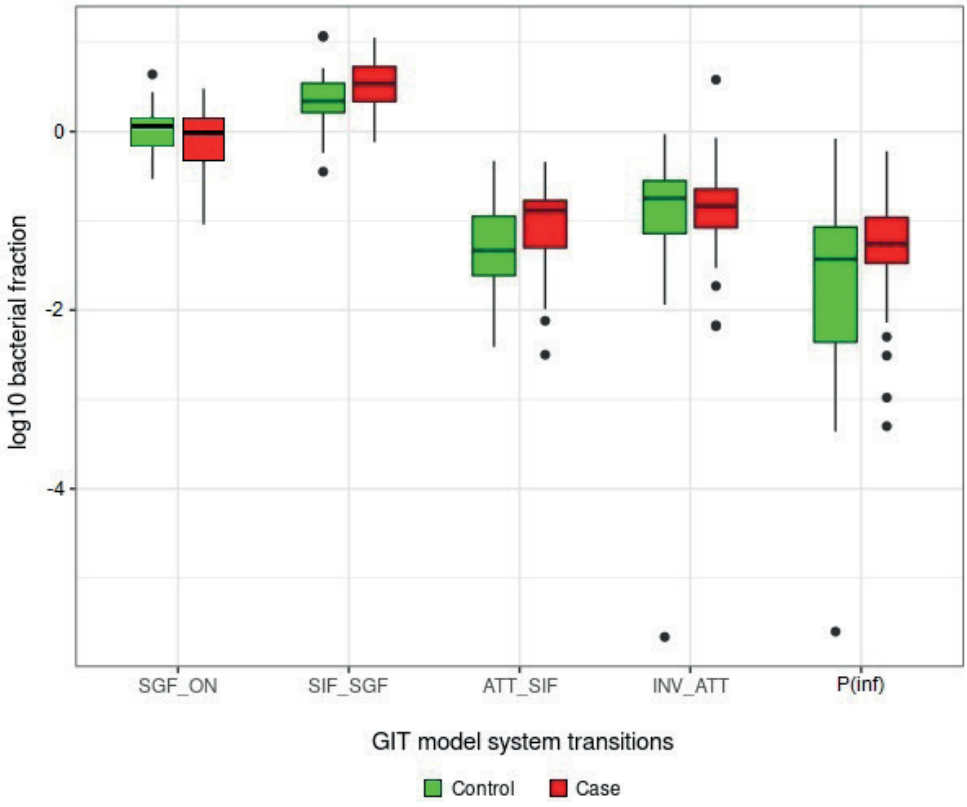


Figure 3. Log transformed fractions of bacterial counts in each of the transitions (survival of gastric fluid and intestinal fluid, attachment and invasion) in the gastrointestinal tract (GIT) model system for nontyphoidal *Salmonella* isolates obtained from cases vs. controls. SGF_ON: simulated gastric fluid vs. overnight bacterial culture; SIF_SGF: simulated intestinal fluid vs. simulated gastric fluid; ATT_SIF: attachment vs. simulated intestinal fluid; INV_ATT: invasion vs. attachment; P(inf): mean in vitro infectivity (INV/ON).

Infection and transformation assays

In the *in vitro* analyses, the infection and transformation capacity of the 60 NTS strains was assessed in *Arf-/-* + c-MYC MEFs. Infection and transformation capacities were normalized against the *S*. Typhimurium reference strain. Ten isolates, belonging to the serovars Enteritidis (n=4), Typhimurium (n=2), Albany (n=1), Bovismorbificans (n=1), Hadar (n=1) and Newport (n=1) failed to infect the MEFs (Figure 4). These included eight control isolates and two case isolates. Case isolates had a lower average infection efficiency compared to control isolates (0.55 *vs.* 0.58), whereas the opposite was true for the transformation efficiency (1.62 *vs.* 1.17) (Figure 4, Figure 5). Twenty-four isolates showed a higher infection efficiency as compared to the *S*. Typhimurium reference strain, 14 of these were case isolates (n=6 Typhimurium, n=4 Enteritidis, n=4 other serovars) (Figure 4, Supplementary Figure S4a, S4c). With regard to transformation efficiency, 18 case isolates (60%) and 10 control isolates (33%) showed a higher transformation efficiency as compared to the reference strain (Figure 5, Supplementary Figure S4b, S4d). Yet, the differences in infection and transformation capacity did not reach the level of significance (t-tests, both the infection and transformation capacities were not significantly higher for case isolates as compared to control isolates). Nonetheless, there was a tendency towards higher similar/higher infection efficiency in the case isolates (Figure 4). Likewise, case isolates were associated with higher transformation efficiencies, (Figure 5b, 5c).

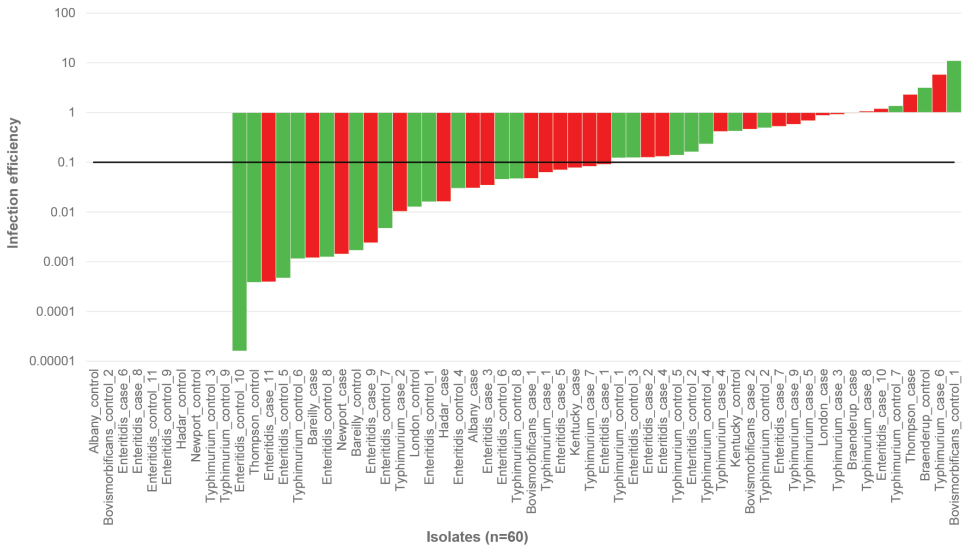


Figure 4. Mean infection potential of the 60 nontyphoidal *Salmonella* isolates obtained from cases (n=30; red bars) and controls (n=30; green bars) expressed in colony forming units (CFU) normalized against the infection efficiency of the *S*. Typhimurium reference strain. Results derive from three independent experiments with technical triplicates. Ten isolates were unable to infect mouse embryonic fibroblasts. Isolates with an infection efficiency >0.1 performed better than the laboratory strain.

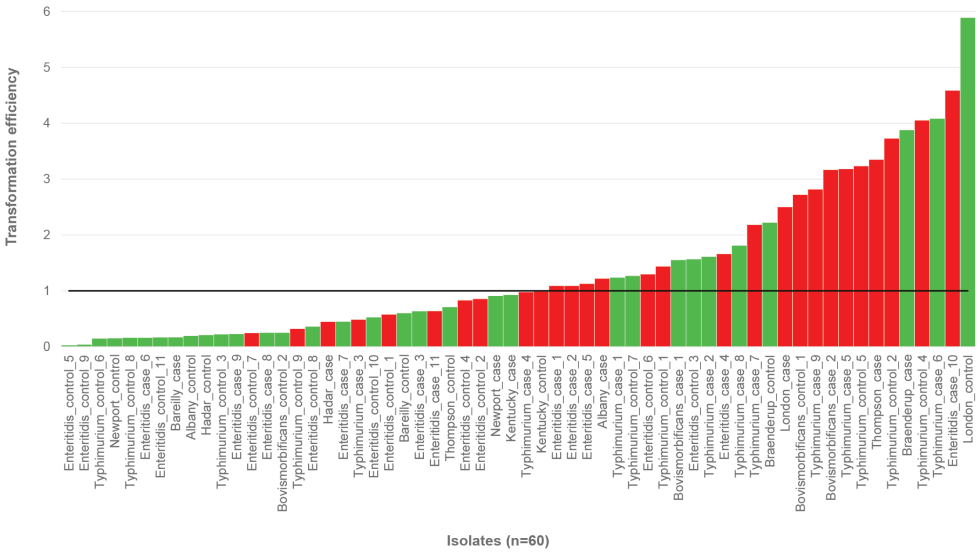


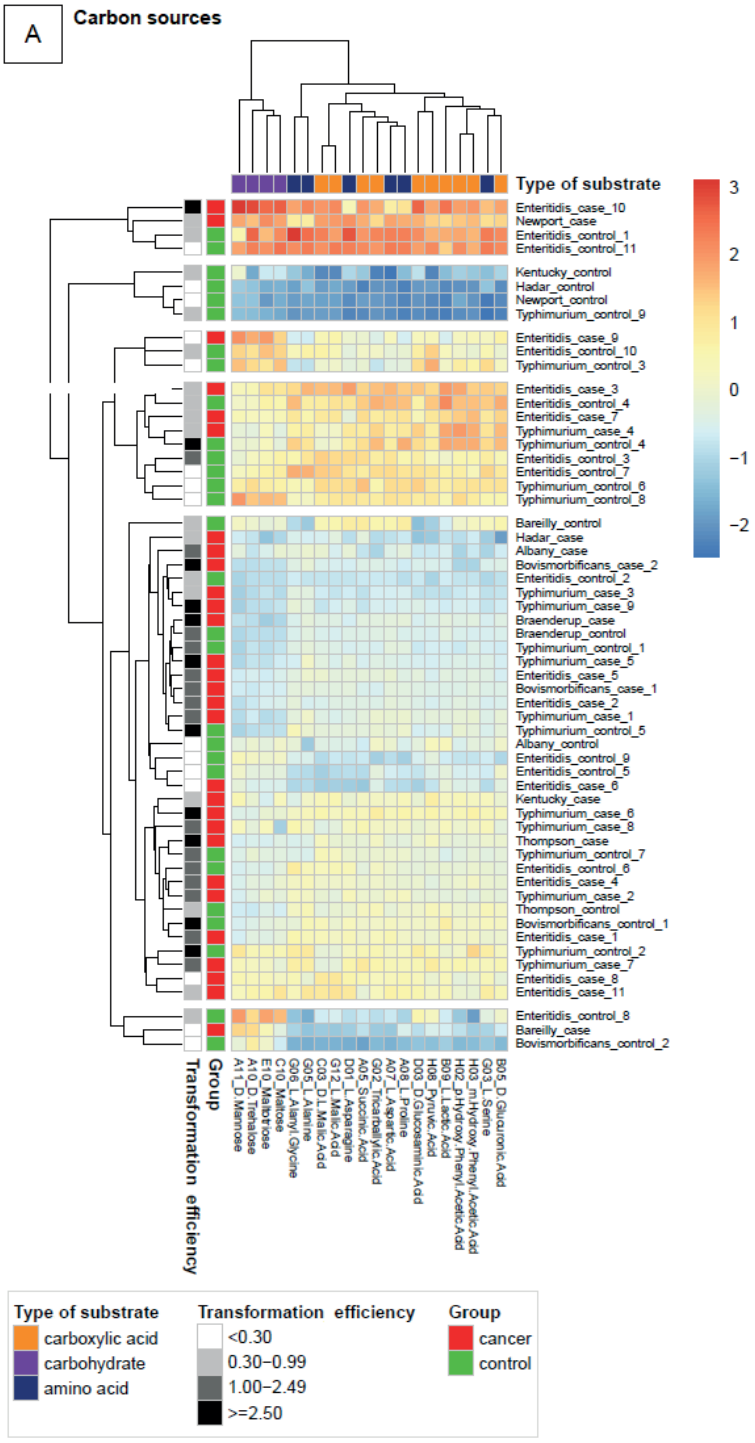
Figure 5. Mean transformation efficiencies of the nontyphoidal *Salmonella* isolates obtained from cases (n=30; red bars) and controls (n=30; green bars) normalized against the transformation efficiency of the *S*. Typhimurium reference strain. Results derive from three independent experiments with technical triplicates. Isolates with a transformation efficiency >1 performed better than the laboratory strain. Isolate 'London_control' shows an unexpected high transformation efficiency, likely as a result of sample contamination.

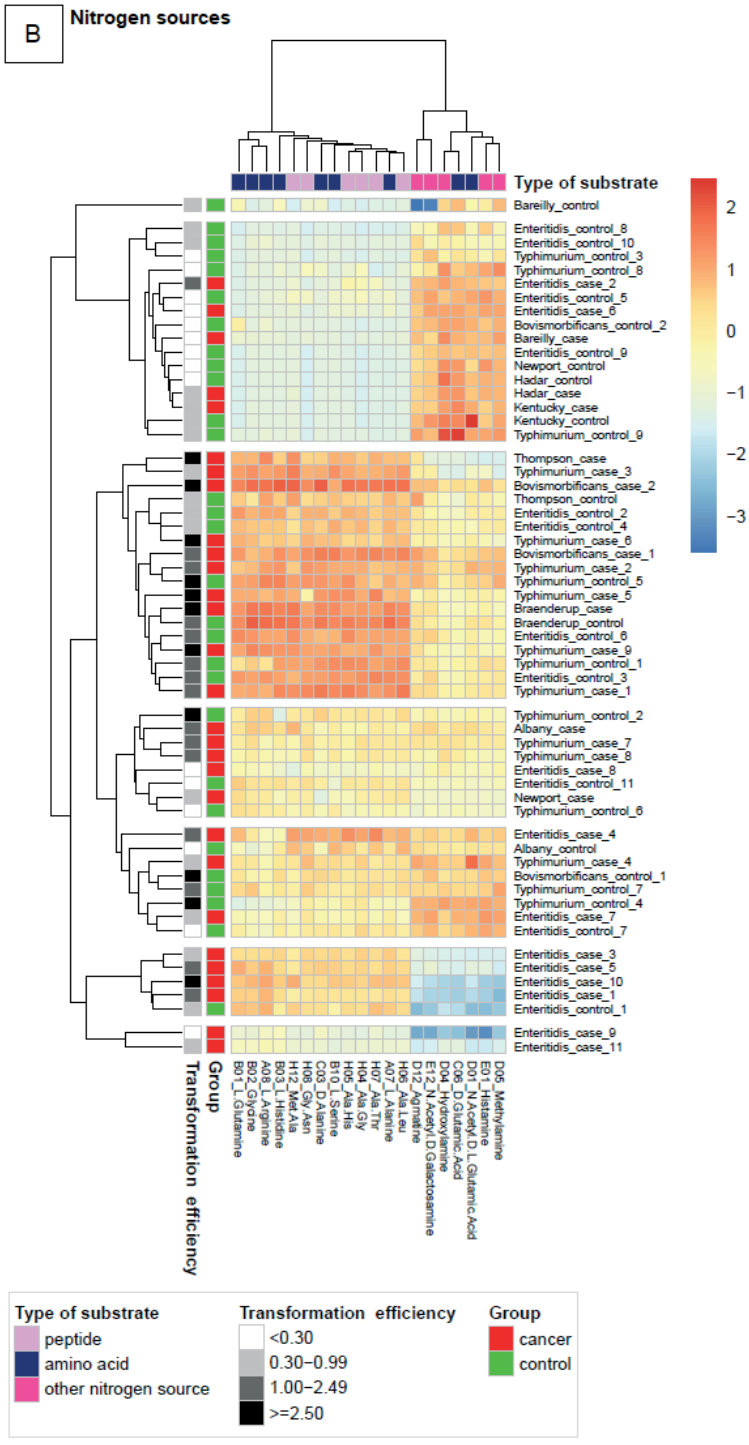
Genomic analysis

The genes in the pangenome of NTS as inferred by Prokka and Roary were used to inform the association tests performed by TreeWAS. We restricted the analysis to *S*. Enteritidis and *S*. Typhimurium isolates as these totaled 22 and 18 isolates respectively (in contrast to all other serovars with 2-4 isolates each). Presence of five genes appeared significantly associated with a higher transformation efficiency in the *S*. Typhimurium subset and four in the combined Enteritidis/Typhimurium subset (Supplementary Table S3). One of the significant genes of the *S*. Typhimurium subset and three from the combined Enteritidis/Typhimurium subset had unknown functions. For the remaining genes, functional annotation revealed three proteins involved in UV protection and mutation of which two are part of the bacterial SOS response to DNA damage (UmuC and UmuD). There were seven SNPs significantly associated with the ranked transformation efficiency of each of *S*. Enteritidis and *S*. Typhimurium subsets, and only one SNP significantly associated with the combined Enteritidis/Typhimurium subset. The genes in which these SNPs are located were associated with several functions including DNA cleavage and transcription activities (Supplementary Table S3).

Metabolic characterization of NTS isolates

In the metabolic analysis of bacterial phenotypes we analyzed 95 sole sources of carbon (PM1), 95 nitrogen sources (PM3) and 59 and 35 sole sources of respectively phosphorus and sulfur (PM4) (Supplementary Table S4). Conditional logistic regression was used to assess the association between substrate utilization and the likelihood of the host being diagnosed with colon cancer later in life. We observed a significant ($p<0.05$) positive association between isolates from cases and utilization of 11 substrates (10 phosphorus sources, 1 nitrogen source) and a negative association for 10 substrates (7 nitrogen sources, 3 carbon sources). Spearman's rank correlation test revealed 135 significant positive correlations and 34 significant negative correlations between substrate utilization and transformation efficiency (Supplementary Table S4). After Bonferroni correction for multiple testing, 76 positive correlations remained, comprising 34, 27 and 15 N- and P-/S-sources respectively. The remaining eight significant inverse correlations included 6 carbon sources and 2 sulfur sources (Supplementary Table S4). The tendency towards increased substrate utilization for isolates with a higher transformation efficiency was particularly pronounced for amino acids and several phosphorus sources. Before conducting a PCA, we defined the optimal clustering method for hierarchical clustering. The average linkage clustering showed the best fit to the data (lowest Gower distance and highest cophenetic correlation). Supplementary Figure S5 shows the heatmap of the scaled utilization scores of all 60 NTS isolates for the 76 positive correlations and eight inverse correlations with isolates being clustered using average linkage. PCAs on the scaled data revealed that the first two principal components (PCs) accounted for 58.7-63.5% of the total variance for the three plates (Supplementary Figures S6-S8). Ninety percent of the variance was explained by 13, 15 and 14 PCs for C-, N- and P-/S-sources respectively. Figure 6 shows the heatmaps of the nutrient utilization of the 60 isolates for the 20 nutrients with the highest contribution to the variance in the data, for the C-, N- and P-S-plate respectively. Particularly for several amino acids, peptides and phosphorus and sulfur sources the degree utilization tends to be higher for isolates with a higher transformation efficiency (part of the sources as also depicted in Supplementary Figure S5).


A
Carbon sources
Type of substrate
3
2
1
0
−1
−2
Enteritidis_case_10
Newport_case
Enteritidis_control_1
Enteritidis_control_11
Kentucky_control
Hadar_control
Newport_control
Typhimurium_control_9
Enteritidis_case_9
Enteritidis_control_10
Typhimurium_control_3
Enteritidis_case_3
Enteritidis_control_4
Enteritidis_case_7
Typhimurium_case_4
Typhimurium_control_4
Enteritidis_control_3
Enteritidis_control_7
Typhimurium_control_6
Typhimurium_control_8
Bareilly_control
Hadar_case
Albany_case
Bovismorbificans_case_2
Enteritidis_control_2
Typhimurium_case_3
Typhimurium_case_9
Braenderup_case
Braenderup_control
Typhimurium_control_1
Typhimurium_case_5
Enteritidis_case_5
Bovismorbificans_case_1
Enteritidis_case_2
Typhimurium_case_1
Typhimurium_control_5
Albany_control
Enteritidis_control_9
Enteritidis_control_5
Enteritidis_case_6
Kentucky_case
Typhimurium_case_6
Typhimurium_case_8
Thompson_case
Typhimurium_control_7
Enteritidis_control_6
Enteritidis_case_4
Typhimurium_case_2
Thompson_control
Bovismorbificans_control_1
Enteritidis_case_1
Typhimurium_control_2
Typhimurium_case_7
Enteritidis_case_8
Enteritidis_case_11
Enteritidis_control_8
Bareilly_case
Bovismorbificans_control_2
Transformation efficiency
Group
A11_D Mannose
A10_D Trehalose
E10_Maltotriose
C10_Maltose
G06_L Alanyl Glycine
G05_L Alanine
C03_D,L Malic Acid
G12_L Malic Acid
D01_L Asparagine
A05_Succinic Acid
G02_Tricarballylic Acid
A07_L Aspartic Acid
A08_L Proline
D03_D Glucosaminic Acid
H08_Pyruvic Acid
B09_L Lactic Acid
H02_p Hydroxy Phenyl Acetic Acid
H03_m Hydroxy Phenyl Acetic Acid
G03_L Serine
B05_D Glucuronic Acid
Type of substrate
carboxylic acid
carbohydrate
amino acid
Transformation efficiency
<0.30
0.30−0.99
1.00−2.49
>=2.50
Group
cancer
control


B
Nitrogen sources
Type of substrate
Bareilly_control
Enteritidis_control_8
Enteritidis_control_10
Typhimurium_control_3
Typhimurium_control_8
Enteritidis_case_2
Enteritidis_control_5
Enteritidis_case_6
Bovismorbificans_control_2
Bareilly_case
Enteritidis_control_9
Newport_control
Hadar_control
Hadar_case
Kentucky_case
Kentucky_control
Typhimurium_control_9
Thompson_case
Typhimurium_case_3
Bovismorbificans_case_2
Thompson_control
Enteritidis_control_2
Enteritidis_control_4
Typhimurium_case_6
Bovismorbificans_case_1
Typhimurium_case_2
Typhimurium_control_5
Typhimurium_case_5
Braenderup_case
Braenderup_control
Enteritidis_control_6
Typhimurium_case_9
Typhimurium_control_1
Enteritidis_control_3
Typhimurium_case_1
Typhimurium_control_2
Albany_case
Typhimurium_case_7
Typhimurium_case_8
Enteritidis_case_8
Enteritidis_control_11
Newport_case
Typhimurium_control_6
Enteritidis_case_4
Albany_control
Typhimurium_case_4
Bovismorbificans_control_1
Typhimurium_control_7
Typhimurium_control_4
Enteritidis_case_7
Enteritidis_control_7
Enteritidis_case_3
Enteritidis_case_5
Enteritidis_case_10
Enteritidis_case_1
Enteritidis_control_1
Enteritidis_case_9
Enteritidis_case_11
Transformation efficiency
Group
B01_L.Glutamine
B02_Glycine
A08_L.Arginine
B03_L.Histidine
H12_Met.Ala
H08_Gly.Asn
C03_D.Alanine
B10_L.Serine
H05_Ala.His
H04_Ala.Gly
H07_Ala.Thr
A07_L.Alanine
H06_Ala.Leu
D12_Agmatine
E12_N.Acetyl.D.Galactosamine
D04_Hydroxylamine
C06_D.Glutamic.Acid
D01_N.Acetyl.D.L.Glutamic.Acid
E01_Histamine
D05_Methylamine
2
1
0
−1
−2
−3
Type of substrate
peptide
amino acid
other nitrogen source
Transformation efficiency
<0.30
0.30−0.99
1.00−2.49
>=2.50
Group
cancer
control

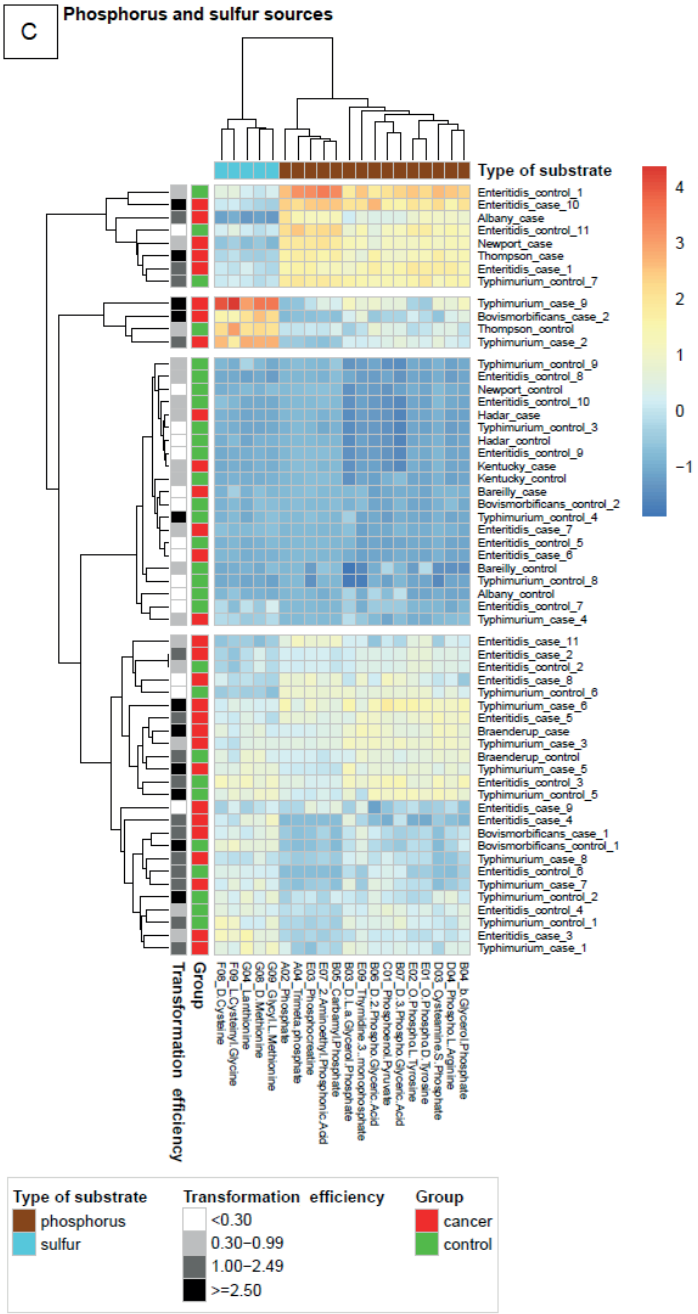


Figure 6. Heatmap of scaled utilization scores of the 60 nontyphoidal *Salmonella* (NTS) isolates for the top 20 sources mostly contributing to the variance in the carbon (A), nitrogen (B) and phosphorus/sulfur (C) data. NTS isolates are clustered based their utilization scores using the average linkage method. Left annotation depicts the group (case *vs.* control) and transformation efficiency (in four categories).

Discussion

Colon cancer ranks among the highest cancer incidences worldwide and the complexity of all microbial factors putatively contributing to the development of colon tumors is gradually being acknowledged. Literature on the role of microorganisms reveals a number of commensal and pathogenic bacteria associated with the induction of malignancies and progression of tumor growth. Yet, from a mechanistical perspective, a lot is unknown about possible pathways involved and whether the oncogenic potential of bacteria might be attributable to certain bacterial genes. For several bacteria, the oncogenic potential is restricted to strains/serotypes expressing specific genes or producing toxins [21-23]. With regard to NTS, multiple (effector) proteins have been identified that manipulate host cell-signaling pathways to escape immunity, reduce inflammation and apoptosis and enhance bacterial proliferation [1]. As collateral damage from host cell manipulation induced in the infection cycle, host cells can enter a cancerous state as part of the multistep process of cancer formation [1]. Whether the possible tumorigenic potential of NTS is attributable to serovar specific traits or genes is not yet investigated.

Here we describe the first study assessing possible genotypic or phenotypic traits of NTS related to cell transformation potential. Also we compared characteristics and markers of tumors derived from proximal colon cancer patients with a notified *Salmonella* infection in the past (*Salmonella*$^+$) to tumors from patients without such reported infection (*Salmonella*$^-$). A higher portion of tumors from *Salmonella*$^+$ patients were well-differentiated as compared to tumors from *Salmonella*$^-$ patients. Generally, well-differentiated (low-grade) tumors have a much better prognosis compared to poorly-differentiated tumors. These outcomes correspond to earlier research which revealed also a slightly higher percentage of well-differentiated proximal colon tumors in patients with a past *Salmonella* infection [3]. Immunohistochemistry of the tumor suppressor p53 and proto-oncogene c-MYC showed excess staining in tumors from *Salmonella*$^+$ patients, though not significant. Yet, many of the tissue samples were old and have been stored at different places throughout the Netherlands which presumably affected the labelling. Mutations in p53 and c-MYC are frequently observed in cancer patients, high levels of p53 staining are indicative for (mutated) inactivated p53 whereas c-MYC overexpression is presumably associated with tumor staging [2, 24-26]. Earlier experimental research showed that *S*. Typhimurium is able to induce tumor formation in mice as well as cellular transformation in gallbladder organoids and MEFs when these organisms/cell harbored inactivated p53 and overexpressed c-MYC [2]. The outcomes of a previous study corroborated this and we here demonstrate that also other NTS strains are able to induce the cellular transformations in MEFs [27].

Substantial variation in infection and transformation efficiency was observed between the isolates. Despite the lack of significant correlations between infectivity and transformation efficiency and disease outcome (cancer or no cancer), there was a tendency towards a higher transformation efficiency for the case-isolates. To our knowledge, this is the first study in which the transformation capacity was assessed for multiple NTS isolates allowing for comparison between isolates. In earlier studies, transformation assays were performed with *S*. Typhimurium [2, 27], yet our study clearly showed that other serovars might have an even higher transformation capacity. The results of the genomic analysis suggest several proteins involved in SOS response to be associated with transformation capacity in the *S*. Typhimurium subset. Nonetheless, these results should be interpreted with caution, as variants of the same gene might be differentially present between *S*. Typhimurium isolates. We observed this for the gene encoding the UmuC protein (data not shown). Whether a higher transformation capacity is related to restricted variants of genes (as could be for the UmuC gene) or the number of copies or variants present is not yet known. Furthermore, we observed several SNPs associated with transformation capacity in the *S*. Enteritidis, *S*. Typhimurium and combined Enteritidis/ Typhimurium subsets. The functions of the genes in which these SNPs were located were diverse and did not consistently indicate a possible role in cellular transformation.

The phenotypic microarray assay showed significant correlations between transformation efficiency and utilization scores of several nitrogen, phosphorus and sulfur sources. A high degree of metabolic flexibility confers a biological advantage for the bacterium as the availability and amount of nutrients changes during its infectious cycle [28]. Several putrefaction pathways have been identified by which commensal and pathogenic bacteria utilize amino acids released during fermentation of undigested proteins by resident bacteria [29]. During amino acid fermentation harmful metabolites such as H_2S, amines, phenol, indole and histidine are produced. Some of these metabolites have been associated with progression of colon cancer [29, 30]. How the metabolic signatures of specific NTS strains might induce a cascade of events and which pathways can be involved, will be the subject of further studies. A limitation of this study, which might have affected the observations, is the large portion (70%) of cases with a reported *Salmonella* infection above the age of 60. The risk of colon cancer increases substantially in older people, mostly as a result of the age-related accumulating of mutations associated with cancer [31]. This might have diluted some of the results, though unfortunately the sample size restrained us from restricting the analysis to the younger subgroups.

The outcomes of this study suggest that the oncogenic potential of *Salmonella* is better explained by phenotypic rather than genotypic traits of the isolates. Despite the absence of

relevant significant associations between gene presence and transformation in our study, we consider supplementing the genomic analyses with RNA sequencing analyses a good step forwards to improve our understanding of the underlying mechanisms. RNA sequencing provides information about the genes actually expressed rather than the presence/absence of genes and gene mutations as identified by DNA sequencing. This method revealed that the degree of pathogenicity among relatively genetically homogeneous strains of *S*. Enteritidis can be attributed to a multitude of genes differentially expressed between high and low pathogenic strains. Those genes were distributed over a range of functional classes including carbohydrate and amino acid metabolism, biogenesis and cell motility [32]. Applying RNA sequencing to the 60 NTS strains in our study might potentially identify differences between case- and control-isolates relevant for tumorigenesis.

Overall, this study revealed that tumors from colon cancer patients with a notified *Salmonella* infection in the past differ from tumors obtained from patients without such reported *Salmonella* infection with regard to tumor differentiation, yet tumor markers were not different between those groups. Moreover, the capacity to induce cellular transformation in MEFs varied between *Salmonella* isolates, with a tendency towards better transformation efficiency for isolates derived from people who were diagnosed with colon cancer later in life. This transformation efficiency was significantly correlated to utilization capacity of multiple nitrogen, phosphorus and sulfur sources. More in depth research is needed to unravel possible mechanisms and metabolic pathways which might be involved in the *Salmonella*-induced colon cancer development/progression.

References

1. Van Elsland D, Neefjes J. Bacterial infections and cancer. EMBO Reports. 2018;19(11):e46632.
2. Scanu T, Spaapen RM, Bakker JM, Pratap CB, Wu LE, Hofland I, et al. *Salmonella* manipulation of host signaling pathways provokes cellular transformation associated with gallbladder carcinoma. Cell Host & Microbe. 2015;17(6):763-74.
3. Mughini-Gras L, Schaapveld M, Kramers J, Mooij S, Neefjes-Borst EA, van Pelt W, et al. Increased colon cancer risk after severe *Salmonella* infection. PloS One. 2018;13(1):e0189721.
4. Van Pelt W, de Wit MA, Wannet WJ, Ligtvoet EJ, Widdowson MA, van Duynhoven YT. Laboratory surveillance of bacterial gastroenteric pathogens in The Netherlands, 1991-2001. Epidemiology and Infection. 2003;130(3):431-41.
5. Oliveira M, Wijnands L, Abadias M, Aarts H, Franz, E. Pathogenic potential of Salmonella typhimurium DT104 following sequential passage through soil, packaged fresh-cut lettuce and a model gastrointestinal tract. International Journal of Food Microbiology. 2011;148:149–55.
6. Wijnands LM, Teunis PF, Kuijpers AF, Asch DV, Ellen HM, Pielaat A. Quantification of Salmonella survival and infection in an in vitro model of the human intestinal tract as proxy for foodborne pathogens. Frontiers in Microbiology. 2017;8:1139.
7. Kuijpers AF, Bonacic Marinovic AA, Wijnands LM, Delfgou-van Asch EH, van Hoek AH, Franz E, et al. Phenotypic prediction: linking in vitro virulence to the genomics of 59 *Salmonella* enterica strains. Frontiers in Microbiology. 2019;9:3182.
8. Méresse S, Steele-Mortimer O, Finlay BB, Gorvel JP. The rab7 GTPase controls the maturation of *Salmonella* typhimurium-containing vacuoles in HeLa cells. The EMBO Journal. 1999;18:4394-4403.
9. Kuijl C, Savage NDL, Marsman M, Tuin AW, Janssen L, Egan DA, et al. Intracellular bacterial growth is controlled by a kinase network around PKB/AKT1. Nature. 2007; 450:725.
10. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research. 2015;25(7):1043-55.
11. Yoshida CE, Kruczkiewicz P, Laing CR, Lingohr EJ, Gannon VP, Nash JH, et al. The Salmonella in silico typing resource (SISTR): an open web-accessible tool for rapidly typing and subtyping draft *Salmonella* genome assemblies. PloS One. 2016;11(1):e0147101.
12. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics, 2014;30(14):2068-9.
13. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691-3.
14. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312-3.
15. Treangen TJ, Ondov BD, Koren S, Phillippy, A. M. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biology. 2014;15(11):1-15.
16. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Research. 2015;43(3):e15-e15.
17. Collins C, Didelot X. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. PLoS Computational Biology. 2018;14(2):e1005958.

18. Shea A, Wolcott M, Daefler S, Rozak DA. Biolog Phenotype Microarrays. In: Navid A. (eds) Microbial Systems Biology. Methods in Molecular Biology (Methods and Protocols), 2012;vol 881.
19. Biolog. (2021, November 23). Phenotype MicroArrays for Microbial Cells – Biolog. https://www.biolog.com/products-portfolio-overview/phenotype-microarrays-for-microbial-cells/
20. Hatakeyama M. Structure and function of Helicobacter pylori CagA, the first-identified bacterial protein involved in human cancer. Proceedings of the Japan Academy, Series B. 2017;93(4):196-219.
21. Cheng WT, Kantilal HK, Davamani, F. The mechanism of Bacteroides fragilis toxin contributes to colon cancer formation. The Malaysian journal of medical sciences: MJMS. 2020;27(4):9.
22. Pleguezuelos-Manzano C, Puschhof J, Huber AR, van Hoeck A, Wood HM, Nomburg J, et al. Mutational signature in colorectal cancer caused by genotoxic pks+ E. coli. Nature. 2020;580(7802):269-73.
23. Kaserer K, Schmaus J, Bethge U, Migschitz B, Fasching S, Walch A, et al. Staining patterns of p53 immunohistochemistry and their biological significance in colorectal cancer. The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland. 2000;190(4):450-6.
24. Toon CW, Chou A, Clarkson A, DeSilva K, Houang M, Chan JC, et al. Immunohistochemistry for myc predicts survival in colorectal cancer. PloS One. 2014;9(2):e87456.
25. Nakayama M, Sakai E, Echizen K, Yamada Y, Oshima H, Han TS, et al. Intestinal cancer progression by mutant p53 through the acquisition of invasiveness associated with complex glandular formation. Oncogene. 2017;36(42):5885-96.
26. Van Elsland DM, Duijster JW, Zhang J, Stévenin V, Zhang YG, Zha L, Xia Y, Franz E, Sun J, Mughini-Gras L, Neefjes JJC. Repetitive exposure to non-typhoidal Salmonellae is an environmental risk factor for colon cancer. Cell Host and Microbe 2021; submitted [see Chapter 4]
27. Reitzer L. Catabolism of amino acids and related compounds. EcoSal Plus. 2005;1(2).
28. Kaur H, Das C, Mande SS. In silico analysis of putrefaction pathways in bacteria and its implication in colorectal cancer. Frontiers in Microbiology. 2017;8:2166.
29. Gerner EW. Impact of dietary amino acids and polyamines on intestinal carcinogenesis and chemoprevention in mouse models. Biochemical Society Transactions. 2007;35(2):322-5.
30. Risques RA, Kennedy SR. Aging and the rise of somatic cancer-associated mutations in normal tissues. PLoS Genetics. 2018;14(1):e1007108.
31. Shah DH. RNA sequencing reveals differences between the global transcriptomes of *Salmonella* enterica serovar enteritidis strains with high and low pathogenicities. Applied and Environmental Microbiology, 2014;80(3):896-906.

Supplementary material

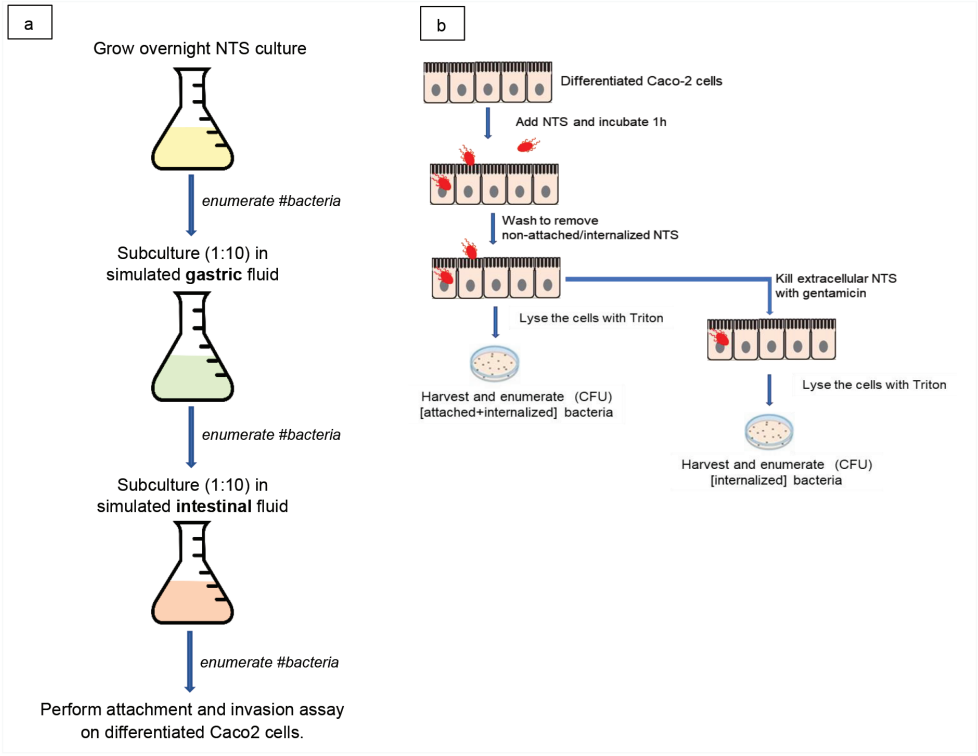


Figure S1. Schematic representation of the simulated gastrointestinal passage (A) and the attachment and invasion assay (B) of the gastrointestinal tract (GIT) model system. NTS: nontyphoidal *Salmonella*. Adapted from: Wijnands LM, Teunis PF, Kuijpers AF, Asch DV, Ellen HM, and Pielaat A. (2017). Quantification of *Salmonella* survival and infection in an in vitro model of the human intestinal tract as proxy for foodborne pathogens. Front. Microbiol., 8, 1139.

Table S1. Immunohistochemistry and tumor staging results of the colon tumor blocks from patients with and without history of severe salmonellosis.

	No history of reported NTS infection	History of reported NTS infection
Differentiation		
Negative (no tumor)	2 (3.13%)	1 (4.17%)
Well differentiated	27 (42.19%)	14 (58.33%)
Intermediately differentiated	17 (26.56%)	7 (29.17%)
Undifferentiated	18 (28.13%)	2 (8.33%)
c-MYC		
Negative	4 (8.51%)	1 (5.56%)
Lightly positive	18 (38.30%)	7 (38.89%)
Intermediately positive	12 (25.53%)	2 (11.11%)
Strongly positive	13 (27.66%)	8 (44.44%)
MAPK/ERK		
Negative	27 (61.36%)	7 (38.89%)
Lightly positive	7 (15.91%)	5 (27.78%)
Intermediately positive	9 (20.45%)	6 (33.33%)
Strongly positive	1 (2.27%)	0 (0.00%)
P53		
Negative	0 (0.00%)	0 (0.00%)
Lightly positive	19 (31.15%)	9 (42.86%)
Intermediately positive	15 (24.59%)	0 (0.00%)
Strongly positive	27 (44.26%)	12 (57.14%)

NTS: nontyphoidal *Salmonella.*

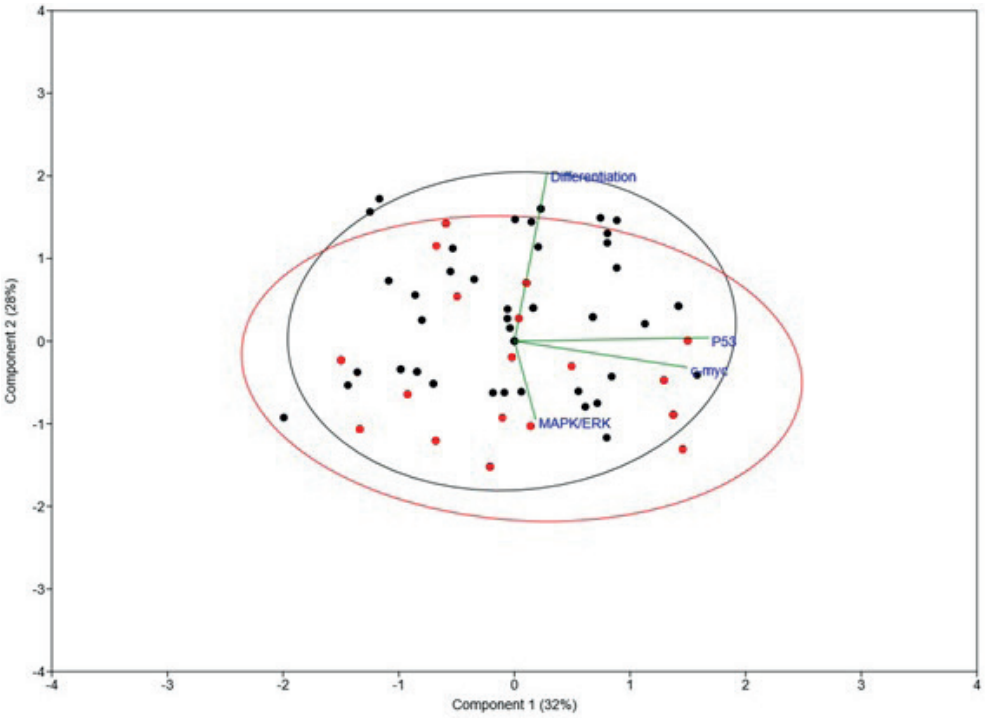


Figure S2. Principal component analysis of colon tumor blocks from colon cancer patients with (red dots) a without (black dots) reported history of *Salmonella* infection as a function of tumor differentiation and immunochemistry (markers p53, c-MYC, and fosfo-AKT/ERK). The principal components 1 and 2 respectively explain 32% and 28% of the total variance in the data.

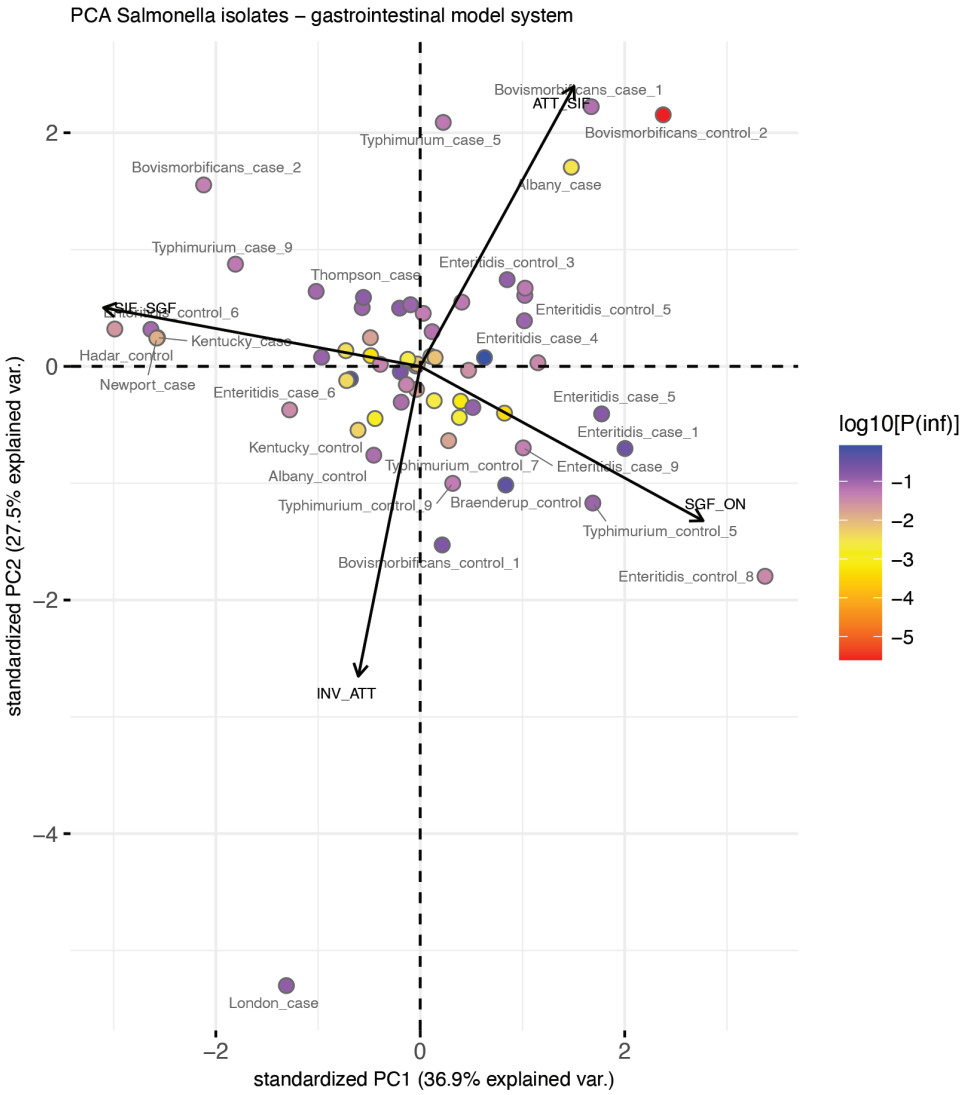


Figure S3. Principal component analysis (PCA) biplot of the fractions of enumerated bacteria in the steps of the gastrointestinal tract (GIT) model system for all 60 nontyphoidal *Salmonella* (NTS) isolates. The color scale of the dots reflects the log P(inf). The principal components (PC) 1 and PC2 explained 36.9% and 27.5% of the variance in the data respectively.

Table S2. Characteristics of the nontyphoidal *Salmonella* infection in individuals who developed colon cancer later in life (i.e. cases) versus those who did not develop cancer (i.e. controls).

Pair	Case/ control	Follow-up (years)	Serotype	Type of infection	Age at infection	Gender
#1	Case	13.6	Enteritidis	Enteric	50-54	Female
	Control	-	Enteritidis	Enteric	50-54	Female
#2	Case	10.4	Enteritidis	Enteric	35-39	Female
	Control	-	Enteritidis	Enteric	30-34	Female
#3	Case	5.4	Enteritidis	Enteric	70-74	Female
	Control	-	Enteritidis	Enteric	65-69	Female
#4	Case	2.6	Enteritidis	Enteric	40-44	Female
	Control	-	Enteritidis	Enteric	40-44	Female
#5	Case	9.9	Enteritidis	Other	65-69	Male
	Control	-	Enteritidis	Other	65-69	Male
#6	Case	1.3	Enteritidis	Enteric	45-49	Female
	Control	-	Enteritidis	Enteric	45-49	Female
#7	Case	3.4	Enteritidis	Enteric	55-59	Female
	Control	-	Enteritidis	Enteric	55-59	Female
#8	Case	7.8	Enteritidis	Enteric	75-79	Female
	Control	-	Enteritidis	Enteric	75-79	Female
#9	Case	4.0	Enteritidis	Enteric	75-79	Male
	Control	-	Enteritidis	Enteric	75-79	Male
#10	Case	4.0	Enteritidis	Enteric	75-79	Female
	Control	-	Enteritidis	Enteric	70-74	Female
#11	Case	1.9	Enteritidis	Enteric	70-74	Female
	Control	-	Enteritidis	Enteric	70-74	Female
#12	Case	5.4	Typhimurium	Enteric	60-64	Male
	Control	-	Typhimurium	Enteric	60-64	Male
#13	Case	13.7	Typhimurium	Enteric	35-39	Male
	Control	-	Typhimurium	Enteric	35-39	Male
#14	Case	3.4	Typhimurium	Enteric	65-69	Male
	Control	-	Typhimurium	Enteric	70-74	Male
#15	Case	2.1	Typhimurium*	Other	75-79	Male
	Control	-	Typhimurium*	Other	75-79	Female
#16	Case	1.1	Typhimurium	Enteric	50-54	Male
	Control	-	Typhimurium	Enteric	50-54	Male
#17	Case	4.3	Typhimurium	Septicemic	65-69	Male
	Control	-	Typhimurium*	Septicemic	60-64	Female
#18	Case	3.0	Typhimurium*	Enteric	70-74	Female
	Control	-	Typhimurium*	Enteric	70-74	Female
#19	Case	1.5	Typhimurium	Enteric	75-79	Female
	Control	-	Typhimurium*	Enteric	75-79	Female
#20	Case	1.1	Typhimurium*	Enteric	75-79	Female
	Control	-	Typhimurium*	Enteric	75-79	Female

Pair	Case/ control	Follow-up (years)	Serotype	Type of infection	Age at infection	Gender
#21	Case	3.1	Albany	Enteric	60-64	Female
	Control	-	Albany	Enteric	50-54	Female
#22	Case	4.3	Bareilly	Enteric	70-74	Male
	Control	-	Bareilly	Enteric	20-24	Female
#23	Case	7.1	Bovismorbificans	Enteric	65-69	Female
	Control	-	Bovismorbificans	Enteric	65-69	Female
#24	Case	8.0	Bovismorbificans	Enteric	60-64	Male
	Control	-	Bovismorbificans	Enteric	50-54	Female
#25	Case	11.5	Braenderup	Enteric	60-64	Female
	Control	-	Braenderup	Enteric	50-59	Male
#26	Case	1.1	Hadar	Other	75-79	Female
	Control	-	Hadar	Other	70-74	Female
#27	Case	1.9	Kentucky	Enteric	50-54	Male
	Control	-	Kentucky	Enteric	50-54	Male
#28	Case	4.2	London	Enteric	65-69	Female
	Control	-	London	Enteric	65-69	Female
#29	Case	4.3	Newport	Enteric	50-54	Female
	Control	-	Newport	Enteric	50-54	Female
#30	Case	2.5	Thompson	Enteric	75-79	Male
	Control	-	Thompson	Enteric	75-79	Male

*monophasic variant as shown by WGS analysis.

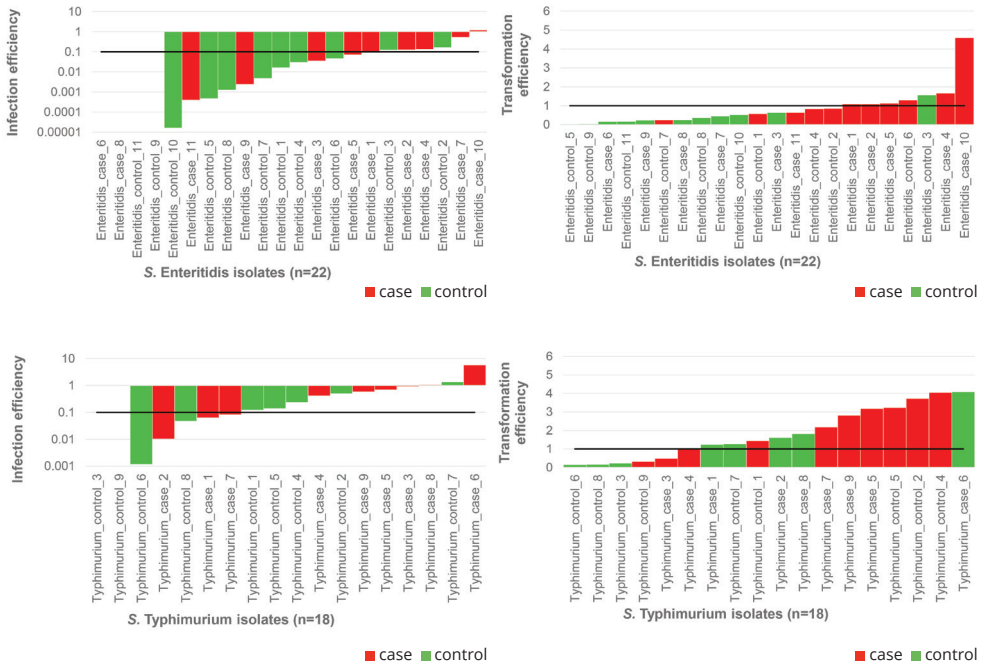


Figure S4. Mean infection and transformation efficiency of the 22 *S*. Enteritidis isolates (a, b) and 18 *S*. Typhimurium isolates (c, d) obtained from cases (red bars) and controls (green bars). Infection and transformation efficiencies are expressed in colony forming units normalized against the infection efficiency of the *S*. Typhimurium reference strain. Four Enteritidis isolates and two Typhimurium isolates were unable to infect mouse embryonic fibroblasts. Isolates with an infection efficiency of 0.1 and a transformation efficiency of 1.0 (black reference lines) performed better than the laboratory strain.

Table S3. Genes and SNPs significantly associated with transformation efficiency in *S*. Enteritidis, *S*. Typhimurium and combined Typhimurium/Enteritidis subsets.

Gene name	Annotation	Gene function
***S*. Typhimurium subset**		
group_2788	DNA-invertase hin [https://www.uniprot.org/uniprot/P03013]	Synthesis of phase-2 flagellin.
group_2807	protein ImpC [https://www.uniprot.org/uniprot/P0A1G0]	Belongs to the imp operon which has a function in UV protection and mutation.
group_1096	protein UmuC	Functions in UV protection and mutation, induced/SOS mutation.
group_420	*unknown*	*unknown*
umuD_2	protein umuD [https://www.uniprot.org/uniprot/P22493]	Functions in UV protection and mutation, induced/SOS mutation.
Combined *S*. Typhimurium and *S*. Enteritidis subset		
group_2065	*unknown*	*unknown*
group_2540	*unknown*	*unknown*
group_4752	*unknown*	*unknown*
group_4753	regulatory protein rop [https://www.uniprot.org/uniprot/P03051]	Regulatory role in plasmid DNA replication.
SNP locus name [gene name]	**Annotation**	**Gene function**
***S*. Enteritidis subset**		
502 [ehaB]	Autotransporter/virulence factor	Cell surface protein. Biofilm formation.
503 [res]	Type III restriction-modification system endonuclease	DNA cleavage.
504 [res]	Type III restriction-modification system endonuclease	DNA cleavage.
505 [res]	Type III restriction-modification system endonuclease	DNA cleavage.
506 [hsdR]	Type I restriction enzyme R protein	DNA cleavage. Nuclease and ATPase activities.
507 [*unknown*]	Cytoplasmic protein	*unknown*
508 [wcaI]	Glycosyltransferase, group 1 family protein	Glycosyltransferase activity.
***S*. Typhimurium subset**		
1926 [yegE]	Anti-FlhC(2)FlhD(4) factor YdiV	Transcription regulation. Virulence.
2058 [YacL]	UPF0231 protein YacL	*unknown*
2154 [*unknown*]	ISNCY family transposase	*unknown*
2367 [hxlA]	Putative hexulose 6 phosphate synthase	Formaldehyde fixation (ribulose monophosphate pathway).

Gene name	Annotation	Gene function
3604 [rfaB]	Lipopolysaccharide core heptose(I) kinase	Adding glycosyl residue to the core lipopolysaccharide. Detergent resistance.
3751 [rpoB]	DNA-directed RNA polymerase subunit beta	Catalyzation of the transcription of DNA into RNA.
4579 [*unknown*]	RpoE-regulated lipoprotein	*unknown*
Combined *S*. Typhimurium and *S*. Enteritidis subset		
20679 [oppA]	Periplasmic oligopeptide-binding protein	Peptide transmembrane transporter activity.

SNP: single nucleotide polymorphism.
IQR: interquartile range."

Table S4. Median and interquartile range of carbon, nitrogen, phosphorus and sulfur source utilization by case isolates (n=30) and control isolates (n=30) and Spearman correlation coefficient (rho) and p-value for the correlation between source utilization and transformation efficiency.

	Control isolates	Case isolates	Spearman		
	Median (IQR)	Median (IQR)	*rho*	p-value	Sig
PM1 – carbon sources					
L-Arabinose	1.72 (1.26-2.31)	1.62 (1.30-2.22)	-0.2350	0.071	
N-Acetyl-D-Glucosamine	2.44 (1.48-3.02)	2.06 (1.50-2.58)	-0.3716	<0.01	
D-Saccharic Acid	2.45 (1.01-3.56)	2.89 (1.27-3.22)	0.4127	<0.01	
Succinic Acid	3.03 (2.44-3.94)	3.01 (2.73-3.62)	0.1880	0.150	
D-Galactose	2.71 (1.88-3.56)	2.32 (2.03-3.22)	-0.3101	<0.05	
L-Aspartic Acid	2.91 (2.34-3.71)	3.03 (2.62-3.47)	0.1431	0.276	
L-Proline	3.13 (2.35-3.93)	3.10 (2.76-3.79)	0.2184	0.094	
D-Alanine	2.58 (1.60-3.08)	2.58 (2.24-3.02)	0.3207	<0.05	
D-Trehalose	2.79 (1.82-3.83)	2.41 (1.88-3.20)	-0.2954	<0.05	
D-Mannose	2.35 (1.73-3.04)	2.04 (1.72-2.84)	-0.3079	<0.05	
Dulcitol	2.62 (2.04-3.45)	2.49 (1.98-3.16)	-0.2152	0.099	
D-Serine	3.52 (2.68-4.25)	3.32 (2.97-3.60)	0.0699	0.595	
D-Sorbitol	2.97 (2.45-3.47)	2.78 (2.20-3.12)	-0.1900	0.146	
Glycerol	3.31 (2.56-4.32)	3.24 (2.88-3.74)	0.1173	0.372	
L-Fucose	2.37 (2.04-3.14)	2.25 (1.94-3.04)	-0.1807	0.167	
D-Glucuronic Acid	3.95 (3.07-4.58)	3.70 (3.28-4.47)	0.0440	0.739	
D-Gluconic Acid	3.88 (3.28-4.79)	3.97 (3.55-4.79)	0.1120	0.394	
D,L-α-Glycerol-Phosphate	3.49 (2.66-4.06)	3.31 (2.93-3.86)	0.1332	0.310	
D-Xylose	2.03 (1.36-2.83)	2.41 (1.68-3.05)	0.1695	0.195	
L-Lactic Acid	3.08 (2.56-3.58)	3.03 (2.72-3.61)	0.1313	0.317	
Formic Acid	0.98 (0.93-1.13)	0.99 (0.92-1.10)	0.3365	<0.05	
D-Mannitol	2.50 (1.74-3.20)	1.97 (1.65-2.74)	-0.3204	<0.05	
L-Glutamic Acid	2.18 (1.85-2.88)	2.00 (1.75-2.44)	-0.0356	0.787	
D-Glucose-6-Phosphate	4.19 (3.72-4.89)	4.38 (3.65-5.09)	-0.0166	0.900	
D-Galactonic Acid-y-Lactone	1.92 (1.29-2.50)	1.45 (1.06-1.91)	-0.5263	<0.001	+

	Control isolates	Case isolates	Spearman		
	Median (IQR)	Median (IQR)	*rho*	p-value	Sig
D,L-Malic Acid	2.77 (2.14-3.60)	2.86 (2.53-3.44)	0.1081	0.411	
D-Ribose	2.46 (1.48-3.19)	2.44 (2.01-2.84)	0.0543	0.680	
Tween 20	2.24 (1.72-2.95)	2.19 (1.95-2.84)	0.2136	0.101	
L-Rhamnose	2.06 (1.54-2.54)	1.66 (1.43-2.20)	-0.3300	<0.05	
D-Fructose	2.83 (1.88-3.26)	2.31 (1.84-2.88)	-0.3292	<0.05	
Acetic Acid	2.13 (1.37-2.52)	2.11 (1.83-2.60)	0.1909	0.144	
α-D-Glucose	2.57 (1.93-3.16)	2.02 (1.62-2.83)	-0.2838	<0.05	
Maltose	2.74 (2.13-3.76)	2.63 (2.19-3.07)	-0.2182	0.094	
D-Melibiose	2.76 (2.16-4.35)	2.87 (2.07-3.39)	-0.2008	0.124	
Thymidine	3.34 (2.69-3.78)	3.15 (2.66-3.56)	0.1426	0.277	
L-Asparagine	2.87 (2.28-3.77)	2.93 (2.64-3.42)	0.2035	0.119	
D-Aspartic Acid	1.91 (1.56-2.31)	2.01 (1.80-2.33)	0.2859	<0.05	
D-Glucosaminic Acid	3.09 (2.43-4.05)	3.08 (2.67-3.63)	0.1207	0.358	
1,2-Propanediol	0.98 (0.91-1.05)	0.99 (0.86-1.04)	-0.0462	0.726	
Tween 40	2.48 (1.72-3.17)	2.41 (2.25-2.86)	0.3645	<0.01	
α-Keto-Glutaric Acid	1.18 (1.11-1.34)	1.23 (1.12-1.33)	0.3115	<0.05	
α-Keto-Butyric Acid	1.96 (1.44-2.57)	2.02 (1.80-2.31)	0.0471	0.721	
α-Methyl-D-Galactoside	3.13 (2.49-4.22)	2.74 (2.31-3.52)	-0.1918	0.142	
α-D-Lactose	1.18 (1.06-1.32)	1.23 (1.12-1.28)	0.3368	<0.05	
Lactulose	0.86 (0.63-0.97)	0.70 (0.57-0.85)	-0.4206	<0.01	
Sucrose	1.05 (1.01-1.14)	1.13 (0.98-1.22)	0.1119	0.395	
Uridine	3.95 (2.76-5.17)	3.97 (3.48-4.96)	0.1195	0.363	
L-Glutamine	1.95 (1.54-2.21)	1.85 (1.59-2.34)	-0.0308	0.815	
M-Tartaric Acid	2.07 (1.23-2.83)	2.12 (1.55-2.34)	0.0718	0.585	
D-Glucose-1-Phosphate	3.06 (1.90-3.90)	3.41 (2.34-3.90)	0.2860	<0.05	
D-Fructose-6-Phosphate	3.78 (2.66-4.51)	3.82 (3.43-4.64)	0.3266	<0.05	
Tween 80	1.98 (1.47-2.45)	2.04 (1.87-2.45)	0.3719	<0.01	
α-Hydroxy Glutaric Acid-γ-Lactone	1.13 (0.92-1.31)	1.03 (0.78-1.18)	-0.2633	<0.05	
α-Hydroxy Butyric Acid	1.88 (1.56-2.64)	2.05 (1.83-2.28)	0.1966	0.132	
ß-Methyl-D-Glucoside	1.45 (1.29-1.76)	1.66 (1.33-1.80)	0.2906	<0.05	
Adonitol	1.10 (1.05-1.18)	1.10 (1.01-1.16)	-0.0123	0.926	
Maltotriose	3.01 (2.48-4.01)	2.79 (2.48-3.59)	-0.2102	0.107	
2-Deoxy Adenosine	4.01 (3.39-5.13)	3.81 (3.33-4.54)	-0.1582	0.227	
Adenosine	4.37 (3.40-5.05)	3.84 (3.38-4.57)	-0.127	0.335	
Glycyl-L-Aspartic Acid	2.47 (1.98-2.95)	2.18 (2.05-2.75)	0.138	0.294	
Citric Acid	2.95 (1.37-3.63)	2.99 (2.68-3.44)	0.359	<0.01	
M-Inositol	0.99 (0.89-1.21)	0.90 (0.72-1.52)	0.010	0.940	
D-Threonine	0.95 (0.88-1.16)	0.93 (0.80-1.05)	-0.268	<0.05	
Fumaric Acid	2.88 (2.26-3.37)	2.85 (2.37-3.20)	0.015	0.913	
Bromo Succinic Acid	2.35 (1.71-2.98)	2.35 (1.96-2.80)	0.155	0.237	
Propionic Acid	2.29 (1.30-2.92)	2.35 (2.04-2.71)	0.278	<0.05	
Mucic Acid	2.11 (1.05-3.45)	2.52 (1.11-3.11)	0.298	<0.05	
Glycolic Acid	0.94 (0.80-0.99)	0.86 (0.75-1.01)	-0.135	0.303	

	Control isolates	Case isolates	Spearman		
	Median (IQR)	Median (IQR)	*rho*	p-value	Sig
Glyoxylic Acid	1.37 (0.98-1.76)	1.39 (1.07-1.66)	0.223	0.087	
D-Cellobiose	1.29 (1.16-1.86)	1.28 (1.19-1.90)	-0.044	0.739	
Inosine	4.26 (3.37-5.42)	3.95 (3.33-5.46)	-0.108	0.411	
Glycyl-L-Glutamic Acid	2.45 (1.76-2.87)	2.28 (2.20-2.83)	0.137	0.296	
Tricarballylic Acid	3.41 (2.48-4.07)	3.39 (2.85-3.85)	0.204	0.118	
L-Serine	3.40 (2.84-4.32)	3.44 (2.96-3.97)	0.102	0.436	
L-Threonine	1.34 (1.02-1.76)	1.21 (1.04-1.77)	-0.118	0.370	
L-Alanine	2.44 (1.56-3.08)	2.63 (2.30-3.00)	0.345	<0.05	
L-Alanyl-Glycine	2.67 (2.02-3.32)	2.64 (2.41-3.36)	0.283	<0.05	
Acetoacetic Acid	1.00 (0.84-1.19)	0.99 (0.88-1.14)	-0.125	0.340	
N-Acetyl-ß-D-Mannosamine	3.27 (2.73-3.91)	3.01 (2.72-3.47)	0.056	0.672	
Mono Methyl Succinate	1.10 (0.93-1.33)	1.09 (0.86-1.36)	-0.390	<0.01	
Methyl Pyruvate	3.21 (2.67-4.01)	2.89 (2.53-3.73)	-0.171	0.192	
D-Malic Acid	0.81 (0.63-0.96)	0.67 (0.55-0.86)	-0.523	<0.001	+
L-Malic Acid	2.87 (2.24-3.56)	2.71 (2.46-3.44)	0.076	0.562	
Glycyl-L-Proline	2.60 (1.88-3.09)	2.56 (2.24-2.91)	0.314	<0.05	
p-Hydroxy Phenyl Acetic Acid	3.02 (2.31-3.77)	2.92 (2.64-3.47)	0.146	0.266	
m-Hydroxy Phenyl Acetic Acid	3.01 (2.36-3.73)	2.98 (2.65-3.51)	0.182	0.164	
Tyramine	2.76 (1.31-3.58)	2.95 (2.62-3.35)	0.375	<0.01	
D-Psicose	1.51 (1.27-1.80)	1.47 (1.31-1.79)	0.028	0.829	
L-Lyxose	0.78 (0.70-0.98)	0.75 (0.62-0.88)	-0.347	<0.05	
Glucuronamide	1.01 (0.76-1.10)	0.82 (0.68-0.97)	-0.518	<0.001	+
Pyruvic Acid	3.64 (2.85-4.59)	3.46 (2.95-4.31)	0.025	0.848	
L-Galactonic Acid-γ-Lactone	0.92 (0.69-1.01)	0.72 (0.65-0.94)	-0.485	<0.001	+
D-Galacturonic Acid	0.99 (0.79-1.07)	0.85 (0.75-1.03)	-0.427	<0.01	
Phenylethylamine	0.77 (0.64-0.95)	0.69 (0.59-0.81)	-0.482	<0.001	+
2-Aminoethanol	0.84 (0.64-0.98)	0.71 (0.56-0.89)	-0.534	<0.001	+
PM3 – Nitrogen sources					
Ammonia	1.00 (0.98-1.12)	1.00 (0.98-1.72)	-0.044	0.738	
Nitrite	0.88 (0.85-0.96)	0.87 (0.80-0.92)	-0.367	<0.01	
Nitrate	0.96 (0.93-0.99)	0.94 (0.87-0.99)	-0.072	0.584	
Urea	0.95 (0.92-1.00)	0.93 (0.85-0.98)	-0.126	0.337	
Biuret	0.98 (0.93-1.02)	0.95 (0.85-1.00)	-0.023	0.862	
L-Alanine	2.77 (1.02-3.91)	3.43 (2.13-3.84)	0.639	<0.001	+
L-Arginine	1.83 (1.10-3.40)	2.94 (1.72-3.58)	0.649	<0.001	+
L-Asparagine	2.54 (1.01-4.26)	3.71 (2.04-4.18)	0.607	<0.001	+
L-Aspartic Acid	1.47 (1.01-3.77)	2.74 (1.65-3.96)	0.570	<0.001	+
L-Cysteine	6.56 (1.07-8.18)	6.69 (3.09-8.41)	0.541	<0.001	+
L-Glutamic Acid	1.94 (1.12-3.03)	2.36 (1.58-3.20)	0.467	<0.001	+
L-Glutamine	3.19 (1.17-4.34)	3.97 (2.73-4.72)	0.632	<0.001	+
Glycine	2.20 (1.04-3.56)	3.31 (1.97-3.87)	0.682	<0.001	+
L-Histidine	2.83 (1.09-3.92)	3.43 (2.07-4.46)	0.595	<0.001	+
L-Isoleucine	0.99 (0.97-1.01)	1.00 (0.91-1.04)	0.187	0.152	

	Control isolates	Case isolates	Spearman		
	Median (IQR)	Median (IQR)	*rho*	p-value	Sig
L-Leucine	0.99 (0.96-1.07)	1.01 (0.91-1.10)	0.100	0.447	
L-Lysine	1.05 (1.01-1.14)	1.14 (1.01-1.44)	0.657	<0.001	+
L-Methionine	1.01 (0.97-1.07)	1.02 (0.98-1.12)	0.430	<0.01	
L-Phenylalanine	1.08 (1.00-1.47)	1.25 (1.05-1.61)	0.428	<0.01	
L-Proline	4.23 (1.22-5.30)	4.40 (3.04-5.66)	0.648	<0.001	+
L-Serine	2.98 (1.00-4.52)	3.91 (2.02-4.78)	0.646	<0.001	+
L-Threonine	1.23 (1.04-1.98)	1.40 (1.09-2.17)	0.527	<0.001	+
L-Tryptophan	1.01 (0.95-1.05)	0.98 (0.90-1.03)	-0.103	0.434	
L-Tyrosine	1.08 (1.04-1.14)	1.03 (0.96-1.09)	-0.118	0.369	
L-Valine	1.00 (0.98-1.03)	0.98 (0.92-1.02)	0.204	0.118	
D-Alanine	2.84 (1.22-3.91)	3.66 (1.80-4.54)	0.689	<0.001	+
D-Asparagine	1.05 (1.00-1.18)	1.08 (1.02-1.58)	0.323	<0.05	
D-Aspartic Acid	1.08 (1.04-1.36)	1.09 (1.04-1.33)	0.342	<0.05	
D-Glutamic Acid	0.86 (0.79-0.93)	0.81 (0.70-0.89)	-0.332	<0.05	
D-Lysine	1.09 (1.02-1.15)	1.07 (1.00-1.14)	0.323	<0.05	
D-Serine	4.07 (1.51-4.71)	4.21 (3.52-5.08)	0.657	<0.001	+
D-Valine	1.09 (1.01-1.16)	1.16 (1.00-1.33)	0.508	<0.001	+
L-Citrulline	1.03 (0.99-1.14)	1.08 (0.99-1.26)	0.520	<0.001	+
L-Homoserine	0.85 (0.80-0.94)	0.83 (0.74-0.92)	-0.187	0.160	
L-Ornithine	1.11 (1.00-1.18)	1.14 (1.02-1.42)	0.435	<0.01	
N-Acetyl-D,L-Glutamic Acid	0.95 (0.92-1.01)	0.93 (0.85-1.01)	-0.139	0.289	
N-Phthaloyl-LGlutamic Acid	1.03 (0.94-1.12)	0.98 (0.85-1.06)	-0.013	0.920	
L-Pyroglutamic Acid	0.97 (0.93-1.00)	0.96 (0.86-1.01)	-0.132	0.317	
Hydroxylamine	0.80 (0.75-0.90)	0.79 (0.66-0.85)	-0.325	<0.05	
Methylamine	0.96 (0.91-1.02)	0.91 (0.81-1.00)	-0.271	<0.05	
N-Amylamine	1.00 (0.96-1.03)	0.98 (0.88-1.05)	-0.088	0.503	
N-Butylamine	1.03 (1.00-1.06)	1.04 (0.93-1.06)	-0.033	0.804	
Ethylamine	1.04 (1.02-1.09)	1.03 (0.97-1.09)	0.153	0.242	
Ethanolamine	1.03 (0.99-1.04)	1.01 (0.93-1.07)	0.118	0.368	
Ethylenediamine	0.98 (0.93-1.07)	0.96 (0.85-1.03)	-0.115	0.381	
Putrescine	1.01 (0.98-1.05)	1.01 (0.92-1.05)	0.140	0.285	
Agmatine	1.00 (0.96-1.03)	0.99 (0.89-1.03)	-0.023	0.863	
Histamine	1.04 (0.98-1.08)	0.99 (0.91-1.08)	-0.237	0.069	
ß-Phenylethyl-amine	0.97 (0.94-1.03)	0.96 (0.90-1.01)	-0.094	0.476	
Tyramine	3.34 (1.27-4.21)	3.76 (2.60-4.54)	0.653	<0.001	+
Acetamide	1.06 (1.03-1.11)	1.05 (1.01-1.09)	0.141	0.282	
Formamide	1.07 (1.03-1.10)	1.04 (1.01-1.09)	0.147	0.261	
Glucuronamide	1.71 (1.26-2.10)	1.66 (1.46-1.92)	0.333	<0.05	
D,L-Lactamide	1.07 (1.02-1.11)	1.05 (0.97-1.11)	0.030	0.821	
D-Glucosamine	4.77 (1.88-6.01)	5.20 (4.01-5.63)	0.553	<0.001	+
D-Galactosamine	0.95 (0.89-1.00)	0.95 (0.82-1.00)	-0.118	0.369	
D-Mannosamine	1.14 (1.07-1.28)	1.15 (1.03-1.31)	0.154	0.241	
N-Acetyl-D-Glucosamine	5.10 (2.10-6.45)	5.24 (4.10-6.38)	0.438	<0.001	

	Control isolates	Case isolates	Spearman		
	Median (IQR)	Median (IQR)	*rho*	p-value	Sig
N-Acetyl-D-Galactosamine	1.02 (0.96-1.07)	0.99 (0.89-1.05)	-0.119	0.365	
N-Acetyl-D-Mannosamine	2.22 (1.12-3.24)	2.72 (2.07-3.65)	0.559	<0.001	+
Adenine	1.15 (0.96-1.57)	1.35 (1.00-1.70)	0.307	<0.05	
Adenosine	6.26 (5.31-7.17)	6.09 (5.33-7.43)	0.449	<0.001	
Cytidine	6.22 (3.34-7.47)	6.51 (5.63-7.32)	0.539	<0.001	+
Cytosine	1.04 (1.01-1.10)	1.04 (0.99-1.08)	0.101	0.441	
Guanine	2.64 (2.02-3.58)	2.33 (1.93-3.26)	0.002	0.987	
Guanosine	1.27 (1.16-1.78)	1.29 (1.07-1.87)	-0.146	0.267	
Thymine	1.03 (0.99-1.11)	1.00 (0.92-1.12)	-0.040	0.762	
Thymidine	1.02 (0.95-1.09)	0.98 (0.88-1.11)	0.102	0.439	
Uracil	1.00 (0.96-1.02)	0.98 (0.88-1.04)	-0.141	0.282	
Uridine	1.04 (0.98-1.11)	1.03 (0.89-1.10)	0.006	0.963	
Inosine	1.05 (1.00-1.10)	1.03 (0.88-1.12)	-0.110	0.405	
Xanthine	0.98 (0.90-1.13)	0.98 (0.87-1.06)	-0.028	0.833	
Xanthosine	1.14 (1.02-1.42)	1.14 (1.03-1.41)	0.459	<0.001	
Uric Acid	1.48 (1.23-1.78)	1.40 (1.26-1.58)	0.077	0.561	
Alloxan	1.17 (1.07-1.25)	1.17 (1.01-1.20)	0.223	0.086	
Allantoin	1.01 (0.97-1.05)	1.01 (0.93-1.04)	0.208	0.111	
Parabanic Acid	1.03 (0.97-1.06)	1.03 (0.95-1.06)	0.223	0.087	
D,L-α-Amino-N-Butyric Acid	0.86 (0.78-0.93)	0.81 (0.73-0.87)	-0.253	0.051	
γ-Amino-N-Butyric Acid	1.04 (1.01-1.12)	1.04 (0.99-1.21)	0.436	<0.001	
ε-Amino-N-Caproic Acid	1.00 (0.98-1.04)	1.03 (0.96-1.07)	0.391	<0.01	
D,L-a-Amino-Caprylic Acid	1.44 (1.26-1.88)	1.43 (1.15-1.87)	-0.061	0.642	
δ-Amino-N-Valeric Acid	1.07 (0.99-1.32)	1.08 (1.00-1.26)	0.556	<0.001	+
α-Amino-N-Valeric Acid	0.86 (0.81-0.92)	0.84 (0.76-0.90)	-0.294	<0.05	
Ala-Asp	4.40 (1.30-5.41)	4.69 (3.29-5.63)	0.629	<0.001	+
Ala-Gln	4.93 (1.25-6.29)	5.15 (4.28-6.30)	0.578	<0.001	+
Ala-Glu	4.38 (1.37-5.56)	4.77 (3.81-5.93)	0.653	<0.001	+
Ala-Gly	3.45 (1.11-4.42)	4.02 (2.67-4.72)	0.672	<0.001	+
Ala-His	2.77 (1.03-3.61)	3.17 (2.14-4.06)	0.668	<0.001	+
Ala-Leu	2.88 (1.08-3.72)	3.07 (1.83-3.95)	0.651	<0.001	+
Ala-Thr	3.04 (1.07-3.60)	3.19 (2.25-4.13)	0.654	<0.001	+
Gly-Asn	4.16 (1.45-5.27)	4.61 (3.38-5.68)	0.629	<0.001	+
Gly-Gln	4.50 (1.43-5.77)	4.70 (3.42-5.94)	0.594	<0.001	+
Gly-Glu	1.57 (1.16-3.13)	2.28 (1.29-2.90)	0.493	<0.001	+
Gly-Met	1.64 (1.01-2.73)	2.35 (1.22-3.21)	0.650	<0.001	+
Met-Ala	2.07 (1.03-3.25)	2.50 (1.35-3.58)	0.645	<0.001	+
PM4 – Phosphorus sources					
Phosphate	1.13 (1.01-2.40)	2.10 (1.14-3.26)	0.412	<0.01	
Pyrophosphate	3.35 (2.35-4.03)	3.26 (2.74-4.20)	0.070	0.598	
Trimeta-phosphate	1.05 (1.00-2.13)	1.59 (1.15-2.89)	0.406	<0.01	
Tripoly-phosphate	1.04 (0.98-1.85)	1.40 (1.04-2.63)	0.417	<0.01	
Triethyl Phosphate	1.00 (0.99-1.04)	1.03 (1.01-1.08)	0.137	0.296	

	Control isolates	Case isolates	Spearman		
	Median (IQR)	Median (IQR)	*rho*	p-value	Sig
Hypophosphite	1.00 (0.97-1.02)	0.99 (0.96-1.01)	-0.275	<0.05	
Adenosine-2'-monophosphate	3.43 (1.14-4.80)	4.70 (3.85-5.41)	0.347	<0.05	
Adenosine-3'-monophosphate	5.38 (1.93-6.79)	6.30 (4.99-7.00)	0.484	<0.001	+
Adenosine-5'-monophosphate	5.01 (2.50-6.53)	5.25 (4.53-6.58)	0.390	<0.01	
Adenosine-2',3'-cyclic monophosphate	4.77 (1.33-6.02)	5.77 (4.48-6.34)	0.456	<0.001	
Adenosine-3',5'-cyclic monophosphate	1.17 (1.14-1.21)	1.20 (1.17-1.24)	0.253	0.051	
Thio-phosphate	1.37 (1.00-2.71)	2.15 (1.04-3.16)	0.552	<0.001	+
Dithio-phosphate	1.22 (1.06-2.36)	1.85 (1.08-2.87)	0.581	<0.001	+
D,L-α-Glycerol Phosphate	2.77 (1.12-4.04)	3.63 (3.15-4.80)	0.613	<0.001	+
ß-Glycerol Phosphate	1.77 (1.06-3.24)	2.95 (1.70-4.05)	0.587	<0.001	+
Carbamyl Phosphate	1.17 (1.07-2.34)	2.03 (1.20-2.98)	0.379	<0.01	
D-2-Phospho-Glyceric Acid	2.37 (1.25-3.32)	2.90 (1.74-3.73)	0.602	<0.001	+
D-3-Phospho-Glyceric Acid	3.04 (1.85-4.41)	3.99 (3.03-4.67)	0.489	<0.001	+
Guanosine-2'-monophosphate	2.46 (1.10-3.96)	3.72 (2.65-4.56)	0.455	<0.001	
Guanosine-3'-monophosphate	5.06 (1.55-6.75)	6.13 (4.87-7.15)	0.585	<0.001	+
Guanosine-5'-monophosphate	2.85 (1.12-5.61)	3.91 (2.73-5.63)	0.491	<0.001	+
Guanosine-2',3'-cyclic monophosphate	2.82 (1.24-4.75)	4.52 (2.58-5.35)	0.483	<0.001	+
Guanosine-3',5'-cyclic monophosphate	1.26 (1.20-1.30)	1.25 (1.22-1.29)	-0.055	0.675	
Phosphoenol Pyruvate	2.76 (1.40-4.27)	3.71 (2.37-4.50)	0.493	<0.001	+
Phospho-Glycolic Acid	1.20 (1.05-2.73)	2.38 (1.45-3.31)	0.411	<0.01	
D-Glucose-1-Phosphate	4.70 (1.32-5.60)	5.32 (4.30-5.98)	0.531	<0.001	+
D-Glucose-6-Phosphate	4.59 (1.89-5.62)	5.64 (4.13-6.04)	0.619	<0.001	+
2-Deoxy-D-Glucose 6-Phosphate	1.11 (1.08-1.14)	1.13 (1.11-1.20)	0.066	0.616	
D-Glucos-amine-6-Phosphate	3.61 (2.46-4.97)	4.59 (3.79-5.47)	0.630	<0.001	+
6-Phospho-Gluconic Acid	1.23 (1.11-1.50)	1.32 (1.26-1.77)	0.294	<0.05	
Cytidine-2'-monophosphate	5.43 (2.38-7.11)	6.22 (4.81-7.12)	0.591	<0.001	+
Cytidine-3'-monophosphate	1.14 (1.09-1.20)	1.17 (1.14-1.26)	0.319	<0.05	
Cytidine-5'-monophosphate	3.10 (1.14-4.65)	3.01 (1.72-4.45)	0.362	<0.01	
Cytidine-2',3'-cyclic monophosphate	4.33 (1.29-6.37)	5.51 (4.17-6.40)	0.596	<0.001	+
Cytidine-3',5'-cyclic monophosphate	1.23 (1.20-1.26)	1.26 (1.24-1.31)	0.338	<0.05	
D-Mannose-1-Phosphate	2.93 (1.08-5.15)	4.61 (2.35-5.57)	0.614	<0.001	+
D-Mannose-6-Phosphate	3.69 (1.58-4.82)	4.74 (3.81-5.37)	0.606	<0.001	+
Cysteamine-S-Phosphate	1.20 (1.06-2.71)	2.27 (1.34-3.52)	0.545	<0.001	+
Phospho-L-Arginine	1.62 (1.09-3.45)	2.98 (1.81-3.94)	0.532	<0.001	+
O-Phospho-D-Serine	1.15 (1.12-1.58)	1.43 (1.13-1.99)	0.440	<0.001	
O-Phospho-L-Serine	1.27 (1.16-1.72)	1.33 (1.24-2.24)	0.346	<0.05	
O-Phospho-L-Threonine	1.45 (1.17-1.79)	1.72 (1.35-2.10)	0.481	<0.001	+
Uridine-2'-monophosphate	2.87 (1.31-4.55)	4.02 (3.28-5.17)	0.500	<0.001	+
Uridine-3'-monophosphate	5.10 (1.92-6.63)	5.94 (4.56-6.52)	0.550	<0.001	+

	Control isolates	Case isolates	Spearman		
	Median (IQR)	Median (IQR)	*rho*	p-value	Sig
Uridine-5'-monophosphate	2.56 (1.12-4.68)	3.61 (2.67-4.87)	0.519	<0.001	+
Uridine-2',3'-cyclic monophosphate	5.20 (1.57-6.84)	6.11 (4.76-6.92)	0.574	<0.001	+
Uridine-3',5'-cyclic monophosphate	1.13 (1.09-1.15)	1.14 (1.12-1.18)	0.279	<0.05	
O-Phospho-D-Tyrosine	2.16 (1.07-3.50)	2.98 (1.91-3.90)	0.425	<0.01	
O-Phospho-L-Tyrosine	1.62 (1.06-3.37)	2.93 (2.12-3.69)	0.442	<0.001	
Phosphocreatine	1.21 (1.08-2.26)	2.07 (1.24-3.28)	0.399	<0.01	
Phosphoryl Choline	1.13 (1.08-1.22)	1.18 (1.14-1.33)	0.266	<0.05	
O-Phosphoryl-Ethanolamine	1.23 (1.14-1.31)	1.27 (1.22-1.46)	0.385	<0.01	
Phosphono Acetic Acid	1.24 (1.17-1.29)	1.23 (1.19-1.29)	0.009	0.946	
2-Aminoethyl Phosphonic Acid	1.43 (1.25-2.28)	2.25 (1.40-2.84)	0.416	<0.01	
Methylene Diphosphonic Acid	1.70 (1.65-1.92)	1.87 (1.67-2.02)	0.013	0.920	
Thymidine-3'-monophosphate	2.03 (1.19-3.70)	3.74 (2.83-4.36)	0.487	<0.001	+
Thymidine-5'-monophosphate	2.84 (1.33-5.03)	3.92 (3.35-5.03)	0.537	<0.001	+
Inositol Hexaphosphate	2.35 (1.25-3.63)	2.74 (1.25-3.32)	0.042	0.752	
Thymidine 3',5'-cyclic monophosphate	1.20 (1.17-1.25)	1.21 (1.19-1.25)	0.130	0.323	
PM4 – Sulfur sources					
Sulfate	1.05 (1.01-1.12)	1.05 (1.02-1.27)	0.293	<0.05	
Thiosulfate	1.11 (1.01-1.53)	1.36 (1.11-1.52)	0.459	<0.001	
Tetrathionate	1.07 (1.02-1.25)	1.20 (1.05-1.50)	0.360	<0.01	
Thiophosphate	1.21 (1.02-1.79)	1.38 (1.15-1.76)	0.547	<0.001	+
Dithiophosphate	1.16 (1.01-1.70)	1.44 (1.08-1.70)	0.506	<0.001	+
L-Cysteine	1.11 (1.04-1.38)	1.26 (1.05-1.40)	0.343	<0.05	
D-Cysteine	1.33 (1.04-1.81)	1.59 (1.25-1.74)	0.597	<0.001	+
L-Cysteinyl-Glycine	1.17 (1.05-1.61)	1.41 (1.26-1.59)	0.590	<0.001	+
L-Cysteic Acid	1.08 (1.02-1.28)	1.10 (1.02-1.26)	0.389	<0.01	
Cysteamine	1.07 (1.01-1.11)	1.05 (1.03-1.11)	0.302	<0.05	
L-Cysteine Sulfinic Acid	1.53 (1.09-2.03)	1.60 (1.25-1.82)	0.571	<0.001	+
N-Acetyl-L-Cysteine	1.01 (0.91-1.04)	0.98 (0.93-1.01)	-0.496	<0.001	+
S-Methyl-L-Cysteine	1.01 (0.95-1.04)	1.00 (0.93-1.06)	-0.276	<0.05	
Cystathionine	1.22 (1.07-1.48)	1.42 (1.18-1.59)	0.622	<0.001	+
Lanthionine	1.50 (1.08-1.85)	1.60 (1.36-1.85)	0.588	<0.001	+
Glutathione	1.32 (1.11-1.82)	1.60 (1.25-1.78)	0.552	<0.001	+
D,L-Ethionine	0.97 (0.62-1.08)	0.82 (0.57-1.03)	-0.511	<0.001	+
L-Methionine	1.34 (1.12-1.57)	1.51 (1.32-1.66)	0.496	<0.001	+
D-Methionine	1.41 (1.11-1.81)	1.69 (1.34-1.96)	0.621	<0.001	+
Glycyl-L-Methionine	1.44 (1.14-1.85)	1.60 (1.46-2.00)	0.591	<0.001	+
N-Acetyl-D,L-Methionine	1.29 (1.11-1.61)	1.54 (1.35-1.69)	0.568	<0.001	+
L-Methionine Sulfoxide	1.33 (1.07-1.61)	1.45 (1.28-1.71)	0.572	<0.001	+
L-Methionine Sulfone	0.99 (0.63-1.12)	0.80 (0.50-1.02)	-0.464	<0.001	
L-Djenkolic Acid	1.40 (1.07-1.60)	1.45 (1.16-1.66)	0.557	<0.001	+
Thiourea	1.08 (1.05-1.23)	1.14 (1.08-1.27)	0.508	<0.001	+
1-Thio-ß-D-Glucose	1.08 (1.04-1.12)	1.07 (1.04-1.12)	-0.302	<0.05	

	Control isolates	Case isolates	Spearman		
	Median (IQR)	Median (IQR)	*rho*	p-value	Sig
D,L-Lipoamide	1.08 (1.04-1.12)	1.09 (1.03-1.12)	0.151	0.249	
Taurocholic Acid	1.09 (1.04-1.15)	1.10 (1.05-1.16)	-0.130	0.324	
Taurine	1.04 (0.99-1.09)	1.01 (0.94-1.07)	-0.223	0.087	
Hypotaurine	1.07 (1.04-1.10)	1.03 (0.97-1.09)	-0.107	0.417	
p-Amino Benzene Sulfonic Acid	1.09 (1.04-1.13)	1.08 (0.99-1.12)	-0.298	<0.05	
Butane Sulfonic Acid	1.06 (0.97-1.15)	1.04 (0.93-1.10)	-0.295	<0.05	
2-Hydroxyethane Sulfonic Acid	1.07 (1.01-1.11)	1.03 (0.97-1.07)	-0.228	0.080	
Methane Sulfonic Acid	1.07 (1.01-1.10)	1.02 (0.97-1.07)	-0.212	0.105	
Tetra-methylene Sulfone	1.07 (0.95-1.14)	1.03 (0.96-1.09)	-0.344	<0.05	

Sig: significance after Bonferroni correction (empty cells: not significant, +: significant).

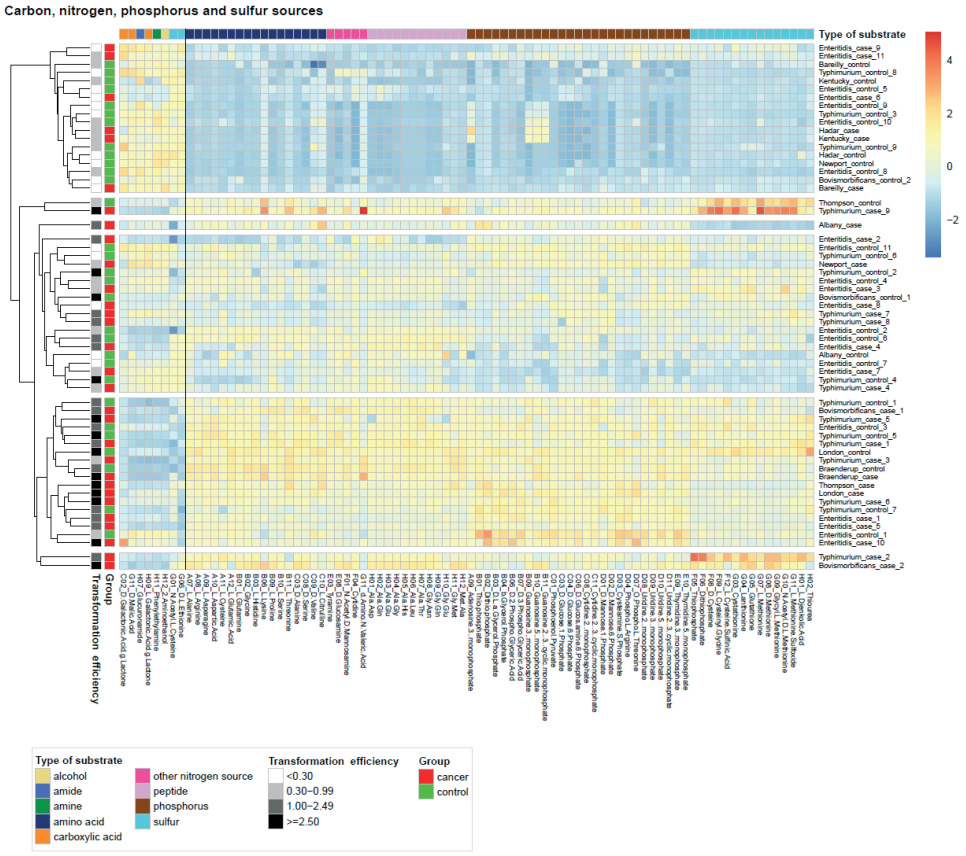


Figure S5. Heatmap of scaled utilization scores of the 60 nontyphoidal *Salmonella* (NTS) isolates for the sources with a significant positive or inverse correlation with transformation efficiency. Left eight columns depict the sources with inverse correlation (left of black line), right 76 columns depict the source with a positive correlation. NTS isolates are clustered based their utilization scores using the average linkage method.

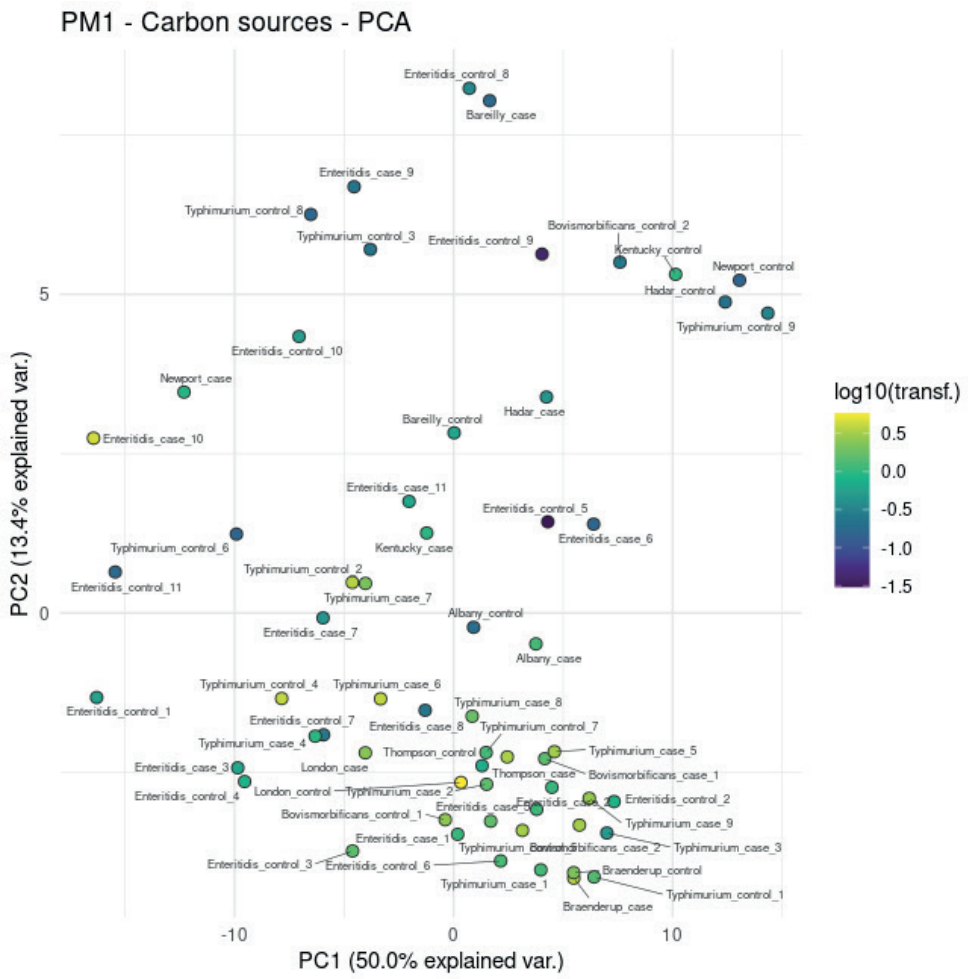


Figure S6. Principal component analysis (PCA) plot of the carbon source utilization of 60 nontyphoidal *Salmonella* isolates. The color scale of the dots reflects the log transformation efficiency. Principal component 1 (PC1) and PC2 collectively accounted for 63.4% of the variance in the nutrient utilization data. Transf: transformation efficiency.

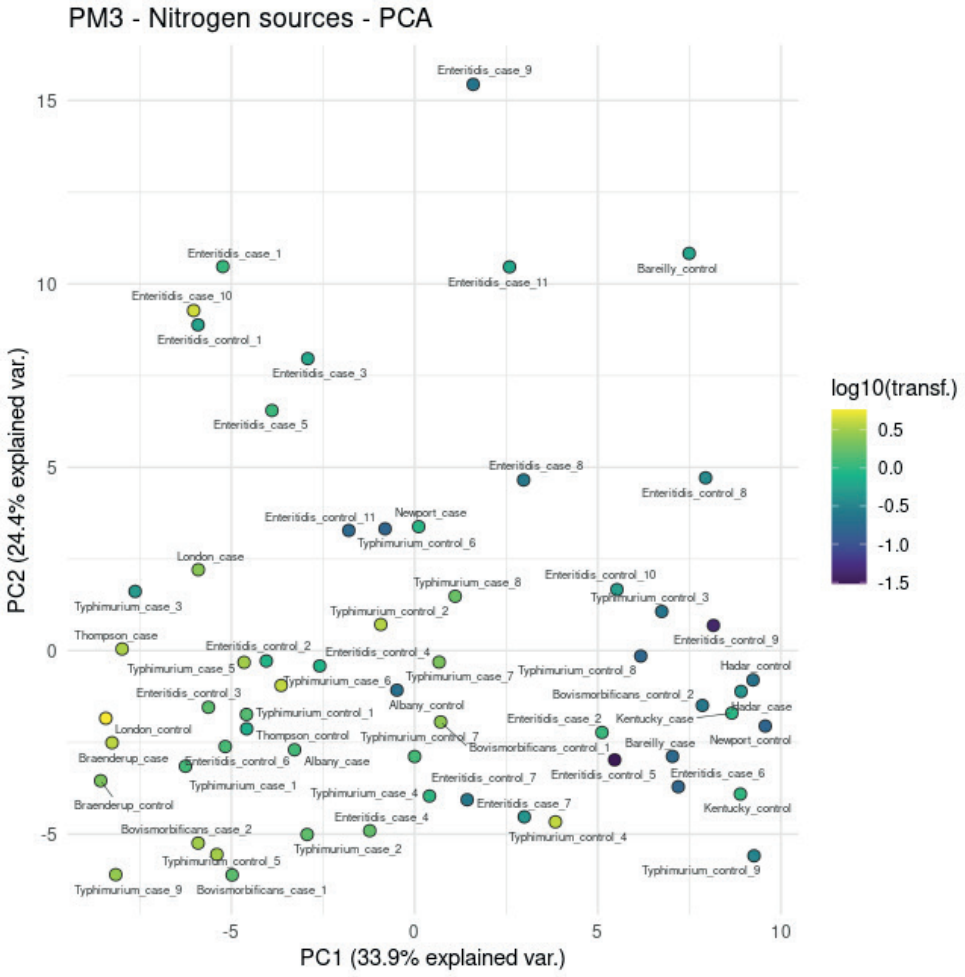


Figure S7. Principal component analysis (PCA) plot of the nitrogen source utilization of 60 nontyphoidal *Salmonella* isolates. The color scale of the dots reflects the log transformation efficiency. Principal component 1 (PC1) and PC2 collectively accounted for 58.3% of the variance in the nutrient utilization data. Transf.: transformation efficiency.

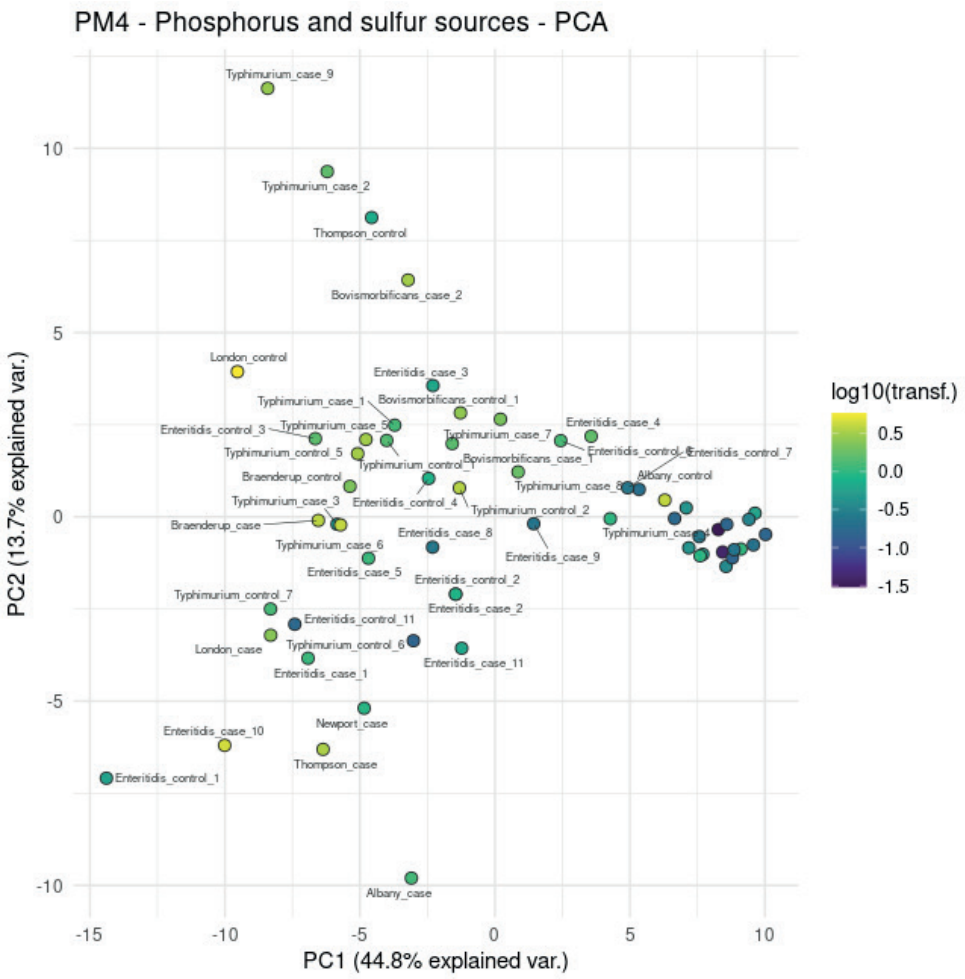


Figure S8. Principal component analysis (PCA) plot of the phosphorus and sulfur source utilization of 60 nontyphoidal *Salmonella* isolates. The color scale of the dots reflects the log transformation efficiency. Principal component 1 (PC1) and PC2 collectively accounted for 58.5% of the variance in the nutrient utilization data. Transf.: transformation efficiency.

Chapter 9

GENERAL DISCUSSION

Worldwide, the incidence of colon cancer ranks among the highest of all cancers with 1.1 million diagnoses annually [1]. Apart from hereditary causes of colon cancer, the main risk factors are related to lifestyle and dietary habits. However, the impact of microorganisms in the initiation and progression of cancers has been the subject of study for decades. Nowadays, over 20% of all cancers is estimated to be attributable to microorganisms [2], which therefore constitute a significant risk factor for a number of malignancies. While the earliest reports of an association between viral infections and cancer date back to the beginning of the 20th century, the oncogenic potential of bacteria has long been neglected. This might be related to the fact that bacteria, in contrast to viruses, do not leave a genetic imprint in human cells after infection, thereby hampering causal inference. Fortunately, the oncogenic potential of pathogenic and commensal bacteria is gradually being acknowledged, as an ever expanding list of bacteria are being associated with the onset and progression of cancer [3]. Few *in vitro* and epidemiological studies suggested a role of nontyphoidal *Salmonella* in the development of colon cancer. In this thesis, we aimed to explore the role of bacteria in the onset and progression of cancers in the gastrointestinal tract, with particular focus on the association between *Salmonella* and colon cancer. Broadly, the objectives of the thesis were covered by three main themes. First, we aimed to investigate whether repeated exposure to (lower doses of) *Salmonella* yields a similar risk of colon cancer development as compared to a single severe or high-dose infection. Second, we attempted to strengthen the evidence of an association between *Salmonella* and gastrointestinal tract cancers from an epidemiological perspective by conducting two registry-based studies and reviewing the current worldwide knowledge on these associations. This was done also for other correlations between bacteria/parasites and gastrointestinal tract cancers. The third theme covers the question whether the *Salmonella* phenotype in terms of its oncogenic potential can be explained by its genotype and which mechanisms might play a role in the *Salmonella*-induced tumorigenesis. In the following paragraphs, the main findings within these themes will be summarized and discussed. Also, we put the results from this thesis in a broader perspective of existing literature. This is followed by a paragraph elaborating the overall lessons learnt, as well as implications and recommendations for further study.

Experimental evidence for an association between repeated exposure to *Salmonella* and colon cancer is generally stronger than epidemiological data

Over a thousand culture-confirmed *Salmonella* infections are reported annually in the Netherlands, mostly representing salmonellosis patients with symptoms severe enough to require medical attention, including symptoms lasting at least 1-2 weeks, and travel-related infections [4]. In an earlier Dutch cohort study, an association was found between these reported *Salmonella* infections and proximal colon cancer, with an almost 3-fold increased risk of proximal colon cancer after infection with *S*. Enteritidis [5]. However, people acquire multiple *Salmonella* infections throughout life via consumption of contaminated food or water, contact with live animals, the environment and to a lesser extent human-to-human transmission [6, 7]. The vast majority of these infections are asymptomatic or present themselves with mild and self-limiting symptoms that do not require medical attention, thereby remaining undiagnosed and unreported. Whether the cumulative effect of multiple, a- or pauci-symptomatic infections also contributes to the initiation or progression of colon cancer, was unknown, and therefore became the subject of the studies in **Chapters 2-4**. Here, we approached this research question from both an experimental and an epidemiological angle.

In the experiments described in **Chapter 4**, mice were orally infected with either a single high dose or multiple low doses of a laboratory *S*. Typhimurium strain to assess the effect on cancer development. To this end, a frequently used mouse model, mimicking human colitis-associated colon cancer, was used in which the mice received the pro-carcinogen azoxymethane (AOM) and colitis inducing dextran sulfate sodium (DSS) to induce the formation of colon tumors. The group of mice infected with a single high dose received 10,000 colony forming units (CFU) of *Salmonella*, whereas the group receiving multiple low doses were three times infected with 10 CFU of *Salmonella* with an interval of 4 weeks. Both groups infected with *Salmonella* (high or low dose), as well as a control group that received the AOM+DSS without infection, developed colon tumors in the 16 weeks of follow-up. This was in contrast to control mice that were neither exposed to AOM+DSS nor received *Salmonella* infection, indicating that the AOM+DSS treatment was the main driver of tumor formation. Overall, no differences were observed in the number of tumors, tumor volume, number of proliferating cells or colonization with *Salmonella* between the mice that received a single high dose of infection and the mice that received multiple low doses. However, the tumors were significantly larger in the *Salmonella*-infected mice as compared to the AOM+DSS control group. In addition, the tissues in both groups of *Salmonella*-infected mice

showed high grade dysplasia and signs of invasive carcinoma, which was not observed in the AOM+DSS control group. Also, higher amounts of proliferating cells were observed in the *Salmonella*-infected mice in both the tumor tissue, as well as the adjacent tissue, whereas the degree of colonization was higher in the tumor tissues as compared to adjacent tissue. The observation that low doses of *Salmonella* are sufficient to induce an (oncogenic) effect are in line with earlier experiments in rats where a low dose of *S*. Enteritidis caused colonic lesions, even in the absence of symptoms of illness [8]. Next to the mice experiments, the effects of a low- or high-dose infection were examined in a simple cell model consisting of mouse embryonic fibroblasts (MEFs) with a tumorigenic predisposition inflicted by overexpression of the proto-oncogene c-MYC and inactivation of tumor suppressor gene TP53 (**Chapter 4**). Exposure of the MEFs to either a high (multiplicity of infection [MOI] 25) or low (MOI 5) dose of a laboratory *S.* Typhimurium strain led to transformation, as characterized by colony formation in soft agar assays. Culturing of these transformed MEFs (i.e. after the first *Salmonella* infection) and subsequent reinfection with the same high or low dose of *Salmonella* resulted in the formation of more and larger colonies. This was particularly true for the MEFs infected with a high dose of *Salmonella.* Still, the number and size of colonies was larger after reinfection of MEFs with a low dose as compared to a single infection with a high dose. Similarly, *Salmonella* showed a tropism for MEFs with the highest level of transformation. This is in line with the *in vivo* experiment, as well as prior studies in which *Salmonella* showed preferential accumulation and proliferation in tumor tissue as compared to normal tissue at a ratio of >1000:1 [9]. Hence, both *in vitro* and *in vivo* experiments showed that infections with a low dose of *Salmonella* trigger a similar tumorigenic effect as a single high-dose infection, though in MEFs a high-dose infection positively impacts the number and size of the colonies formed. All in all, the experimental evidence for an association between *Salmonella* infection and colon cancer obtained in **Chapter 4** was substantial and generally indicated that, under certain conditions that may occur realistically in nature, *Salmonella* is able to promote colon carcinogenesis.

To complement these experimental findings, we also conducted several epidemiological studies. First, we investigated whether the presumed higher levels of (repeated) exposure to *Salmonella* are reflected by a higher incidence of salmonellosis. To this end, we compared the incidence of reported salmonellosis among different occupational groups in a large nationwide population-based registry study (**Chapter 2**). We also included reported *Campylobacter* infections in this study, as this pathogen is a leading cause of gastroenteritis as well, albeit scarcely associated with cancer. We thus studied the incidence of salmonellosis and campylobacteriosis among the whole spectrum of occupations grouped into 'divisions'

according to an internationally agreed upon occupational classification system. Moreover, we defined three risk groups or divisions with a presumedly higher degree of occupational exposure to zoonotic pathogens, including *Salmonella*, due to contact with animals and products thereof. As anticipated, a significantly higher incidence of salmonellosis (1.8-fold) and campylobacteriosis (1.7-fold) was observed among people working with live animals or animal manure (e.g. farmers, veterinarians, abattoir workers) as compared to the incidence in the total employed population. Among people involved in the sale of animal-derived products (e.g. butcher's and cheese shops), the reported incidences were also significantly higher for both salmonellosis (1.6-fold) and campylobacteriosis (1.4-fold), whereas the incidences among people involved in the processing of foods of animal origin (e.g. cooks and chefs) were not significantly different from those in the total employed population. These results agree to a major extent with a previous Dutch study based on a combined analysis of microbial subtyping and epidemiological (case-control) data on human *Salmonella* infections in which occupational exposure to raw meat or animals was found to be associated with an over 6-fold increased incidence of cattle-borne human salmonellosis [6]. Excess cases of salmonellosis and campylobacteriosis (as compared to the total employed population) were observed among health care associated occupations, as well as some industrial divisions. The observed differences in incidence of reported infections might reflect a genuinely higher level of exposure to *Salmonella* and/or *Campylobacter*. However, the analysis based on reported infections only includes symptomatic cases of disease and is affected by one's propensity to seek medical care. Inequalities in the utilization of medical care are relatively minor in the Netherlands, though several studies found a significant higher probability of visiting a general practitioner (GP) among people with a lower socio-economic status (SES) as compared to those with a higher SES, also after correction for differences in overall health status between the groups [10-12]. Hence, differences in GP visitation rates between occupational groups are inevitable. Apart from the virulence determinants of the specific bacterial strain affecting its pathogenicity, the probability of developing severe gastroenteritis (i.e. requiring medical attention) also incorporates host-dependent factors related to susceptibility to infection and the ability to clear the infection before severe or long-term complaints appear. Exposure to pollutants such as chemicals, heavy metals and nanoparticles, as observed in some industrial occupations, is suggested to alter the composition of the microbiome, potentially rendering people more susceptible to (severe) *Salmonella* or *Campylobacter* infection [13]. As the propensity to use medical care and an individual's likelihood to develop a severe infection are only partially related to the extent, type and frequency of exposure to the pathogen, we analyzed other types of data, namely serological data, to compare the sero-incidence of *Salmonella* and *Campylobacter* in

a subset of the employed population (**Chapter 2**). The anti-*Salmonella*/-*Campylobacter* IgM, IgG and IgA antibody titers in an individual's serum sample served as input for calculating the time since last seroconversion (i.e. time since last exposure to the pathogen) using an established Bayesian back-calculation model [14]. The time since last seroconversion can subsequently be translated into an estimated average number of infections per person-year, i.e. the seroincidence [15]. Since the seroincidence is based on an antibody response and does not discriminate between symptomatic and asymptomatic disease, it constitutes a less biased measure of infection pressure. In our study, we observed minor variation in the seroincidences of *Salmonella* and *Campylobacter* in different occupational groups, albeit that the seroincidences were assessed at a lower hierarchical level of occupational coding due to sample size constraints. Also, the seroincidences among people with an occupation included in one of the three risk groups did not differ from the average seroincidences in the whole subset. Such a disconcordance between reported infections and seroincidence has been observed in several European countries [16, 17]. Registered type of occupation is just one proxy for the level of exposure to *Salmonella*. Yet, this covers only a part of the myriad of possible exposures in an individual's life. Consumption of raw or undercooked meat and eggs, ownership of companion animals, use of proton pump inhibitors and poor kitchen hygiene practices leading to cross-contamination have been found to be the main risk factors for salmonellosis in the Netherlands, while occupational exposure to live animals or meat was only a risk factor for cattle-associated *Salmonella* infection, with a relatively minor contribution to the total set of risk factors [6]. Moreover, other studies have found higher incidences of salmonellosis and campylobacteriosis among people living in close proximity to broiler farms or areas with a high density of dairy/cattle farms [18, 19]. Our study clearly indicated an increased risk of salmonellosis (and campylobacteriosis) in people occupationally exposed to 'the source' of the pathogen, i.e. livestock and products thereof, but the relation between occupation and frequency of exposure, as depicted by serology, did not help much disentangling further the deeper mechanistic process.

We then explored the incidence of colon cancer among different professions in order to assess whether the occupational groups with increased salmonellosis incidence were also those more prone to develop cancer. To this end, in **Chapter 3** we presented the results of another nationwide occupational registry-based study with a comparable design and occupational (risk) group classification as in the study presented in **Chapter 2**. Overall, the differences in colon cancer incidence among occupational divisions, including the pre-defined risk groups, was relatively minor. Indeed, no excess risk of colon cancer was observed among people with a history of employment in the three risk groups as compared to the general

population, despite significantly higher incidences of salmonellosis. Significantly increased risks of up to 45% were observed for some divisions including mainly 'blue-collar' and manual occupations (e.g. occupations involved in printing, manufacturing of rubber, plastics, machinery and equipment and the sale of motorcycles/-vehicles), as well as some divisions pertaining mainly to the 'white-collar' and non-manual occupations (e.g. information service and real estate activities, and (re)insurance and pension funding). Yet, the significantly higher incidence of colon cancer among people with (a history of) employment in the real estate and (re)insurance and pension funding divisions corresponded to the *Salmonella* seroincidence results, as the seroincidence among people working in these divisions was amongst the highest ones. In general, however, the study presented in **Chapter 3** helped us ascertain that occupation in itself provides little differences in colon cancer incidence and that a direct link between (occupational) exposure to *Salmonella* and (increased) colon cancer risk was not possible to make with this approach.

To overcome the limitations of using occupation as a proxy for *Salmonella* infection pressure, we then studied whether colon cancer risk is correlated to increased exposure to *Salmonella* earlier in life, regardless of occupation but still using seroincidence data (**Chapter 4**). By linking individual-level data from a Dutch serosurvey to colon cancer diagnosis data, we found 36 individuals who provided a serum sample in the survey and were diagnosed with proximal colon cancer ≥1 year later (hereafter referred to as 'cases'). In a matched case-control analysis, we then observed that, overall, *Salmonella* seroincidence did not differ significantly between cases and controls. However, upon stratification, we noted that *Salmonella* seroincidence was significantly higher among cases younger than 60 years at the time of the serosurvey as compared to controls. Similarly, a higher seroincidence was observed among cases living in a neighborhood with a high socio-economic status at time of the serosurvey. These findings suggested that the effect of (increased exposure to) *Salmonella* on colon cancer risk would be (epidemiologically) appreciable only in absence of other (and generally much stronger) known risk factors for colon cancer, such as older age and unhealthy lifestyle (which low socio-economic is a proxy for), which might thus mask the relatively smaller effect of *Salmonella* itself. A proximal colon cancer diagnosis preceded by a higher seroincidence before the age of 60 is in agreement with the results of an earlier Dutch cohort study where salmonellosis reported between 20-60 years of age was associated with a significant higher risk of proximal cancer [5].

Several factors might explain the absence of a correlation between the incidence of reported salmonellosis and colon cancer among occupational groups. The most predominant one is the fact that colon cancer is, to a large extent, attributable to lifestyle and dietary habits.

Despite substantial variations in literature concerning the magnitude of risks and population attributable fractions (PAFs), a number of well-established risk factors account for a significant portion of colon and colorectal cancer diagnoses. These include the consumption of red and processed meat (PAF 5-12%), a low intake of dietary fibers (PAF 16-18%), alcohol consumption (PAF 1-13%), low calcium intake (PAF 6-10%), excess body weight (PAF 1-17%), low levels of physical activity (PAF 5-18%) and tobacco use (9-12%) [20-23]. Although the risk of colon cancer due to (repeated) exposure to *Salmonella* could be partially contained in the PAF for the consumption of red and processed meat, this would be limited to a very minor contribution, given that a larger portion of human *Salmonella* infections were attributable to layers, eggs and broilers as compared to pigs and cattle in the study period [6]. Of note, during the last decade, the sources of human *Salmonella* infections have been changing gradually, with pigs being nowadays the main source of human infections and eggs decreasing in importance as attributable source relative to pigs [4]. Adherence to a healthy lifestyle (i.e. the absence of the aforementioned major risk factors for colon cancer) differs strongly between occupational groups. This is partially for reasons inherent to specific jobs, such as a more sedentary lifestyle for non-manual workers (i.e. occupations at a more managerial or administrative level) and exposure to chemical substances in some manual occupations [24, 25]. An example is the relatively low incidence of colon cancer observed in the risk group involving contact with live animals in **Chapter 3**, which is consistent with earlier research reporting a reduced risk of colon and colorectal cancer among farmers, presumably owing to high levels of occupational physical activity [26, 27]. On the other hand, adherence to a healthy lifestyle is to a certain extent correlated to socio-economic status unrelated to the content of the occupation itself [28, 29]. Examples hereof are an average higher intake of red/processed meat, lower levels of leisure time physical activity and higher smoking rates among people with a lower socio-economic status [30-32]. Yet, whether socio-economic status can serve as a proxy for colon cancer risk remains arguable since inconsistent results are documented in literature [33].

Both the *in vitro* and *in vivo* experiments suggested that multiple low doses of *Salmonella* infection suffice to induce a tumorigenic effect in predisposed cells/organisms. Yet, extrapolating these experimental conditions to human infections under natural conditions raises some difficulties. First, the seroincidence and the reported salmonellosis as used in the registry-based studies provide an estimation of the annual number of infections and roughly the severity and longevity of disease symptoms (requiring medical care) but lack information about the dose of the infection. Literature about the dose-response relationship for *Salmonella* in the human population is relatively scarce. Of note, the implementation

of *Salmonella* dose-response studies during foodborne outbreaks is often subject to data availability constraints, such as lack of the contaminated product consumed by the time that people visited the GP for persistent symptoms of gastroenteritis. Estimating the number of ingested bacteria is therefore practically impossible. Moreover, the size of the population exposed to the contaminated product is often unknown, as people with an asymptomatic or paucisymptomatic infection remain unidentified [8]. *Salmonella* dose-response models of outbreak data showed that the average ingested dose of bacteria that corresponds to an infection- or illness-probability of 50% (ID50) was 7 CFUs for infection and 36 CFUs for illness [34]. However, the dose-response relation strongly depends on the virulence profile of the bacterial strain, as well as on the hosts' susceptibility and immune response. Hence, the dose of infection cannot be deduced from the course of disease and a simple threshold for a low- *versus* high-dose infection is hard to establish in humans. The dose of infection can also not be estimated from the seroincidence, although the dynamics of the serological response in terms of the time until the peak of antibody production, the maximum concentrations of antibodies and the shape of the decay curve are to an unknown extent affected by the dose, as well as the serovar. Under experimental conditions, the production of IgG antibodies against *S*. Enteritidis was positively correlated with dose of infection in rats [35], whilst infection of piglets with different *S*. Typhimurium strains led to differences in onset of seroconversion [36]. Therefore, the epidemiological analyses in this thesis did not allow for disentanglement of low- *versus* high-dose infections, nor the actual frequency/degree of exposure to *Salmonella* in all groups of the population under study, nor the definition of a mild *versus* severe infection. However, we consider the contribution of *Salmonella* to colon cancer development to be the product of several factors. Ultimately, the probability that a predisposed or pre-transformed cell is infected by a *Salmonella* bacterium leading to replication of mutated cells and the development of a tumor is determined by a) the frequency of infection, b) the bacterial load as defined by the number of bacteria ingested, the number of bacteria surviving the gastric acid, bacterial replication and killing, as well as c) the duration of *Salmonella* bacteria being present in the colon before being eliminated by the hosts' immune system [37].

Epidemiological findings from other cohorts and malignancies confirm only small effects of *Salmonella* infection

As the epidemiological evidence of a possible association between severe (nontyphoidal) *Salmonella* infection and colon cancer was based on a single registry study in a Dutch cohort [5], we aimed to substantiate these results using an independent but comparable cohort. To this end, we performed a cohort study based on linked registries from the Danish population (**Chapter 5**). The design of the study resembled the Dutch cohort study with the exception that the Danish *Salmonella* surveillance system covers virtually the whole population (in contrast to the ~64% population coverage of the Dutch surveillance system), allowing for comparison of colon cancer incidence in individuals with a history of reported culture-confirmed salmonellosis *versus* individuals without such reported infection. Cox regression revealed no overall increased risk of colon cancer (both proximal and distal) ≥1 year after *Salmonella* infection. An 1.4-fold significant risk of proximal colon cancer was observed among people with a history of infection with a serovar other than Enteritidis and Typhimurium (including its monophasic variant), which are the most frequently occurring serovars among human salmonellosis cases. Besides, the hazard ratios (HRs) for proximal colon cancer in subgroup analyses were consistently higher for the group infected with other serovars as compared to Enteritidis and Typhimurium. The incidence of distal colon cancer diagnosis before the age of 50 after infection with a non-Enteritidis/Typhimurium infection was over three-fold higher. Remarkably, the incidence of diagnosed colon cancer was over two-fold increased among individuals with a recent (i.e. <1 year interval) *Salmonella* infection as compared to those without reported infection. While a causal relationship in terms of initiation of tumor development is implausible for reported *Salmonella* infections less than one year before cancer diagnosis, we cannot rule out that *Salmonella* infection in the early (prediagnostic) phase of colon/colorectal cancer leads to enhanced tumor progression. The tropism for infecting transformed MEFs and the higher rates of colonization in murine tumor tissue compared to adjacent tissue, as observed in **Chapter 4**, support this hypothesis. As the Dutch and Danish cohort used a comparable study design, we meta-analyzed the main outcomes of both studies. Analysis of heterogeneity of the two studies showed a low between-study variation ($p=0.958$; $I^2=0.01\%$). A priori, we decided to use a random-effects model when significant heterogeneity would be observed ($p<0.1$ or $I^2>50\%$), whereas a fixed-effects model would be used if $p>0.1$ or $I^2<50$. Hence, the fixed-effect model was used to obtain a pooled study estimate. An overall pooled risk ratio (RR) of 0.81 (95%CI 0.73-0.89) was obtained with a higher weight assigned to the Danish study (74.1%) compared to the Dutch study (25.9%) (Table 1).

Table 1. Risk ratios of colon cancer after Salmonella infection in the Dutch and Danish cohort studies.

	Reported *Salmonella* infection		**No reported *Salmonella* infection**		
	Colon cancer	**No cancer**	**Colon cancer**	**No cancer**	**RR (95%CI)**
Netherlands§	96	14,168	140,458	16,619,770	0.803 (0.658-0.980)
Denmark	245	47,578	54,624	7,544,498	0.808 (0.719-0.909)
Pooled estimate					0.807 (0.729-0.893)

RR: risk ratio. § Data derived from Mughini-Gras et al. (2018) [5].

The lack of correspondence between the outcomes of the two studies is presumably attributable to several effect modifying or diluting factors not equally present in both cohorts. One such factor is the epidemiology of human salmonellosis. In both the Netherlands and Denmark, control programs have been implemented aiming to reduce the prevalence of *Salmonella* in animal production systems. In Denmark, the implementation of three major control programs in 1988, 1993 and 1997 targeting *Salmonella* reduction in broiler chickens, pigs/pork and laying hens respectively, have led to a marked decrease of human *Salmonella* infections and a shift in the distribution of serovars [38, 39]. During the study period (1994-2015) the total number of reported salmonellosis decreased from a median of 2100 infections annually in 1994-2005 to 1129 in 2006-2016. *S*. Enteritidis accounted for the largest reduction (-61.1% in 2006-2016 *vs.* 1994-2005), followed by (monophasic) *S*. Typhimurium (-36.3%) and other *Salmonella* serovars (-19.6%) [40-43]. The reduction of notified salmonellosis in the Netherlands was less pronounced during the study period of the Dutch cohort study (1999-2015). A median annual number of 1674 infections was reported in the first part of the study period (1999-2007) compared to 1264 in the second part (2008-2015). *S*. Enteritidis infections reduced by 49.1% (2008-2015 *vs* 1999-2007), whilst the reduction of infections with *S*. Typhimurium and other serovars was limited to 5.4% and 3.6% respectively. Considering that the reported infections are only the tip of the iceberg of the true burden of disease, changes in infection pressure of *Salmonella* over time are presumably not sufficiently comparable between the two countries. In 2009, the incidence of human salmonellosis based on reported infections was 38.5 per 100,000 inhabitants in Denmark *versus* 11.6 per 100,000 inhabitants in the Netherlands [44, 45]. However, data from serosurveys conducted in 2006-2007 showed a higher infection pressure in the Netherlands compared to Denmark, with respective seroincidences of 0.149 and 0.084 infections per person-year [16]. In addition to a dissimilar degree of exposure, both countries have a distinct distribution of serovars and strains/clones responsible for human

salmonellosis, partly driven by travel and import of food products. For instance, the portion of sporadic human *Salmonella* infections related to travel is considerably higher in Denmark compared to the Netherlands (30-40% *vs.* 12%) [39, 44, 46]. Both the extent of exposure to *Salmonella* as well as serovar distributions might have contributed to the differences in observed outcomes between the Dutch and Danish cohort studies. Besides distinct exposures, the differences may also have been influenced by factors directly affecting the outcome probability, such as consumption behavior, physical activity and body weight. In 2014, the percentage overweight in young adults (18-24 years) was 25% in Denmark *versus* 20% in the Netherlands, while an opposite pattern was observed among individuals aged ≥65 (48-57% in Denmark, 55-62% in the Netherlands) [47]. Likewise, the consumption of red and processed meat per capita in Denmark is about 1.5 times higher as compared to Dutch inhabitants (840 grams/week *vs.* 560 grams/week) [48, 49]. The consumption of red and processed meat can act as confounder in the two studies, given that in both Denmark and the Netherlands a substantial part of the *Salmonella* infections is attributable to pigs or pork and the heme iron in red meat can modify the microbiome in favor of *Salmonella* colonization whilst it also directly increases colon cancer risk [50].

Looking at other malignancies than colon cancer, while a significant body of literature shows an association between *S*. Typhi and gallbladder carcinoma, the association between nontyphoidal *Salmonella* infection and cancers in the biliary tract has barely been studied. In contrast to *S*. Typhi that causes a systemic disease and is frequently associated with chronic infection, nontyphoidal *Salmonella* generally causes a local inflammation of the intestine [51]. Hence, direct manipulation of the epithelial cells of the gallbladder or bile ducts by nontyphoidal *Salmonella* as part of the infectious cycle is implausible for the vast majority of infections. Nonetheless, *in vitro S*. Typhimurium infection of gallbladder organoids and MEFs induced cellular transformation and the formation of colonies in soft agar [52]. Whether this tumorigenic potential could also be observed at the population level in an epidemiological study was addressed in **Chapter 6**. Here we performed a registry-based cohort study in the Dutch population, investigating the risk of cancers of the biliary tract after reported nontyphoidal *Salmonella* or *Campylobacter* infection. Cancers in the biliary tract include the gallbladder, the proximal and distal bile ducts and the extrahepatic bile ducts. The study design resembled the designs of the Danish study (**Chapter 4**) and the earlier Dutch study assessing the risk of colon cancer after salmonellosis [5]. Biliary tract cancer was diagnosed in nine individuals with a history of reported salmonellosis and seven individuals with a history of reported campylobacteriosis. These low numbers are not surprising given that biliary tract cancer is a rare malignancy in the Netherlands, with only

~800 diagnoses annually. Although none of the outcomes reached the level of significance, a clear tendency towards increased biliary tract cancer risk was observed for people with a history of *Salmonella* infection (standardized incidence ratios [SIRs] ranging from 1.22-1.88), whilst no such tendency was observed for *Campylobacter* (SIRs ranging from 0.75-1.24). The small sample size did not allow us to analyze subgroups, such as splitting by serovar/species or age at infection. Yet, the observed higher risk of cancer after *Salmonella* infection has been corroborated by others in a Taiwanese cohort, in which a HR of 1.78 was observed for biliary tract cancer after reported salmonellosis [53]. Likewise, traces of *S*. Typhi and *S*. Typhimurium were found in 9/26 and 10/26 samples respectively of tumor and adjacent normal tissue [54].

In parallel with the start of the research activities of this thesis, we noticed a substantial increase in literature addressing the association between microorganisms and cancers owing to rapidly evolving sequencing techniques. As a consequence, the need for a summary of the current epidemiological efforts and knowledge on this topic increased as well. We therefore conducted a literature review summarizing current epidemiological reports for the association between cancer in the gastrointestinal tract and bacterial or parasitic infections (**Chapter 7**). Viruses and the bacterium *Helicobacter pylori* were not included in this review, as these have already been addressed in many studies and reviews before. The majority of the 158 included studies and abstracts, covering 10 different bacteria and three parasites, focused on colon/colorectal cancer. A total of seven publications were found that studied the association between nontyphoidal *Salmonella* and gastrointestinal cancer, including the ones that are part of this thesis. A registry study based on data from the Taiwanese population found no evidence of an increased risk of colorectal cancer after *Salmonella* infection. In contrast, two other studies revealed significant higher antibody levels against *Salmonella* flagellin in Dutch colorectal cancer patients *versus* healthy controls, and a higher abundance of *Salmonella* in colon tissue adjacent to the tumor as compared to tissue from healthy controls in the USA [50, 54].

Only a few years ago, the scientific community focused mostly on microorganisms inducing the formation of a malignancy, whereas nowadays there is accumulating evidence recognizing a role for a larger group of (predominantly) bacteria that aid the progression of tumor growth. This has been described in a so-called driver-passenger model for colorectal cancer in which bacterial species that trigger the initiation of cancer are considered 'drivers', whereas opportunistic bacteria benefitting from the altered intestinal (tumor micro) environment and facilitating tumor development are considered 'passengers' [55]. The driver-passenger model/theory demands some flexibility though, as many bacteria have

potentially carcinogenic capacities (e.g. stimulation/inhibition of an inflammatory response, production of toxins or effector proteins) leading to manipulation of cell signaling pathways and ultimately the growth and proliferation of malignant cells, while they are also enriched in advanced/late stage tumors. According to the model, Enterobacteriaceae (the family to which *Salmonella* belongs) are considered driver bacteria given their abundance in off-tumor tissue as compared to tumor tissue and their perceived resemblance to enterotoxigenic *Bacteroides fragilis* in terms of prolonged inflammation [55]. While this hypothesis is supported by the results in this thesis, as well as outcomes of prior research showing that *Salmonella* is able to induce transformation in naïve non-malignant cells, *Salmonella* also shows passenger behavior by exploiting the metabolic niche arisen in the tumor microenvironment for its survival an proliferation.

The oncogenic potential of nontyphoidal *Salmonella* is difficult to explain from both a genotypic and phenotypic perspective

In previous *in vivo* and *in vitro* research, different *Salmonella* strains were used to assess their capacity to induce tumorigenic changes [**Chapter 4**; 52; 56]. Attempts to infect mice with three different *S*. Typhimurium and three *S*. Enteritidis isolates originating from human clinical samples (obtained from the *Salmonella* laboratory surveillance system) resulted in over 50% mortality rates in the mice for two of the Typhimurium isolates and one Enteritidis isolate (**Chapter 4**, data not shown). Moreover, the estimated risk of colon cancer after infection with *S*. Enteritidis, *S*. Typhimurium and other serovars were not consistent among the epidemiological studies [**Chapter 5**; 5]. Hence the aim of **Chapter 8** was to unravel possible virulence factors and other mechanisms responsible for the oncogenic potential of *Salmonella* and to examine whether colon tumors from patients with a reported *Salmonella* infection are different in nature to tumors from patients without a reported infection. With regard to the tumor characteristics, no significant differences were present between the two groups in terms of tumor markers, although the percentage of poorly differentiated tumors was lower in patients with a past reported *Salmonella* infection. For the other analyses, we used a matched case-control design based on 60 *Salmonella* isolates derived from patients suffering from gastroenteritis and visiting a GP. Thirty of these patients had developed colon cancer ≥1 year after the *Salmonella* infection (i.e. the 'case isolates'), whereas the other 30 did not develop cancer (i.e. 'control isolates'). Assessing the *in vitro* infectivity and capacity to induce cellular transformation of MEFs for each of the 60 isolates individually, showed a tendency towards higher infectivity and transformation capacity for case isolates as

compared to control isolates. However, subjecting the individual strains to a gastrointestinal model system resembling the human intestinal tract did not reveal a significant difference in infectivity between case and control isolates. Similarly, no consistent differences were observed in the presence of bacterial genes between case and control isolates. Analysis of the *S*. Typhimurium and *S*. Enteritidis subset revealed some bacterial genes and single nucleotide polymorphisms (SNPs) to be associated with transformation capacity, although the biological relevance of these genes/SNPs in terms of *Salmonella*'s oncogenic potential were unclear. Investigating the isolates' ability to utilize a range of different nutrient sources yielded several differences between isolates with a high *versus* a low transformation capacity. It therefore seems that the tumorigenic potential of *Salmonella* is better explained by phenotypic/metabolic differences rather than genotypic differences. Of note, transcriptional research not only revealed a distinct expression of genes associated with cell cycle, cytokine signaling and immune signaling at 2.5 *versus* 8 hours post infection, but also significant differences in the activation of (metabolic) pathways between the serovars Typhimurium and Enteritidis [57]. Hence, RNA sequencing might comprise a useful tool to examine the transcriptional signature of isolates with a high *versus* low transformation capacity [58]. Part of the isolates with a lower transformation capacity revealed a reduced ability to utilize several amino acids and peptides, whereas a contrasting utilization pattern was observed for isolates with a higher transformation capacity. Likewise, the utilization of several phosphorus and sulfur sources was more consistently low for isolates with low transformation capacity as compared to isolates with high transformation capacity which showed mixed results. Of note, both the transformation assay in MEFs and the nutrient utilization arrays are based on simple models unlike the human gut. Whether the observed variation in transformation capacity and the corresponding patterns in nutrient utilization are still valid in the complex colonic environment and confer a biological advantage on these strains in terms of oncogenic potential, remains to be elucidated. Despite this need for further research deciphering the possible biological relevance of these outcomes, we could generate some hypotheses of underlying mechanisms that may contribute to them. For instance, the development of colonic adenomas and carcinomas is accompanied by significant changes in metabolic pathways to comply with the increasing energy demand of the growing polyp/tumor allowing tumor progression and metastasis. Such metabolic changes include higher demands for nitrogen, amino acids and glucose, and a shift to aerobic glycolysis (i.e. the Warburg effect) [59]. Regarding this point, the review in **Chapter 7** summarized a myriad of studies that found a distinct presence of bacteria in tumor and off-tumor tissues as compared to tissue from healthy controls. In our previous experiments, *Salmonella* showed a tropism for transformed cells. However, this implies that the bacterium has to adjust its nutrient uptake

to the altered metabolic profile of the (early) tumor/polyp microenvironment. As for many bacteria in the gut, glucose is the preferred source of energy for *Salmonella* when available [60]. Yet, *Salmonella* has a profound adaptive response, using host- and microbiota-derived metabolites (produced as an intermediate or by-product during inflammation) to assure its metabolism even under suboptimal conditions [61, 62]. Upon invasion of the host cell, most *Salmonella* bacteria reside in the *Salmonella*-containing vacuole (SCV) without direct access to host cytosolic metabolites/nutrients. Literature suggest that *Salmonella* recruits host cell transporters to obtain nutrients such as arginine and glutamine in the SCV [63, 64]. Since developing colorectal tumors are accompanied with an accumulation of L-Arginine, the observed increased L-arginine utilization by strains with a higher transformation capacity supports the hypothesis of *Salmonella* as a passenger bacteria [65]. *Salmonella* strains with a higher capacity to utilize a broad spectrum of nutrients, as observed in **Chapter 8**, therefore have an advantage of surviving and replicating in the altered microenvironment stimulating tumor progression, particularly when it also outcompetes other bacterial species lacking such metabolic capacities or when microbe-microbe interactions play a role [66].

Several strategies have been postulated by which *Salmonella* might induce colonic tumorigenesis. Briefly, these include the upregulation of Wnt signaling and suppression of JNK and NF-kB by the *Salmonella* AvrA protein and the activation of STAT3 and MAPK/Akt pathways by the *Salmonella* effector proteins leading to epithelial cell proliferation, suppression of host immune responses and cell transformation [67]. A possible complementing strategy might include hyperactivation of the serine-threonine kinase mammalian target of rapamycin (mTOR). mTOR is part of the upstream phosphatidylinositol-3-kinase (PI3K)/Akt pathway, which is involved in a myriad of cellular functions, including metabolism, growth and survival [68]. Aberrant Akt signaling is frequently associated with cancers and constitutes one of the causes of the downstream overactivation of mTOR. On the other hand, the mTOR gene itself is identified as proto-oncogene with strong tumorigenic capacities. A myriad of pathways leading to stimulation of mTOR have been linked to tumorigenesis and proliferation of cancer cells [69, 70]. Overactivation of mTOR has been associated with cellular transformation of MEFs (tested in a soft agar assay) and rapid development of tumors after subcutaneous administration in mice [69]. mTOR consists of two complexes, activation of mTOR complex 1 (mTORC1) is mainly regulated by nutrient availability (particularly amino acid levels, including glutamine, leucine and arginine), growth factors, oxygen levels and reactive oxygen species (ROS) [71]. Under normal conditions, active mTOR complex 1 (mTORC1) induces cell growth and proliferation by the phosphorylation of the downstream protein S6 kinase (S6K) and eukaryotic translation initiation factor 4E-binding protein (4EBP1). In case

of *Salmonella* infection, a transient damage of the SCV membrane shortly after infection (1-2h), induces a rapid decline in cytosolic amino acids (mainly L-leucine and L-isoleucine), a phenomenon also observed in reaction to infection by other intracellular bacteria [72]. This decline appears to be host-driven rather than the result of consumption of amino acids by *Salmonella*. The low levels of amino acids lead to autophagy due to mTORC1 inhibition (i.e. a catabolic response of cells to cope with stress conditions) [72]. However, *Salmonella* triggers a rapid normalization of the amino acid levels in the cytosol (3-4h post infection), as well as relocation of mTOR to the maturing SCV [72]. The mechanism behind the rapid restoration of amino acid levels during *Salmonella* infection as compared to other intracellular bacterial infections is not yet known. Nonetheless, by reactivating mTOR, *Salmonella* escapes from autophagy, thereby favoring its growth and replication within the host [72].

Another route of mTOR activation involves ROS production by mitochondrial activity. Upon infection, mitochondria produce higher levels of ROS as part of the innate host defense against pathogens [73]. Yet, the amount of mitochondrial ROS (mtROS) produced was shown to be dependent on the *Salmonella* serovar [57]. Infection of human intestinal organoids (HOI) with *S*. Enteritidis revealed upregulated gene expression of mitochondrial related processes (including mitochondrial translation, protein import and oxidative phosphorylation) as compared to infection by *S*. Typhimurium. Consistently, the increase in mtROS production between 1 and 24 hours post infection was significantly higher for Enteritidis compared to Typhimurium [57]. Remarkably, enhanced mitochondrial activity has been observed in transformed predisposed MEFs (i.e. harboring $Arf^{-/-}$ and c-MYC mutations) with a history of *S*. Typhimurium infection as compared to predisposed MEFs without a prior *Salmonella* infection. This mitochondrial activity appeared crucial for maintaining the transformed state of the MEFs as inhibition of mitochondrial activity led to abolishment of the *Salmonella*-induced transformation [D.M. van Elsland, personal communication, December 5th, 2019]. These results suggest that the induction of enhanced mitochondrial activity (and the subsequent production of mtROS) is not restricted to *S*. Enteritidis and might vary within serovars. How *Salmonella* reprograms the mitochondria towards sustained activity even when bacterial infection is cleared, remains to be elucidated. Mitochondrial dysfunction has been associated with a multitude of human pathologies, including cancers [74]. Aberrations in mitochondrial metabolism play a role in the oncogenesis cascade from malignant transformation to tumor progression [75]. The overactivated mitochondria continuously produce ROS. As a consequence, the activity of the tumor suppressor PTEN (phosphatase and tensin homolog deleted on chromosome 10) is restrained due to oxidation by ROS [76], which subsequently leads to activation of the Akt pathway and the downstream mTOR

pathway [68]. Remarkably, the effect of ROS on mTORC1 is dose-dependent, as low doses are associated with mTORC1 activation whereas high-dose or long-term ROS exposure leads to decreased mTORC1 activity (as a result of AMPK-induced phosphorylation of Raptor) [77]. Whether mTOR is one of the missing links in the causality of *Salmonella*-induced cancer progression, warrants further study but it is found activated in human gallbladder carcinoma samples from Indian patients that have been chronically infected by *S*. Typhi.

Association between bacteria and gastrointestinal cancer – lessons learned and a way forward

This thesis aimed to contribute to the existing knowledge about the association between *Salmonella* and colon cancer. Overall, the outcomes show a mixed picture. On the one hand, the *in vitro* and *in vivo* experiments showed results clearly supporting an association between *Salmonella* and the development of colon cancer. On the other hand the epidemiological outcomes were less consistent and challenging to interpret. Amongst the possible reasons are the lack of data concerning other possible risk factors, effect modifiers and confounders that warrant attention in the analyses, as well as the inherent risk of (left and right) truncation of observations in the relative short study periods. Moreover, since people are estimated to acquire, on average, a *Salmonella* infection every ~7 years (in the Netherlands), a truly unexposed population does not exist [16]. This implies that epidemiological studies using indicators on a more continuous scale, such as abundance of bacterial DNA in tumor *versus* off-tumor tissue or concentrations of anti-*Salmonella* antibodies, rather than the dichotomous (and sometimes indirect) design of our registry studies, might be recommended approaches for future research, particularly when combined with a multiyear follow-up period. Ideally, one would aim for the implementation of a large-scale prospective cohort study with a long-term follow-up, such as in the LifeLines cohort in the northern part of the Netherlands to assess phenotypic, environmental and genomic parameters related to development of chronic diseases and healthy aging by following people for ≥30 years [78]. The design of such cohort study, including repeated follow-up questionnaires and collection of biomaterials (e.g., blood, feces), as well as linkage to medical/environmental registries, allows for the assessment of the risk of cancer(s) in relation to diet, physical activity, body weight, occupational and environmental exposures, medication use, and (fluctuating) microbiome compositions. Although this might provide better insights into the correlations between risk factors and their relation with microbiome, such study would be costly, time consuming and is highly dependent on the perseverance of the participants. Alternatively, future research can aim at using existing data sources and the application of

machine learning techniques to reveal patterns/associations in the tremendous amount of data generated from high-throughput sequencing methods worldwide.

Overall, the epidemiological outcomes of this thesis seem to indicate the contribution of *Salmonella* infection to the burden of colon cancer in the Netherlands (North West Europe) as negligible, yielding no urgent implications for colon cancer screening policies or *Salmonella* control programs. However, the increased incidence of colon cancer diagnosis within one year after *Salmonella* infection, as well as the elevated rates of proliferation in tumors compared to adjacent normal tissue, and the propensity of *Salmonella* to infect cells with the highest level of transformation all confirm that *Salmonella* has a strong tropism for infecting transformed/malignant cells. This implies that, regardless of the magnitude of *Salmonella* infection as driver for cancer induction at the population level, the bacterium might substantially enhance or accelerate tumor growth with the inherent risk of cancer being diagnosed in a late stage. Therefore, *Salmonella* could be added to the list of microorganisms, including *F. nucleatum* and *Clostridium hathewayi* amongst others, warranting further study into their potential as biomarker for (early) colon cancer diagnosis [79]. Owing to the chemotaxis of *Salmonella* towards tumors, attenuated avirulent *Salmonella* strains are a topic of interest in several animal models and clinical trials to assess their applicability in bacterial-mediated cancer therapy [9, 80]. The mechanisms by which *Salmonella* exerts the anti-tumor effects remain elusive [9]. Yet, despite the contrasting roles of *Salmonella* as oncogenic *versus* therapeutic agent, we can continuously learn from the gained knowledge in the therapeutic discipline, particularly concerning the mechanisms by which *Salmonella* exerts its effects on tumor cells and the interaction with the immune system [9].

The pathogenesis of colon cancer is driven by a tripartite relationship between the microbiome, the mucosal immune balance in the colon and the colonic epithelial cells [81]. However, the vast majority of research on the association between microorganisms and cancer focused on the role of a single microorganism in cancer development or a snapshot of the distinct abundances of bacterial phyla/families in relation to cancer stage. Presence and magnitude of microbe-microbe interactions with either counteracting or amplifying effects on cancer development have hardly been studied. Several studies observed cooperative or antagonistic behaviors between bacterial species or between gut bacteria and the gut virome, such as protection of *S*. Typhimurium against β-lactam antibiotics by *Staphylococcus aureus*, *Bacteroides* species and resistant *E. coli* [82-84], and the inflammation-induced SopEΦ bacteriophage transfer to *S*. Typhimurium, which fosters disease progression [85-87]. Likewise, influenza virus infection has been associated with enhanced *Salmonella* colonization in the gut through the action of type I interferons and an

altered gut microbiome composition [88], highlighting the interplay between microbiome compartments and the immune response. Despite the obvious complexity, future research could aim at deciphering polymicrobial interactions between microbes (in the gut) that potentially exert synergistic effects on tumor formation.

As for many intestinal bacteria, research on the association between *Salmonella* and colon cancer is still in its infancy, and the studies in this thesis were rather explorative in nature. The absence of obvious genetic or metabolic markers or dose-dependent effects explaining the tumorigenic potential of *Salmonella*, suggests a more subtle contribution of *Salmonella* when certain conditions or prerequisites are met. Next steps could involve experimental research focusing on these possible conditions, such as the comparison of transcriptional profiles between serovars/strains and the effects of gene silencing on cell transformation potential. In this thesis, the experimental *in vivo* and *in vitro* research was performed in mice, MEFs and Caco-2 cells. Innovations such as intestines-on-a-chip and cell models with a higher resemblance to the human colon such as HOI might offer an opportunity to capture more realistically the putative tumorigenic role of *Salmonella* in the complex intestinal environment [89]. Moreover, the multitude of bacteria being associated with gastrointestinal cancer initiation and progression suggests a plethora of pathways involved that remain to be unraveled in the coming decades.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a Cancer Journal for Clinicians. 2021;71(3):209-49.
2. Parkin DM. The global health burden of infection-associated cancers in the year 2002. International Journal of Cancer. 2006;118:3030-44.
3. Hansen JP, Ali WM, Sivadasan R, Rajeeve K. Bacteria–Cancer Interface: Awaiting the Perfect Storm. Pathogens. 2021;10(10),1321.
4. Vlaanderen F, Cuperus T, Keur I, De Rosa M, Rozendaal H, Friesema I, et al. Staat van Zoönosen 2020. Rijksinstituut voor Volksgezondheid en Milieu (RIVM). 2021;0190.
5. Mughini-Gras L, Schaapveld M, Kramers J, Mooij S, Neefjes-Borst EA, Pelt WV, et al. Increased colon cancer risk after severe *Salmonella* infection. PloS one. 2018;13(1):e0189721.
6. Mughini-Gras L, Enserink R, Friesema I, Heck M, van Duynhoven Y, van Pelt W. Risk factors for human salmonellosis originating from pigs, cattle, broiler chickens and egg laying hens: a combined case-control and source attribution analysis. PloS one. 2014;9(2):e87933.
7. Waldner LL, MacKenzie KD, Köster W, White AP. From exit to entry: long-term survival and transmission of *Salmonella*. Pathogens. 2012;*1*(2):128-155.
8. Havelaar AH, Garssen J, Takumi K, Koedam MA, Dufrenne JB, Van Leusden FM, et al. A rat model for dose–response relationships of *Salmonella* Enteritidis infection. Journal of Applied Microbiology. 2001;91(3):442-52.
9. Mi Z, Feng ZC, Li C, Yang X, Ma MT, Rong PF. *Salmonella*-mediated cancer therapy: an innovative therapeutic strategy. Journal of Cancer. 2019;10(20):4765.
10. Bongers IM, van der Meer JB, van den Bos J, Mackenbach JP. Socio-economic differences in general practitioner and outpatient specialist care in The Netherlands: a matter of health insurance?. Social Science & Medicine. 1997;44(8):1161-68.
11. Smits JPJM, Droomers M, Westert GP. Sociaal-economische status en toegankelijkheid van zorg in Nederland. RIVM rapport. 2003;279601002/2002.
12. Denktaş S, Koopmans G, Birnie E, Foets M, Bonsel G. Ethnic background and differences in health care use: a national cross-sectional study of native Dutch and immigrant elderly in the Netherlands. International Journal for Equity in Health. 2009;8(1):1-9.
13. Tsiaoussis J, Antoniou MN, Koliarakis I, Mesnage R, Vardavas CI, Izotov BN, et al. Effects of single and combined toxic exposures on the gut microbiome: current knowledge and future directions. Toxicology Letters. 2019;312:72-97.
14. Simonsen J, Strid MA, Molbak K, Krogfelt KA, Linneberg A, Teunis P: Sero-epidemiology as a tool to study the incidence of *Salmonella* infections in humans. Epidemiology and Infection. 2008;136(7):895-902.
15. Teunis PFM, van Eijkeren JCH, Ang CW, van Duynhoven YTHP, Simonsen JB, Strid MA, et al. Biomarker dynamics: estimating infection rates from serological data. Statistics in Medicine. 2012;31:2240–8.
16. Falkenhorst G, Simonsen J, Ceper TH, van Pelt W, de Valk H, Sadkowska-Todys M, et al. Serological cross-sectional studies on *Salmonella* incidence in eight European countries: no correlation with incidence of reported cases. BMC Public Health. 2012;12(1),1-9.
17. Mølbak K, Simonsen J, Jørgensen CS, Krogfelt KA, Falkenhorst G, Ethelberg S, et al. Seroincidence of human infections with nontyphoid *Salmonella* compared with data from public health surveillance and food animals in 13 European countries. Clinical Infectious Diseases. 2014;59(11):1599-1606.

18. Shaw KS, Cruz-Cano R, Jiang C, Malayil L, Blythe D, Ryan P, et al. Presence of animal feeding operations and community socioeconomic factors impact salmonellosis incidence rates: An ecological analysis using data from the Foodborne Diseases Active Surveillance Network (FoodNet), 2004–2010. Environmental Research. 2016;150:166-72.
19. Poulsen MN, Pollak J, Sills DL, Casey JA, Rasmussen SG, Nachman KE, et al. Residential proximity to high-density poultry operations associated with campylobacteriosis and infectious diarrhea. International Journal of Hygiene and Environmental Health. 2018;221(2):323-33.
20. Aleksandrova K, Pischon T, Jenab M, Bueno-de-Mesquita HB, Fedirko V, Norat T, et al. Combined impact of healthy lifestyle factors on colorectal cancer: a large European cohort study. BMC Medicine. 2014;12(1),1-15.
21. Behrens G, Gredner T, Stock C, Leitzmann MF, Brenner H, Mons U. Cancers due to excess weight, low physical activity, and unhealthy diet: estimation of the attributable cancer burden in Germany. Deutsches Aerzteblatt International. 2018;115(35-36):578.
22. Cho S, Shin A. Population attributable fraction of established modifiable risk factors on colorectal cancer in Korea. Cancer Research and Treatment: Official Journal of Korean Cancer Association. 2021;53(2):480.
23. Kim H, Wang K, Song M, Giovannucci EL. A comparison of methods in estimating population attributable risk for colorectal cancer in the United States. International Journal of Cancer. 2021;148(12):2947-53.
24. Simons CC, Hughes LA, van Engeland M, Goldbohm RA, van den Brandt PA, Weijenberg MP. Physical activity, occupational sitting time, and colorectal cancer risk in the Netherlands cohort study. American Journal of Epidemiology. 2013;177(6):514-30.
25. Montano D. Chemical and biological work-related risks across occupations in Europe: a review. Journal of Occupational Medicine and Toxicology. 2014;9(1),1-13.
26. Pukkala E, Martinsen JI, Lynge E, Gunnarsdottir HK, Sparén P, Tryggvadottir L, et al. Occupation and cancer–follow-up of 15 million people in five Nordic countries. Acta Oncologica. 2009;48(5):646-790.
27. Blair A, Freeman LB. Epidemiologic studies in agricultural populations: observations and future directions. Journal of Agromedicine. 2009;14:125–31.
28. Smith L, McCourt O, Sawyer A, Ucci M, Marmot A, Wardle J, et al. A review of occupational physical activity and sedentary behaviour correlates. Occupational Medicine. 2016;66(3):185-92.
29. Vanthomme K, van den Borre L, Vandenheede H, Hagedoorn P, Gadeyne S. Site-specific cancer mortality inequalities by employment and occupational groups: a cohort study among Belgian adults, 2001–2011. BMJ Open. 2017;7(11):e015216.
30. Bosdriesz JR, Willemsen MC, Stronks K, Kunst AE. Tobacco control policy and socio-economic inequalities in smoking in 27 European countries. Drug and Alcohol Dependence. 2016;165:79-86.
31. Clonan A, Roberts KE, Holdsworth M. Socioeconomic and demographic drivers of red and processed meat consumption: implications for health and environmental sustainability. Proceedings of the Nutrition Society. 2016;75(3):367-73.
32. O'Donoghue G, Kennedy A, Puggina A, Aleksovska K, Buck C, Burns C, et al. Socio-economic determinants of physical activity across the life course: A" DEterminants of DIet and Physical ACtivity"(DEDIPAC) umbrella literature review. PloS one. 2018;13(1):e0190737.
33. Leufkens AM, van Duijnhoven FJ, Boshuizen HC, Siersema PD, Kunst AE, Mouw T, et al. Educational level and risk of colorectal cancer in EPIC with specific reference to tumor location. International Journal of Cancer. 2012;130(3):622-30.

34. Teunis PF, Kasuga F, Fazil A, Ogden ID, Rotariu O, Strachan NJ. Dose–response modeling of *Salmonella* using outbreak data. International Journal of Food Microbiology. 2010;44(2):243-9.
35. Havelaar A, Garssen J, Takumi K, Koedam M, Ritmeester W, de la Fonteyne L, et al. Intraspecies variability in the dose-response relationship for *Salmonella* Enteritidis associated with genetic differences in cellular immune response. Journal of Food Protection. 2004;67(9):2008-15.
36. Van Parys A, Boyen F, Leyman B, Verbrugghe E, Maes D, Haesebrouck F, et al. Induction of seroconversion and persistence of *Salmonella* Typhimurium in pigs are strain dependent. Comparative Immunology, Microbiology and Infectious Diseases. 2013;36(5):465-71.
37. Helaine S, Thompson JA, Watson KG, Liu M, Boyle C, Holden DW. Dynamics of intracellular bacterial replication at the single cell level. Proceedings of the National Academy of Sciences. 2010;107(8):3746-51.
38. Wegener HC, Hald T, Wong LF, Madsen M, Korsgaard H, Bager F, et al. *Salmonella* control programs in Denmark. Emerging Infectious Diseases, 2003;9(7):774.
39. Anonymous. Annual Report on Zoonoses in Denmark 2019. 2020. National Food Institute, Technical University of Denmark.
40. Anonymous. Annual Report on Zoonoses in Denmark 2002. 2003. Ministry of Food, Agriculture and Fisheries.
41. Anonymous. Annual Report on Zoonoses in Denmark 2007. 2009. National Food Institute, Technical University of Denmark.
42. Anonymous. Annual Report on Zoonoses in Denmark 2012. 2013. National Food Institute, Technical University of Denmark.
43. Anonymous. Annual Report on Zoonoses in Denmark 2018. 2019. National Food Institute, Technical University of Denmark
44. Anonymous. Annual Report on Zoonoses in Denmark 2009. 2010. National Food Institute, Technical University of Denmark.
45. Havelaar AH, Friesema IHM, van Pelt W. Disease burden of food-related pathogens in the Netherlands, 2010. RIVM Letter Report. 2012;330331004/2012.
46. Mughini-Gras L, Pijnacker R, Duijster J, Heck M, Wit B, Veldman K, et al. Changing epidemiology of invasive non-typhoid *Salmonella* infection: a nationwide population-based registry study. Clinical Microbiology and Infection. 2020;26(7):941-e9.
47. Eurostat. Share of overweight population by sex and age, 2014 (%) FP20. Accessed at 29 December 2021. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Share_of_overweight_population_by_sex_and_age,_2014_(%25)_FP20.png#filelinks
48. Biltoft-Jensen A, Kørup K, Christensen T, Eriksen K, Hess Ygil K, Fagt S. 2016. Køds rolle i kosten. DTU Fødevareinstituttet. Retrieved from: koedets-rolle-i-kosten.pdf
49. RIVM. 2020. Consumptie van vlees in Nederland – SHIFT-DIETS-project. Retrieved from: Consumptie van vlees in Nederland (rivm.nl)
50. Kato I, Boleij A, Kortman GA, Roelofs R, Djuric Z, Severson RK, et al. Partial associations of dietary iron, smoking and intestinal bacteria with colorectal cancer risk. Nutrition and. Cancer. 2013;65(2):169-77.
51. Gal-Mor O, Boyle EC, Grassl GA. Same species, different diseases: how and why typhoidal and nontyphoidal *Salmonella* enterica serovars differ. Frontiers in Microbiology. 2014;5:391.

52. Scanu T, Spaapen RM, Bakker JM, Pratap CB, Wu LE, Hofland I, et al. *Salmonella* manipulation of host signaling pathways provokes cellular transformation associated with gallbladder carcinoma. Cell Host & Microbe. 2015;17(6):763-74.
53. Chang R, Wei JCC, Lin MC, Hung YM. The potential role of nontyphoidal salmonellosis in gastric cancer: a nationwide matched cohort study. Gastric Cancer. 2021;24(2):292-301.
54. Iyer P, Barreto SG, Sahoo B, Chandrani P, Ramadwar MR, Shrikhande SV, et al. Nontyphoidal *Salmonella* DNA traces in gallbladder cancer. Infectious Agents and Cancer. 2016;11(1):1-4.
55. Tjalsma H, Boleij A, Marchesi JR, Dutilh BE. A bacterial driver–passenger model for colorectal cancer: beyond the usual suspects. Nature Reviews Microbiology. 2012;10(8):575-82.
56. Lu R, Bosland M, Xia Y, Zhang YG, Kato I, Sun J. Presence of *Salmonella* AvrA in colorectal tumor and its precursor lesions in mouse intestine and human specimens. Oncotarget. 2017;8(33):55104.
57. Abuaita BH, Lawrence ALE, Berger RP, Hill DR, Huang S, Yadagiri VK, et al. Comparative transcriptional profiling of the early host response to infection by typhoidal and nontyphoidal *Salmonella* serovars in human intestinal organoids. PLoS pathogens, 2021;17(10):e1009987.
58. Shah DH. RNA sequencing reveals differences between the global transcriptomes of *Salmonella* enterica serovar enteritidis strains with high and low pathogenicities. Applied and Environmental Microbiology. 2014;80(3):896-906.
59. Pavlova NN, Thompson CB. The emerging hallmarks of cancer metabolism. Cell Metabolism. 2016;23(1):27-47.
60. Eisenreich W, Dandekar T, Heesemann J, Goebel W. Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. Nature Reviews Microbiology. 2010;8(6):401-12.
61. Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, et al. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. Nature. 2010;467(7314):426-9.
62. Taylor SJ, Winter SE. *Salmonella* finds a way: Metabolic versatility of *Salmonella* enterica serovar Typhimurium in diverse host environments. PLoS pathogens. 2020;16(6):e1008540.
63. Das P, Lahiri A, Lahiri A, Sen M, Iyer N, Kapoor N, et al. Cationic amino acid transporters and *Salmonella* Typhimurium ArgT collectively regulate arginine availability towards intracellular *Salmonella* growth. PloS one. 2010;5(12):e15466.
64. Popp J, Noster J, Busch K, Kehl A, Zur Hellen G, Hensel M. Role of host cell-derived amino acids in nutrition of intracellular *Salmonella* enterica. Infection and Immunity. 2015;83(12):4466-75.
65. Lu Y, Wang W, Wang J, Yang C, Mao H, Fu X, et al. Overexpression of arginine transporter CAT-1 is associated with accumulation of L-arginine and cell growth in human colorectal cancer tissue. PloS one. 2013;8(9):e73866.
66. Zhao L, Cho W, Nicolls M. Colorectal Cancer-Associated Microbiome Patterns and Signatures. Frontiers in Genetics. 2012;12.
67. van Elsland D, Neefjes J. Bacterial infections and cancer. EMBO reports. 2018;19(11):e46632.
68. Francipane MG, Lagasse E. mTOR pathway in colorectal cancer: an update. Oncotarget. 2014;5(1):49.
69. Murugan AK, Alzahrani A, Xing M. Mutations in critical domains confer the human mTOR gene strong tumorigenicity. Journal of Biological Chemistry. 2013;*288*(9):6511-21.
70. Zou Z, Tao T, Li H, Zhu X. mTOR signaling pathway and mTOR inhibitors in cancer: progress and challenges. Cell & Bioscience. 2020;10(1):1-11.
71. Rabanal-Ruiz Y, Otten EG, Korolchuk VI. mTORC1 as the main gateway to autophagy. Essays in Biochemistry. 2017;61(6):565-584.

72. Tattoli I, Sorbara MT, Vuckovic D, Ling A, Soares F, Carneiro LA, et al. Amino acid starvation induced by invasive bacterial pathogens triggers an innate host defense program. Cell Host & Microbe. 2012;11(6):563-75.

73. Silwal P, Kyung Kim J, Jae Kim Y, Jo EK. Mitochondrial reactive oxygen species: double-edged weapon in host defense and pathologic inflammation during infection. Frontiers in Immunology. 2020;11:1649.

74. Gammage PA, Frezza C. Mitochondrial DNA: the overlooked oncogenome?. BMC Biology. 2019;17(1):1-10.

75. Srinivasan S, Guha M, Kashina A, Avadhani NG. Mitochondrial dysfunction and mitochondrial dynamics-The cancer connection. Biochimica et Biophysica Acta (BBA)-Bioenergetics. 2017;1858(8):602-14.

76. Zhang Y, Park J, Han SJ, Yang SY, Yoon HJ, Park I, et al. Redox regulation of tumor suppressor PTEN in cell signaling. Redox Biology. 2020;34:101553.

77. Li M, Zhao L, Liu J, Liu A, Jia C, Ma D, et al. Multi-mechanisms are involved in reactive oxygen species regulation of mTORC1 signaling. Cellular Signalling. 2010;22(10):1469-76.

78. Scholtens S, Smidt N, Swertz MA, Bakker SJ, Dotinga A, Vonk JM, et al. Cohort Profile: LifeLines, a three-generation cohort study and biobank. International Journal of Epidemiology. 2015;44(4):1172-1180.

79. Liang Q, Chiu J, Chen Y, Huang Y, Higashimori A, Fang J, et al. Fecal bacteria act as novel biomarkers for noninvasive diagnosis of colorectal cancer. Clinical Cancer Research. 2017;23(8):2061-70.

80. Silva-Valenzuela CA, Desai PT, Molina-Quiroz RC, Pezoa D, Zhang Y, Porwollik S, et al. Solid tumors provide niche-specific conditions that lead to preferential growth of *Salmonella*. Oncotarget. 2016;7(23):35169.

81. Sears CL, Pardoll DM. Perspective: alpha-bugs, their microbial partners, and the link to colon cancer. Journal of Infectious Diseases. 2011;203(3):306-11.

82. Perlin MH, Clark DR, McKenzie C, Patel H, Jackson N, Kormanik C, et al. Protection of *Salmonella* by ampicillin-resistant *Escherichia coli* in the presence of otherwise lethal drug concentrations. Proceedings of the Royal Society B: Biological Sciences. 2009;276(1674):3759-68.

83. Lee J, Lee EY, Kim SH, Kim DK, Park KS, Kim KP, et al. *Staphylococcus aureus* extracellular vesicles carry biologically active β-lactamase. Antimicrobial Agents and Chemotherapy. 2013;57(6):2589-95.

84. Stentz R, Horn N, Cross K, Salt L, Brearley C, Livermore DM, et al. Cephalosporinases associated with outer membrane vesicles released by *Bacteroides* spp. protect gut pathogens and commensals against β-lactam antibiotics. Journal of Antimicrobial Chemotherapy. 2015;70(3):701-9.

85. Diard M, Bakkeren E, Cornuault JK, Moor K, Hausmann A, Sellin ME, et al. Inflammation boosts bacteriophage transfer between *Salmonella* spp. Science. 2017;355(6330):1211-15.

86. Wahl, A., Battesti, A., & Ansaldi, M. (2019). Prophages in Salmonella enterica: a driving force in reshaping the genome and physiology of their bacterial host?. Molecular microbiology, 111(2), 303-316.

87. Kato I, Zhang J, Sun J. Bacterial-Viral Interactions in Human Orodigestive and Female Genital Tract Cancers: A Summary of Epidemiologic and Laboratory Evidence. Cancers. 2022;14(2):425.

88. Deriu E, Boxx GM, He X, Pan C, Benavidez SD, Cen L, et al. Influenza virus affects intestinal microbiota and secondary *Salmonella* infection in the gut through type I interferons. PLoS Pathogens. 2016;12(5):e1005572.

89. Puschhof J, Pleguezuelos-Manzano C, Clevers H. Organoids and organs-on-chips: Insights into human gut-microbe interactions. Cell Host & Microbe. 2021;29(6):867-78.

Appendix

SUMMARY

SAMENVATTING

CURRICULUM VITAE

LIST OF PUBLICATIONS

DANKWOORD

Summary

A growing body of scientific literature documents a putative role of commensal and pathogenic bacteria in the initiation and progression of cancers. One such bacterium is nontyphoidal *Salmonella*, which has been associated with colon cancer in a few studies. Yet, a lot is still unknown about the magnitude and underlying mechanisms, including the necessary conditions or 'prerequisites', of the potential colon carcinogenesis promoting effects of *Salmonella*. The main objective of this thesis was to elucidate the role of nontyphoidal *Salmonella* infection in the development of cancers in the gastrointestinal tract, with particular focus on colon cancer. To this end, we performed several complementary analyses based on both experimental and epidemiological study designs.

In **chapter 2, 3 and 4** we investigated whether frequent exposure to zoonotic pathogens such as *Salmonella* is associated with an increased risk of colon cancer. Exposure to *Salmonella* might occur in occupational settings, for instance due to contact with live animals or (raw) products thereof. In **chapter 2** we assessed the association between reported culture-confirmed *Salmonella* and *Campylobacter* infection and occupation in a nationwide registry study in the Netherlands in the period 1999-2016. To this end, we used an internationally agreed upon occupational classification system. Besides analyses of the incidence in occupational divisions within the total dataset, we defined three high-risk groups for zoonotic infection. These high-risk groups include occupations possibly associated with frequent exposure to (low dose) *Salmonella* through 1) contact with live animals and animal manure (e.g., farmers, abattoir workers, veterinarians), 2) processing of foods of animal origin (e.g., cooks and chefs) and 3) sale of products of animal origin (e.g., butchers and cheese store employees). Occupational data was linked to salmonellosis or campylobacteriosis data at the individual person level. We compared the incidence of salmonellosis and campylobacteriosis in the occupational divisions and high-risk groups with the incidence in the matching group (based on *Salmonella* serovar or *Campylobacter* species, calendar year and people's gender and age group) in the total employed population of the Netherlands. The standardized incidence ratios of both salmonellosis and campylobacteriosis were significantly increased in the risk groups involving contact with live animals or animal manure and food sale (1.4-1.8-fold increased). Also, significant excess incidence of reported salmonellosis and campylobacteriosis was observed in healthcare-related occupations and several industrial occupations. In addition to reported infections, we compared the serological incidence (or seroincidence) of *Salmonella* or *Campylobacter* in a subset of the employed population. The seroincidence is defined as the estimated number of (in this case *Salmonella* or *Campylobacter*) infections per person per year, which provides

a less biased method for the estimation of the infection pressure (i.e., force of infection), as it is based on seroconversion (i.e., immune response-eliciting infection), rather than clinical disease alone. Little variation was found in the seroincidences between occupational groups or high-risk groups, hence, the observed differences in incidence of reported infections were only slightly reflected by the infection pressure. In **chapter 3**, we used a comparable study design to explore the incidence of colon cancer among different professions and assessed whether the occupational groups with increased salmonellosis incidence were also those more prone to develop colon cancer. In accordance with previous literature, occupation in itself provided little differences in colon cancer incidence, as relatively minor differences in incidences were observed among occupational divisions. Likewise, no increased risk of colon cancer was observed in the three pre-defined high-risk groups with higher incidences of salmonellosis. Significant higher incidences of colon cancer were found for several industrial divisions, as well as a few non-manual occupational divisions associated with more sedentary tasks. The contribution of the major lifestyle-related risk factors of colon cancer and the heterogeneity of these risk factors among occupational groups might have diluted the relatively minor putative effect of *Salmonella*.

In **chapter 4,** we assessed whether the risk of colon cancer is associated with *Salmonella* exposure earlier in life by using serum samples from a Dutch serosurvey, which were linked to colon cancer diagnosis data. To this end, we compared the *Salmonella* seroincidence of 36 people who developed proximal colon cancer >1 year after participation in the serosurvey (i.e., cases) to the seroincidence of 72 matched individuals without a colon cancer diagnosis (i.e. controls). Matching was done on gender, age, educational level, socioeconomic status and smoking behavior. The seroincidence was significantly higher in cases *versus* controls in the subgroup aged <60 years at time of the serosurvey and the subgroup with a high socioeconomic status. Besides the epidemiological analyses, we also investigated whether repeated exposure with a low dose of *Salmonella* can induce a tumorigenic response under experimental conditions. Mice with a predisposition for colon cancer received either three low-dose infections or one high-dose infection with *S*. Typhimurium. Both infected groups of mice, as well as the control group (without infection), developed colon tumors. However, the tumors of infected mice were larger and showed more often high-grade dysplasia. Yet, repeated low-dose infections with *Salmonella* sufficed to induce a similar tumorigenic effect, as no differences were observed between the mice with multiple low- *versus* single high-dose infections. Also, mouse embryonic fibroblasts (MEFs) with a predisposition for cancer were infected with a high or a low dose *S*. Typhimurium. After this first infection, cellular transformation was observed in both groups of MEFs in a soft agar assay. Reinfection of

these transformed MEFs (i.e. with prior exposure to *Salmonella*) with the same low or high dose led to the formation of more and larger colonies as compared to the first infection. Still, two-fold infection with a low dose was more successful as compared to a single high-dose infection (i.e. more and larger colonies).

In **chapter 5,** we continued with an epidemiological study in a Danish cohort to complement the findings of an earlier Dutch cohort study which found an increased risk after reported (severe) *Salmonella* infection. In the Danish cohort, the incidence of colon cancer was compared for individuals with a reported *Salmonella* infection in the past *versus* individuals without a reported *Salmonella* infection in the period 1994-2015. Again, the salmonellosis surveillance data was linked to colon cancer diagnosis data and demographic data. Cox regression showed no overall increased risk of colon cancer after *Salmonella* infection. However, the risk of proximal colon cancer was significantly increased (1.4-fold) >1 year after infection with serovars other than Enteritidis and Typhimurium (the two serovars which cause over 70% of the infections). Remarkably, an over two-fold increased incidence of colon cancer was found among individuals with a *Salmonella* infection less than one year before the cancer diagnosis. These results differ from the findings of the earlier Dutch cohort study which found an increased risk of proximal colon cancer in people aged <60 when infected, particularly when infected with *S*. Enteritidis.

The association between chronic *S*. Typhi infection and gallbladder carcinoma is well established and earlier research showed that also *S*. Typhimurium is able to induce tumorigenesis in gallbladder organoids. Whether nontyphoidal *Salmonella* is associated with a higher risk of biliary tract (including gallbladder) cancer was addressed in **chapter 6**. By comparing the incidence of biliary tract cancer >1 year after *Salmonella* or *Campylobacter* infection with the incidence of biliary tract cancer in the general Dutch population, we found a tendency towards a higher risk of biliary tract cancer after *Salmonella* but not *Campylobacter* infection. Yet, as biliary tract cancer is rare in the Netherlands, the numbers were too low to reach the level of significance.

As the list of microorganisms being associated with a variety of cancers continues to grow, we performed a literature review to summarize the epidemiological studies addressing the association between bacteria and parasites and cancers in the gastrointestinal tract (**chapter 7**). Overall, we provided an overview of the study designs and main findings of 158 studies covering 10 bacteria and three parasites. Most studies addressed the association between bacteria and colon or colorectal cancer. Many of these compared the presence or abundance of a bacterium in cancer patients *versus* healthy controls or tumor tissue *versus*

off-tumor tissue, which does not allow for assessing cause-effect relationships. A relative small portion of studies included a follow-up time of several years.

In **chapter 8,** we aimed to further unravel a causal link between *Salmonella* infection and colon cancer from a bacterial perspective. We therefore performed a 'case-control study' using 60 *Salmonella* isolates selected based on the linkage of national *Salmonella* surveillance and cancer diagnosis data in the Netherlands. Indeed, through this linkage, 30 isolates were identified that were obtained from patients with a *Salmonella* infection who developed proximal colon cancer >1 year after the infection (i.e., case isolates), and another 30 matched isolates (based on *Salmonella* serovar, gender, age and calendar year) were obtained from patients who were not diagnosed with colon cancer (i.e., control isolates). All these isolates were sent to the Dutch National Institute for Public Health and the Environment (RIVM) for typing as part of the national (human) salmonellosis surveillance. The selected 60 *Salmonella* isolates were used in several analyses. No difference in infectivity was observed between case isolates and control isolates in a cell model resembling the human gastrointestinal tract. On the other hand, we found substantial variation between isolates in terms of their capacity to infect MEFs and to induce cellular transformation, with a tendency towards higher transformation efficiency in the case isolates. Whole-genome sequencing of the isolates did not reveal biologically relevant genes or single nucleotide polymorphisms significantly associated with transformation efficiency. Assessing the capacity of the isolates to utilize a broad range of carbon, nitrogen, phosphorus and sulfur sources showed a significant positive correlation between transformation efficiency and utilization of a range of sources, mainly amino acids, peptides and phosphorus sources. Isolates with a larger metabolic flexibility in response to nutrient availability possess a biological advantage during human infection. However, the implications of these findings for *Salmonella*'s putative tumorigenic potential and which pathways/mechanisms might be involved need to be further unraveled.

Samenvatting

Dikkedarmkanker (hierna 'darmkanker' genoemd) is de op twee na meest voorkomende vorm van kanker wereldwijd. In Nederland wordt jaarlijks bij meer dan 8,000 personen darmkanker vastgesteld. Hoewel de incidentie van darmkanker in de oudere leeftijdsgroepen al daalt mede door de invoering van darmkankerscreening bij personen 55-75 jaar (sinds 2014), stijgt de incidentie in jongere leeftijdsgroepen (<50 jaar). Een klein deel van de gediagnosticeerde darmkankers (≤5%) heeft een erfelijke oorzaak. Naast de bekende risicofactoren zoals overgewicht, consumptie van rood en bewerkt vlees, een tekort aan lichaamsbeweging, alcoholgebruik, roken en een lage inname van voedingsvezels en calcium, spelen mogelijk andere nog onbekende risicofactoren een rol in de ontwikkeling van darmkanker. Een toenemend aantal micro-organismen wordt geassocieerd met de inductie en ontwikkeling van kanker. Enkele bekende voorbeelden zijn de associatie tussen het humaan papillomavirus en hepatitis virus en de ontwikkeling van respectievelijk baarmoederhalskanker en leverkanker. In tegenstelling tot virussen laten bacteriën geen genetische afdruk achter in cellen waardoor het causale verband tussen bacteriële infectie en kanker lange tijd is onderkend. Tijdens infectie manipuleren bacteriën de bestaande processen/routes van cellulaire communicatie in de gastheer wat mogelijk aantasting van de integriteit van de gastheercel tot gevolg heeft. Dit heeft voor de bacterie als doel om overleving en replicatie in de gastheer te waarborgen en te optimaliseren, echter het kan als nevenschade ook kwaadaardige transformatie van cellen en daarmee het ontstaan van kanker teweegbrengen. De best onderzochte associaties tussen bacteriële infectie en kanker zijn op dit moment *Helicobacter pylori* als veroorzaker van maagkanker en chronische *Salmonella* Typhi infectie als oorzaak van galblaaskanker. Daarnaast wijst experimenteel en epidemiologisch onderzoek op een verband tussen niet-tyfeuze *Salmonella* en darmkanker. Niet-tyfeuze *Salmonella* is een van de meest voorkomende oorzaken van gastro-enteritis wereldwijd en infectie heeft veelal een mild, zelflimiterend beloop. Het bewijs voor een mogelijke associatie tussen niet-tyfeuze *Salmonella* en kanker is slechts gebaseerd op enkele onderzoeken en veel is nog onbekend over de mechanismen en condities voor het mogelijke verband tussen *Salmonella* en darmkanker en mogelijke andere vormen van kanker in het spijsverteringskanaal. Daarom heeft deze thesis als hoofddoel de bestaande kennis over de mogelijke associatie tussen niet-tyfeuze *Salmonella* infectie en kanker in het spijsverteringskanaal verder uit te diepen, met een focus op darmkanker. Hiervoor hebben we een geïntegreerde epidemiologische en experimentele aanpak gebruikt.

In **hoofdstukken 2, 3 en 4** onderzoeken we of frequente blootstelling aan zoönotische pathogenen zoals *Salmonella* een verhoogd risico op darmkanker tot gevolg heeft. Een van de

mogelijke manieren van blootstelling is via werk, bijvoorbeeld door contact met levend vee of (rauwe) producten van dierlijke oorsprong. In **hoofdstuk 2** hebben we in een landelijke registerstudie onderzocht of er een verband is tussen incidentie van gerapporteerde humane salmonellose en campylobacteriose en beroep in de Nederlandse bevolking, daarbij gebruikmakend van een internationaal beroepsclassificatiesysteem. Naast analyse van de incidenties in de verschillende beroepsgroepen in de totale dataset, hebben we drie risicogroepen gedefinieerd. Deze risicogroepen omvatten beroepen mogelijk geassocieerd met een frequente blootstelling aan (een lage dosis) *Salmonella* via 1) contact met levende dieren en dierlijke mest (bijv. veehouders, slachthuismedewerkers en dierenartsen), 2) dierlijke producten tijdens voedselproductie/-bereiding (bijv. koks) en 3) dierlijke producten tijdens verkoop (bijv. slagers en medewerkers van kaas-/delicatessenwinkels). De incidentie van salmonellose en campylobacteriose in de beroepsgroepen en risicogroepen werd vergeleken met de incidentie in de matchende groep (o.b.v. *Salmonella* serovar of *Campylobacter* species, geslacht en leeftijdsgroep van de mensen, en kalenderjaar) in de totale beroepsbevolking (de referentiepopulatie) in de periode 1999-2016. De uitkomstmaat was een gestandaardiseerde incidentie ratio (SIR). SIRs voor zowel salmonellose en campylobacteriose waren significant verhoogd in de risicogroep met blootstelling aan levende dieren of dierlijke mest en de verkoop van dierlijke producten (1,4-1,8 maal verhoogd). De incidentie van gerapporteerde salmonellose/ campylobacteriose was eveneens verhoogd in gezondheidszorg-gerelateerde beroepen en een aantal industriële beroepen.

Tevens hebben we in deze studie de *Salmonella*- en *Campylobacter*-seroincidentie vergeleken tussen verschillende beroepsgroepen in een subgroep van de beroepsbevolking. Gerapporteerde infecties geven het topje van de ijsberg weer, gezien met name infecties met ernstige of langdurige klachten worden gerapporteerd. Seroincidentie daarentegen is gebaseerd de ontwikkeling van een immuunrespons in reactie op infectie ongeacht ziekteverschijnselen. Hiertoe hebben we gebruik gemaakt van serummonsters uit een eerdere seroprevalentiestudie. De anti-*Salmonella*/-*Campylobacter* IgM, IgG en IgA antilichaamtiters in een individueel serum sample werden gebruikt als input om de periode sinds de laatste seroconversie (blootstelling aan pathogeen) te berekenen door middel van een Bayesiaans model. De periode sinds laatste seroconversie kon vervolgens worden geconverteerd naar een seroincidentie, gedefinieerd als het geschatte aantal infecties per persoonsjaar. De spreiding in seroincidentie van *Salmonella* en *Campylobacter* was slechts beperkt binnen verschillende beroepsgroepen en was niet significant hoger in de drie gecombineerde risicogroepen in vergelijking met de seroincidenties in niet-risicogroepen.

In **hoofdstuk 3** hebben we de associatie tussen darmkanker en beroep onderzocht in een landelijke registerstudie in de periode 2000-2016. Om de uitkomsten te kunnen vergelijken

met de eerdere beroepsstudie (hoofdstuk 2) hebben we gebruik gemaakt van dezelfde beroepsclassificatie en risicogroepen. De incidentie van darmkanker (proximaal noch distaal) was niet significant verhoogd in de drie risicogroepen vergeleken met de incidentie in matchende referentiegroep in de algemene Nederlandse bevolking. Over het algemeen waren de verschillen in darmkankerincidentie tussen de onderzochte beroepsgroepen klein. Ook bleek de seroincidentie van *Salmonella* in personen die een aantal jaar later werden gediagnosticeerd met proximale darmkanker niet significant hoger dan in gematchte controles die geen darmkanker ontwikkelden, behalve in de groep die een serummonster afstonden voor hun 60e levensjaar (**hoofdstuk 4**). In een experimentele setting is onderzocht of een herhaalde infectie van gepredisponeerde muizen met een lage dosis *S.* Typhimurium *versus* een enkele infectie met een hoge dosis eenzelfde tumorontwikkeling teweegbrengt. Dit bleek het geval te zijn, echter tumoren van muizen met een *Salmonella*-infectie waren groter en invasiever dan in de tumoren in de controle groep. Ook was *Salmonella* in hogere mate aanwezig in tumorweefsel in vergelijking met het omliggende weefsel. Dit werd bevestigd in een *in vitro* experiment waar gepredisponeerde embryonale muizenfibroblasten (MEFs) met een herhaalde blootstelling aan een lage dosis *S.* Typhimurium meer kolonies vormden in agar dan MEFs met een enkele hoge dosis infectie. Herhaalde blootstelling aan een hoge dosis *Salmonella* leidde echter tot een hoger aantal kolonies. Bovendien had *Salmonella* een voorkeur voor infectie van cellen met de hoogste mate van pretransformatie.

In **hoofdstukken 5 en 6** hebben we de associatie tussen *Salmonella*-infectie geanalyseerd in twee retrospectieve cohortstudies. De studie in hoofdstuk 5 had als doel om de associatie tussen ernstige (d.w.z. gerapporteerde) salmonellose en darmkanker te onderzoeken d.m.v. een landelijke cohortstudie in de Deense populatie in de periode 1994-2016. Door data uit verschillende registratiesystemen op persoonsniveau te koppelen konden we de incidentie van darmkanker vergelijken tussen de groep personen met een gerapporteerde *Salmonella*-infectie in het verleden en de groep zonder gerapporteerde *Salmonella*-infectie (controle groep). Daarbij werd onderscheid gemaakt tussen darmkankerdiagnose <1 jaar na *Salmonella*-infectie en ≥ 1 jaar na infectie. In het algemeen was het risico op darmkanker niet hoger in de groep met *Salmonella*-infectie ≥ 1 jaar geleden ten opzichte van de controle groep, met uitzondering van de subgroep met gerapporteerde infectie door serovars anders dan Enteritidis en Typhimurium waar een 1,4 maal verhoogd risico op proximale darmkanker werd gevonden. Daarentegen was de incidentie van darmkanker in de groep met een recente infectie meer dan tweemaal hoger dan in de controle groep. In **hoofdstuk 6** hebben we een vergelijkbare studie uitgevoerd om het risico op galblaas- en galwegkanker na *Salmonella*- of *Campylobacter*-infectie vast te stellen in de Nederlandse populatie in

de periode 2000-2016. Galblaas-/galwegkanker is relatief zeldzaam in Nederland (~800 diagnoses/jaar), echter de overlevingskans is laag en diagnose kan in veel gevallen niet worden geattribueerd aan de bekende risicofactoren. Hoewel de associatie tussen *S.* Typhi en galblaascarcinoom inmiddels vaak onderwerp is van studie, is er nog weinig bekend over het risico op galblaas-/galwegkanker na niet-tyfeuze *Salmonella* infectie. Kankerdiagnoses <1 jaar na infectie werden geëxcludeerd in de analyse, evenals infecties voor de leeftijd van 20 jaar. Galblaas-/galwegkanker werd vastgesteld bij negen personen met een gerapporteerde *Salmonella*-infectie en zeven personen met een gerapporteerde *Campylobacter*-infectie. Deze aantallen waren te laag om statistische significantie te bereiken, echter voor *Salmonella* was er een indicatie voor een hogere kans op galblaas-/galwegkanker na infectie.

Het afgelopen decennium is het aantal gepubliceerde studies die het verband tussen micro-organismen (met name bacteriën) en kanker in het spijsverteringskanaal onderzoeken, substantieel toegenomen. Dit is onder andere te danken aan de snelle ontwikkeling van sequencing-methoden zoals '*whole-genome sequencing*' dat onderzoek naar onder andere het microbioom in relatie tot gezondheid en ziekte in een stroomversnelling heeft gebracht. **Hoofdstuk 7** bevat een reviewartikel van de epidemiologische studies die het verband onderzoeken tussen bacteriën/parasieten en kanker in de spijsverteringsorganen. *H. pylori* hebben we niet meegenomen in deze literatuurstudie omdat voor deze bacterie al veel reviewartikelen bestaan. Het artikel bevat een samenvatting van de belangrijkste methodologische aspecten en resultaten van 158 studies waarin het verband met kanker wordt onderzocht voor verschillende 10 bacteriën en 3 parasieten. Naast de Deense cohortstudie uit hoofdstuk 5, werd in vier andere studies de associatie tussen niet-tyfeuze *Salmonella* en colorectaalkanker onderzocht. Eén daarvan was een Nederlandse registerstudie waarin een significant verhoogde incidentie van darmkanker werd gevonden in personen met een gerapporteerde *Salmonella*-infectie in de leeftijd 20-60 jaar. De incidentie van proximale darmkanker was bijna drie keer zo hoog na infectie met *S.* Enteritidis in vergelijking met de algemene bevolking. Deze resultaten konden niet worden bevestigd in zowel de Deense cohortstudie als een Taiwanese cohortstudie. Daarentegen werd in twee andere studies een significante associatie gevonden tussen *Salmonella* antilichaamtiter en colorectaalkanker en tussen aanwezigheid van *Salmonella* AvrA eiwit in feces en colorectaalkanker.

De studie in **hoofdstuk 8** had als doel meer inzicht te krijgen in mogelijke genetische factoren die verband houden met de vermeende oncogene rol van *Salmonella* alsmede mogelijke mechanismen die daarbij betrokken zijn. In een case-control studie hebben we 30 *Salmonella*-isolaten van mensen met een gerapporteerde *Salmonella*-infectie die een aantal jaar later werden gediagnosticeerd met darmkanker ('case-isolaten') vergeleken met 30

Salmonella-isolaten van mensen zonder darmkankerdiagnose ('control-isolaten'). De cases en controles werden gematcht op basis van serotype, geslacht, leeftijd en jaar van infectie. De 60 *Salmonella*-isolaten werden vervolgens gebruikt voor verschillende experimenten. In een *in vitro* model bleek de mate waarin isolaten in staat waren MEFs te infecteren en cellulaire transformatie te induceren te verschillen. Een verschil in infectiviteit werd echter niet gevonden in een *in vitro* model systeem wat het humane spijsverteringskanaal nabootst. Ook werd in het genoom van de *Salmonella*-isolaten geen eenduidige aanwijzing gevonden voor betrokkenheid van specifieke genen of '*single-nucleotide polymorphisms*' (SNPs) bij de capaciteit om transformatie te induceren. Wel bleken de isolaten met een hogere transformatie-efficiëntie beter in staat om een variëteit aan nutriënten te verbruiken, zoals aminozuren, peptiden en andere stikstof- en fosforbronnen, in tegenstelling tot de isolaten met een lagere transformatie-efficiëntie. In de studie werden er geen duidelijke verschillen tussen case-isolaten en control-isolaten gevonden op genotypisch en fenotypisch niveau.

Op basis van de uitkomsten van deze thesis kunnen we verschillende conclusies trekken. Allereerst, waren de resultaten van de epidemiologische analyses minder consistent en moeilijker te interpreteren in vergelijking met de experimentele analyses. Een mogelijke reden hiervoor is dat *Salmonella* waarschijnlijk een kleine rol speelt in darmkanker ontwikkeling ten opzichte van de bekende leefstijl-gerelateerde risicofactoren die het risico van *Salmonella* overschaduwen. Correctie voor deze risicofactoren die in tijd variëren is in veel epidemiologische studiedesigns niet mogelijk. Daarnaast wordt elk individu meermaals blootgesteld aan *Salmonella* gedurende zijn/haar leven. Een niet-blootgestelde populatie bestaat in feite niet. Gerapporteerde infecties omvatten slechts een klein deel van het werkelijke aantal infecties, waardoor de kans bestaat dat de associatie tussen *Salmonella* en darmkanker wordt onderschat. Op basis van de uitkomsten van de muisexperimenten en de *in vitro* analyses in deze thesis kunnen we concluderen dat *Salmonella* onder bepaalde omstandigheden (die in werkelijkheid voor kunnen komen) in staat is om de ontwikkeling van darmkanker te initiëren en daarmee een '*driver*' rol kan innemen. Daarnaast blijkt uit de *in vivo* en *in vitro* analyses alsmede de observatie dat de darmkankerdiagnoses met name verhoogd zijn in het eerste jaar na *Salmonella* infectie, dat *Salmonella* een tropisme heeft voor het infecteren van gepretransformeerde cellen en kankercellen. Dit ondersteunt de hypothese dat *Salmonella* een '*passenger*' is die profiteert van het veranderde microklimaat in/om de tumor en mogelijk de progressie van darmkanker stimuleert. Hoewel we in deze thesis geen causale link hebben kunnen vaststellen tussen *Salmonella* en darmkanker, biedt de geobserveerde variatie in cellulaire transformatie capaciteit en de correlatie met het gebruik van nutriënten mogelijkheden voor verder onderzoek naar de verstoring van cellulaire processen in de gastheer die leiden tot tumorontwikkeling.

Curriculum vitae

Janneke Duijster was born on February 13th 1991 in Ede, the Netherlands. In 2009, she graduated from pre-university education (VWO) at the Van Lodenstein College in Amersfoort. Thereafter she enrolled the bachelor program Animal Sciences at the Wageningen University, where she obtained her Bachelor's degree in 2013. During her Master program (Animal Sciences) she did a specialization in Quantitative Veterinary Epidemiology. She performed her final master's internship at the Centre for Monitoring of Vectors (CMV, part of the Netherlands Food and Consumer Product Safety Authority) on the Dutch vector monitoring strategy within the framework of West Nile virus preparedness. After obtaining her Master's degree in 2015, she started working as a junior researcher in the outbreak detection and surveillance research group at the department of Epidemiology and Surveillance at the National Institute for Public Health and the Environment (RIVM). In that period, the focus of her work was mainly the epidemiology of mosquito-borne diseases and syndromic surveillance of infectious diseases. In 2017, she had the opportunity to start her PhD project on the association between *Salmonella* infection and colon cancer at the RIVM in collaboration with Leiden University Medical Center (LUMC).

List of publications

Part of this thesis

- **Duijster JW**, Franz E, Neefjes JJ, Mughini-Gras L. (2019). Occupational risk of salmonellosis and campylobacteriosis: a nationwide population-based registry study. *Occupational and environmental medicine*, *76*(9), 617-624.

- De Savornin Lohman E, **Duijster J**, Groot Koerkamp B, Van der Post R, Franz E, Mughini Gras L, De Reuver P. (2020). Severe *Salmonella* spp. or *Campylobacter* spp. infection and the risk of biliary tract cancer: A population-based study. *Cancers*, *12*(11), 3348.

- **Duijster JW**, Hansen JV, Franz E, Neefjes JJ, Frisch M, Mughini-Gras L, Ethelberg S. (2021). Association between *Salmonella* infection and colon cancer: a nationwide registry-based cohort study. *Epidemiology & Infection*, *149*.

- **Duijster J**, Mughini-Gras L, Neefjes J, Franz E. (2021). Occupational exposure and risk of colon cancer: a nationwide registry study with emphasis on occupational exposure to zoonotic gastrointestinal pathogens. *BMJ Open* 2021;11:e050611.

- **Duijster JW**, Franz E, Neefjes J, Mughini-Gras L. (2021). Bacterial and parasitic pathogens as risk factors for cancers in the gastrointestinal tract: a review of current epidemiological knowledge. *Frontiers in Microbiology*, 2021(12): 3746.

- Van Elsland DM, **Duijster JW**, Zhang J, Stévenin V, Zhang YG, Zha L, Xia Y, Franz E, Sun J, Mughini-Gras L, Neefjes JJC. Repetitive exposure to non-typhoidal Salmonellae is an environmental risk factor for colon cancer. Submitted to: *Cell Reports Medicine (2022)*.

Outside this thesis

- Sondén K, Rolling T, Wångdahl A, Ydring E, Vygen-Bonnet S, Kobbe R, ... & Färnert A. (2019). Malaria in Eritrean migrants newly arrived in seven European countries, 2011 to 2016. *Eurosurveillance*, 24(5), 1800139.

- Monge S, **Duijster J**, Kommer GJ, Van de Kassteele J, Donker GA, Krafft T, ... & Van Asten L. (2020). Use of ambulance dispatch calls for surveillance of severe acute respiratory infections. *Emerging infectious diseases*, 26(1), 148.

- Monge S, **Duijster J**, Kommer GJ, Van de Kassteele J, Krafft T, Engelen P, ... & Van Asten L. (2020). Ambulance dispatch calls attributable to influenza A and other common respiratory viruses in the Netherlands (2014-2016). *Influenza and other respiratory viruses*, 14(4), 420-428.

- **Duijster JW**, Doreleijers SD, Pilot E, Van der Hoek W, Kommer GJ, Van der Sande MA, ... & Van Asten L. C. (2020). Utility of emergency call centre, dispatch and ambulance data for syndromic surveillance of infectious diseases: a scoping review. *European journal of public health*, 30(4), 639-647.

- Mughini-Gras L, Pijnacker R, **Duijster J**, Heck M, Wit B, Veldman K, & Franz E. (2020). Changing epidemiology of invasive non-typhoid *Salmonella* infection: a nationwide population-based registry study. *Clinical Microbiology and Infection*, 26(7), 941-e9.

-

Dankwoord

Onderzoek doen en een proefschrift schrijven doe je niet alleen, daarom wil ik graag iedereen bedanken die direct en/of indirect een steentje heeft bijdragen aan de totstandkoming van dit proefschrift. Een aantal mensen wil ik in het bijzonder noemen in dit beknopte dankwoord.

Allereerst mijn promotor, Sjaak. Jouw aanstekelijke enthousiasme, optimisme en onuitputtelijke stroom aan ideeën gaf me telkens weer een impuls om méér te willen leren en méér onderzoek te doen over dit onderwerp. Heel veel dank voor de leerzame jaren en de kans die ik heb gekregen om dit promotietraject te doen. Lapo en Eelco, wat een goede co-promotoren waren jullie. Naast de inhoudelijke kennis en expertise waar ik enorm veel van heb geleerd ook ontzettend bedankt voor de vrijheid om mijn eigen draai aan het onderzoek te geven en jullie mental support, ook in de tijd dat het wat minder ging.

Daphne en Virginie, bedankt voor de fijne samenwerking en alle celbiologische kennis die jullie me hebben bijgebracht. Claudia, Angela en Lucas, van jullie heb ik zoveel nieuwe onderzoeksmethoden en -technieken geleerd, die verbreding is enorm waardevol voor mij.
Jun and colleagues in Chicago, thank you for performing the mouse experiments, the *in vivo* part is an essential addition to our research. Steen, Jørgen, Karen and Morton, the week at SSI (Denmark) was a great opportunity to assist in some of the laboratory work for the project and starting the research for our joint manuscript. Thank you very much for the fruitful collaboration. Andra, bedankt voor het pathologische werk en het meedenken met ons onderzoek. Elise, ik vond het leuk om samen met jou een artikel te schrijven over een aanverwant onderwerp, dank voor de fijne samenwerking.

Alle (ex-)EPI-collega's, bedankt voor de mooie tijd. De koffiemomentjes in de pantry, lunches en wandelingetjes waren niet alleen een goed moment om te sparren, ze gaven ook de nodige afleiding en ontspanning. In het bijzonder de GEZ-collega's; wat een fijne groep zijn jullie om mee samen te werken. Roan en Ingrid, bedankt voor alle brainstormmomenten tijdens uitbraakonderzoeken, dat was een leerzame afleiding van mijn promotieonderwerp. Kamergenoten Wilma, Sonja en Laura bedankt voor de gezelligheid. Naast collega's en degenen die een directe bijdrage hebben geleverd zijn er natuurlijk veel mensen die indirect hebben bijgedragen. Vriend(inn)en, dankjulliewel voor het meeleven met de vorderingen van mijn onderzoek, de goede gesprekken en vooral de ontspanning tijdens alle gezellige avonden/dagen. Pa en ma, de basis voor dit alles heb ik aan jullie te danken. Heel veel dank voor alle liefde, steun, belangstelling en altijd luisterend oor. Zussen, broers, (schoon)zussen en zwagers, jullie zijn er altijd voor mij, het is niet in woorden uit te

drukken hoe waardevol dat is. Arno en Henk, hoewel ik bij de start van mijn promotietraject me nog geen voorstelling kon maken dat ik mijn proefschrift ook daadwerkelijk ooit zou verdedigen, één ding stond al vast: ik wil mijn beide broers als paranimfen. Bedankt dat jullie dit willen doen. Schoonfamilie, bedankt voor jullie belangstelling voor mijn proefschrift, gezamenlijke klusdagen en alle gezellige familiemomenten.

Lieve Michel, wat een geluk dat ik jou heb leren kennen. Bedankt voor álles wat je hebt gedaan om dit proefschrift mogelijk te maken!