RISK STRATIFICATION IN DUTCH PRIMARY CARE

A Promising Approach to Manage Population Health

Shelley-Ann Girwar

RISK STRATIFICATION IN DUTCH PRIMARY CARE

A Promising Approach to Manage Population Health

Shelley-Ann Girwar

Department of Public Health and Primary Care of the Leiden University Medical Center / Health Campus The Hague

ISBN: 978-94-6483-163-4
Cover: Shelley-Ann Girwar
Lay-out: Publiss | www.publiss.nl
Print: Ridderprint | www.ridderprint.nl

This research was funded by Noaber Foundation.

Financial support for the printing of this thesis was provided by the department of Public Health and Primary Care of the Leiden University Medical Center / Health Campus the Hague

Risk Stratification in Dutch Primary Care

A Promising Approach to Manage Population Health

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 donderdag 22 juni 2023
klokke 13.45 uur

door

Shelley-Ann Maria Girwar
geboren te Paramaribo, Suriname
in 1987

Promotores

Prof. dr. M. E. Numans
Prof. dr. M. Fiocco

Co-promotor

Dr. M. A. Bruijnzeels

Promotiecommissie

Prof. dr. M. R. Spruit
Dr. M. R. Crone
Prof. dr. R. H. H. Groenwold
Prof. dr. H. Schers, Radboudumc
Dr. M. D. Spreeuwenberg, Maastricht University Medical Center

For Micah

Where the writing of this dissertation ended, your life began.

May your generation benefit from Population Health Management approaches including Risk Stratification, both in the Netherlands and in my homeland Suriname.

May innovative and efficient approaches improve Suriname's healthcare system one day.

Mi Kondre tru, mi lobi yu *(My country, I love you)*

TABLE OF CONTENT

CHAPTER 1

CHAPTER 1

GENERAL INTRODUCTION

Our personal "data" are widely being collected, stored and (re-)used in a systematic matter. With registries of websites, app subscriptions and even our browsing history, our data become subject to analyses with or without our knowledge and awareness. Besides official authorities such as governmental and legal authorities, commercial businesses such as grocery - and clothing stores are collecting and storing our daily data. Assuming that privacy regulations are taken into account, storage and analysis of our data can be used for our benefit. How easy is it to use for example, the information we receive from a supermarket bonus card scheme, derived from our grocery shopping habits? How convenient is it, that based on our previous clothing purchases, we get information on sales products that match our clothing style? With systematic analyses of our data, we can be provided with tailored advertisement, which we may appreciate much more than large-scale collective advertisement.

Although we also store a lot of data in healthcare, we are not (re-) using these data as efficiently as the commercial industry does. The majority of provided care is offered in a similar way as large-scale collective advertisement, for instance with standardized care programs to fit an assumed persona or population. We still have a lot to learn about systematically using healthcare data in order to provide patients at specific risk with tailored interventions to reduce that risk. This is what risk stratification is all about: systematic analysis of healthcare data in such a way that interventions can be tailored to individuals' risks of specific adverse outcomes.

Risk stratification

Due to the inequality of healthcare needs within populations, large-scale 'one-size-fits-all' healthcare programs are prone to losing their efficacy. Data-driven approaches making these inequities in healthcare needs visible, are therefore potentially very beneficial and routine registry data are increasingly being used to provide proactive care, attempting to lower health costs by reducing expensive and avoidable care such as hospital admissions. Risk stratification, the assignment of individual risk scores to patients based on registered health profiles, has proven to be an efficient and effective tool in the provision of proactive care. In addition, risk stratification can be used to allocate available resources according to needs instead of demands, as healthcare needs are usually unequally distributed within populations. Various studies have shown the use of risk stratification tools to determine morbidity burden (1, 2) and identify healthcare utilization inequities (3). Risk stratification has also proven to be effective in selecting the right subpopulations for patient-centered interventions aiming at lowering expensive healthcare utilization such as hospitalizations (4).

Risk stratification in Dutch Primary Care

Data-driven approaches using risk stratification are not currently commonly used in the Netherlands. Hospitals are trying to introduce “data-driven healthcare” and “value based healthcare” principles into their internal organization. However, despite the availability of data and the great possibilities of providing proactive care in primary care (which is due to the gatekeeper function of General Practitioners (GPs) in the Netherlands), routine primary care data are not being used systematically for Population Health Management purposes. Regional attempts have been made to establish proactive care interventions based on data, in the primary care context, such as the U-care program in Utrecht. This program was aiming to test the cost-efficacy of early detection of frail elderly, followed by personalized proactive care interventions. It is still used as a means to prioritize care in many practice centers. The early detection in this program concerned a frailty assessment based on patients’ electronic medical record data (mainly physical deficiencies), followed by a more detailed evaluation of cognitive, social and psychological domains when the Frailty Index score was above a certain threshold (5). Although this is a fair attempt to use data to assess the risk of being frail and thus performing risk stratification in Dutch primary care, it required the additional collection of information through interviews and questionnaires rather than fully relying on systematically usable routinely collected registry data. Other approaches are mostly targeted on specific subpopulations or specific diseases. Risk stratification analyses on complete primary care populations using registry data in a systematic way, are lacking.

Built on the issues discussed above, the studies reported in this thesis aim at finding a suitable risk stratification method that can be used on general populations in Dutch primary care.

Risk stratification in the context of Population Health Management

Risk stratification is an important element of Population Health Management (PHM), a revitalized new concept in healthcare. Although many different definitions for PHM exist, the Triple Aim value perspective seems important in most definitions, ranking patients' value and quality of care as equally important elements as lowering of costs (6). With the patient's value in mind, providing personalized care is one of the main pillars of PHM. An efficient way to analyze risks and needs of patients and thus enabling personalized care, is to use data-driven approaches.

Data-driven approaches in the form of risk stratification, also support efficiency and proactive care provision, other important pillars of PHM: performing risk stratification in PHM approaches, enables us to allocate resources to those who will benefit most from intervention. Assigning specific risks to individuals gives opportunities to identify the need for targeted proactive care. With aging populations, the view of the burden of disease is changing from that of single chronic diseases to more complex patterns of diseases, recognizing that multimorbidity is the norm (7). This change in epidemiology is causing the already diminishing healthcare resources to run out at an even faster rate. The use of innovative technology within healthcare and the increasing demands of patients about the way their healthcare should be organized, are adding to this problem. With reduced budgets available for healthcare and a decreasing workforce due to the aging population, our healthcare is becoming too expensive to afford. The effective and proactive nature of PHM, which can be supported by risk stratification, can offer solutions.

Even though risk stratification is possibly one of the strengths of PHM approaches, PHM goes beyond the allocation of resources to the identification of appropriate subpopulations and the provision of proactive care. PHM highlights the importance of the coordination of care delivery across a specified population. In PHM, improvement of both clinical and financial outcomes is gained through a set of not only individual, but also organizational and cultural interventions. (8, 9). This requires multidisciplinary, collaboration of different stakeholders and consideration of the governance structures on different levels. Most healthcare systems are organized in fragmented non-integrated silos, which are built up according to single organization, medical or social conditions. Separate educational programs, payment systems and pressure groups in parliament show the fragmentation of our healthcare systems. The organizational interventions of PHM approaches, which are striving towards interdisciplinarity and collaboration across disciplines, are therefore becoming a necessity rather than a luxury.

PHM provides solutions to renew and improve our healthcare systems, with risk stratifying data-driven approaches as important tools towards more efficient and sustainable healthcare systems. To enable cultural interventions and to provide holistic profiles of patients when performing risk stratification, it is important to also include social and cultural determinants of health rather than sticking to just biomedical determinants. This thesis while less focused on social and cultural elements, concentrated on routinely available data in Dutch GP practices.

Objective of this thesis

The objective of this thesis was to identify a risk stratification tool which can be used in Dutch primary care covering overall general practice populations. Different risk stratification tools are used internationally. However, the suitability for the use in primary care varies between the different tools. In addition, the appropriateness of any tool is dependent on the goal that needs to be achieved. Specifically in primary care, the goal should be to improve the general health and wellbeing of the population. Consequently, a risk stratification tool fit for primary care should be based on complete patients' profiles rather than on specific disease outcomes.

The research question to be answered with this thesis is 'What risk stratification tool is most suitable for Dutch primary care and how can this tool be used with Dutch routine primary care data?

Different sub-questions can be derived:

- What risk stratification tool is most appropriate for use in primary care?
- Is this risk stratification tool applicable to the Dutch primary care data?
- How do we select the right data for proper application of the risk stratification tool in primary care?
- How well does this risk stratification tool perform in Dutch primary care?
- How can we adjust the risk stratification tool to appropriately fit the Dutch primary care setting?

Outline of this thesis

The second chapter of this thesis was aimed at answering the following question *'What risk stratification tool is most appropriate for use in primary care?'* In order to answer this question, a systematic review was performed to identify and assess different risk stratification models used internationally in primary care. Calibration and discrimination properties of the models were employed to evaluate the performance.

Chapter 3 reported results of a pilot study in which a risk stratification tool known to be appropriate for the use in primary care, is applied to a Dutch primary care setting. Results show the applicability of the tool using Dutch primary care data.

The fourth chapter discussed the results of the main study within this thesis. In this study, the performance of the risk stratification model in terms of calibration and discrimination properties, used in Dutch primary care, was statistically assessed. In addition, the model was adjusted in order to best fit the Dutch primary care situation.

To visualize the effects and benefits of using a risk stratification model in primary care, a model which identifies a specific subgroup within the population, that would benefit most from certain interventions, was estimated. This is discussed in chapter 5. The subgroup to be identified was a group of complex patients with problems on multiple health domains and with above average acute care utilization. Identifying this type of subgroup for intensive care management interventions, can allocate healthcare resources efficiently yielding to proper intervention programs and eventually lowering overall healthcare costs.

This thesis ends with recommendations for further research, especially regarding societal and practical impact.

References

1. Kristensen T, Rose Olsen K, Sortso C, Ejersted C, Thomsen JL, Halling A. Resources allocation and health care needs in diabetes care in Danish GP clinics. Health Policy. 2013;113(1–2):206–15.
2. Sibley LM, Moineddin R, Agha MM, Glazier RH. Risk adjustment using administrative data-based and survey-derived methods for explaining physician utilization. Med Care. 2010;48(2):175-182.
3. Shadmi E, Balicer RD, Kinder K, et al. Assessing socioeconomic health care utilization inequity in Israel: impact of alternative approaches to morbidity adjustment. BMC Public Health. 2011;11:609.
4. Soto-Gordoa M, de Manuel E, Fullaondo A, et al. Impact of stratification on the effectiveness of a comprehensive patient-centered strategy for multimorbid patients. Health Serv Res. 2019;54(2):466-473.
5. Bleijenberg N, Drubbel I, Neslo RE, Schuurmans MJ, ten Dam VH, Numans ME, de Wit GA, de Wit NJ. Cost-Effectiveness of a proactive primary care program for frail older people: a cluster-randomized controlled trial. J Am Med Dir Assoc 2017(18);12:1029-36.e3.
6. Steenkamer, B. M., Drewes, H. W., Heijink, R., Baan, C. A., & Struijs, J. N. Defining population health management: a scoping review of the literature. Population health management. 2017 Feb; 20(1), 74-85.
7. Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med. 2007;22(suppl 3):391-395.
8. Hill M. (2002). McGraw Hill Concise Medical Dictionary of Modern Medicine.
9. Hillman, Michael. Testimony before the Subcommittee on Health of the House of Committee on Ways and Means, hearing on promoting disease management in Medicare

CHAPTER 2

CHAPTER 2

A SYSTEMATIC REVIEW OF RISK STRATIFICATION TOOLS INTERNATIONALLY USED IN PRIMARY CARE SETTINGS

Shelley-Ann M. Girwar, Robert Jabroer, Marta Fiocco, Stephen P. Sutch, Mattijs E. Numans, Marc A. Bruijnzeels

Published in Health Sciences Reports. 2021 Jul 23;4(3):e329. doi: 10.1002/hsr2.329

Abstract

Introduction

In our current healthcare situation, burden on healthcare services is increasing, with higher costs and increased utilization. Structured Population Health Management has developed as an approach to balance quality with increasing costs. This approach identifies sub-populations with comparable health risks, to tailor interventions for those that will benefit the most. Worldwide, the use of routine healthcare data extracted from electronic health registries for risk stratification approaches is increasing. Different risk stratification tools are used on different levels of the healthcare continuum. In this systematic literature review, we aimed to explore which tools are used in primary healthcare settings and assess their performance.

Methods

We performed a systematic literature review of studies applying risk stratification tools with health outcomes in primary care populations. Studies in OECD countries published in English language journals were included. Search engines were utilised with keywords e.g. 'primary care', 'risk stratification' and 'model'. Risk stratification tools were compared based on different measures: Area Under the Curve (AUC) and C-statistics for dichotomous outcomes and R2 for continuous outcomes.

Results

The search provided 4718 articles. Specific election criteria such as primary care populations, generic health utilization outcomes, and routinely collected data sources identified 61 articles, reporting on 31 different models. The three most frequently applied models were the Adjusted Clinical Groups (ACG, n=23), the Charlson Comorbidity Index (CCI, n=19) and the Hierarchical Condition Categories (HCC, n=7). Most AUC and C-statistic values above 0.7, with ACG showing slightly improved scores compared to the CCI and HCC (typically between 0.6 and 0.7).

Conclusion

Based on statistical performance, the validity of the ACG was the highest, followed by the CCI and the HCC. The ACG also appeared to be the most flexible, with the use of different international coding systems and measuring a wider variety of health outcomes.

Introduction

For several decades healthcare costs have been rising. This has been attributed to ageing populations and innovative ways of curing and treating diseases, leading to an increased prevalence of chronic illnesses and comorbidities among community dwelling older people (1). Also patients have increased demands regarding increasing choice around the way their healthcare should be organized and have tended to utilize more care. Furthermore, the needs for healthcare are not evenly distributed within populations. In Western countries, the sickest 5% of the population make up for 50% of the total healthcare costs (2). In order to maintain high quality healthcare, resources should be distributed according to the needs of the population instead of the demand. One way of dealing with this is to allocate resources according to the individual care needs in subpopulations. Predicting healthcare utilization and health outcomes based on needs provides opportunities to allocate resources more appropriately. Predictions of health outcomes through risk stratification can be used to tailor proactive clinical care, to install preventive measures, to restructure healthcare and to improve insight for healthcare professionals. In the long run this approach will help improve the quality of care and reduce the costs (3,4).

A way to monitor and predict costly patient outcomes such as hospitalization, high care utilization and emergency department visits, is through the use of structured population health management programs. Population Health Management is an approach that aims to improve the health of a defined group of people and to strive for more equitable distribution of health outcomes within the group. In Population Health Management programs, an important step is to stratify individuals within a specific subpopulation according to the risk of experiencing an adverse event, such as defined undesirable health outcomes or the extent of their healthcare utilization. Stratification analyses are often performed based on the use of routinely collected healthcare data. Typically, the high-risk sub-population usually comprises of a small percentage of the total population. The medium-and low-risk subpopulations are much larger with around 35% of the overall population classified as medium-risk and 60% as low-risk (2). The identification of people classified on their respective risk-estimates is referred to as risk stratification. Preceding risk stratification population segmentation is performed. Segmentation can be performed based on general characteristics such as age, gender and specific diseases, but also on morbidity and healthcare utilization patterns. A discussion of segmentation was outside the scope of this study.

Many methods for risk stratification exist internationally. Current literature regarding risk stratification models prominently focus on stratifying hospital populations, based on readily available hospital data. However, primary care data has a great potential to improve healthcare quality and reduce health costs (5). Especially in countries where primary care registries have nearly 100% coverage of the total population, such as the Netherlands and the United Kingdom (UK), the opportunity arises to assess the whole population by using these routinely collected primary care data. Distribution of risk in a primary care population is different from a hospital or specialized care population. Current literature also mainly focuses on risk stratification models with disease specific outcomes, whereas in this study the focus is on more generic utilization outcomes such as risk on hospitalization, emergency department visits, future high healthcare utilization and high pharmaceutical expenditures.

The aim of this study was to perform a systematic literature review to describe and assess the performance of different risk stratification tools with generic health utilization outcomes using routinely collected data, and with possibilities of application to the European context, such as in Dutch primary care. Based on the description of the performance of the tools, we recommend the risk stratification tool best suited for usage in Dutch primary care.

Methods

The PRISMA statements regarding conduction and reporting systematic literature reviews were followed throughout the literature review process (6).

This review was conducted through searches in the search engines Pubmed and Embase. The search-string which contained both keywords and MeSH terms is shown in supplementary table 1. The most important keywords were ‘primary care’, ‘risk stratification’ and ‘model’. EndNote X8.2 was used as the reference manager for the articles. The search-string was produced in collaboration with the Leiden University Medical Center (LUMC) Walaeus library.

The PRISMA flow diagram displays the numbers of included and excluded articles (figure 1).

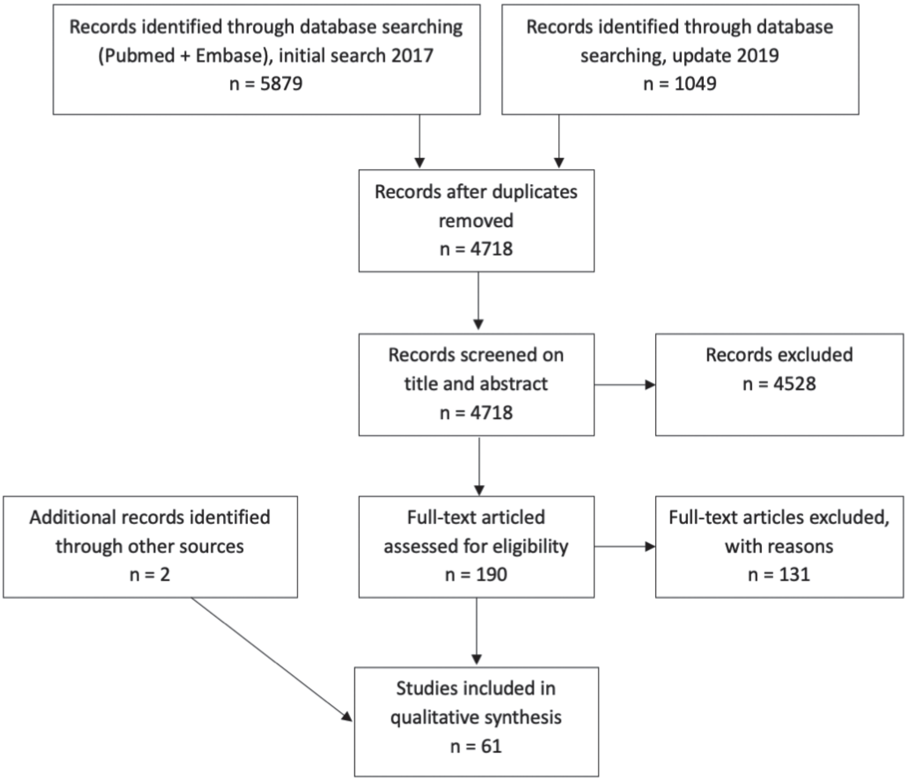


Figure 1: PRISMA flow chart displaying numbers of included and excluded articles

Inclusion criteria

The search characteristics are specified by the **P**opulation, **I**ntervention, **C**ontrol and Outcome (PICO) method. In our research, the *population* is the primary care population. Therefore, we only included articles where models applied to primary care populations are discussed. The *interventions* investigated were the risk stratification approaches and models that are applied to primary care data. *Outcomes* investigated are risks of hospitalization, high healthcare costs, emergency department visits, high pharmaceutical drug expenditure, mortality and other generic health utilization outcomes.

For comparability with a Western-societal environment such as the Dutch situation, only studies performed in countries listed with the Organisation for Economic Cooperation and Development (OECD) (7), were included. Only freely accessible articles in the English language were considered eligible. Articles from January 2007 till August 2019 were reviewed. The inclusion criteria narrowed the search down to a context which was more applicable in a European primary care situation with a gatekeeper's role, such as the Dutch primary care system.

Exclusion criteria

Articles that used risk stratification tools on populations consisting of hospitalized patients, or patients seeking consultation with a specialist (*e.g.* an oncologist or cardiologist) were excluded. These patients were not considered to represent those in a primary care setting. In addition, research looking at specific disease outcome was also excluded, as this review aims at exploring general population outcomes. Articles not freely accessible were excluded as well as articles that were not available in English.

The initial search, conducted in December 2017, yielded 5879 articles. In September 2019, an update of the search was conducted, resulting in an additional 1049 articles. After removing duplicates according to the manual of the Free University (VU) library (8), 4718 articles remained. Articles were screened on both title and abstract, based on the criteria mentioned earlier. 78% of the screening, based on title and abstract, was performed by two researchers independently (R.J. & S.G.). Their results were compared, and in the case of disagreement (2%), the articles were discussed until consensus was achieved. The main causes for disagreement concerned indistinct and misunderstood study populations and model outcomes. As the percentage of disagreement was low, the remaining 22% of the titles and abstracts where only screened by one researcher. After screening on title and abstract, 190 articles remained to be screened on their full text. Screening of all 190 full papers was performed by the same two researchers independently and results were compared. Again, in case of disagreement (21%), the article was discussed until consensus was achieved. After exclusion of 131 articles, including 17 titles which were either not freely accessible or where no English versions of the full papers were available, 59 studies remained to be included in this review. Two further articles were added through the snowball method, resulting in 61 articles.

Assessing performance of models

The different models were compared on three aspects: frequency of use, statistical diagnostic validity, and performance in primary care.

For each identified risk stratification model, **the frequency of use** of the model was presented, taking into account all included studies.

For the assessment of the **statistical diagnostic validity**, reviewed studies were divided into *application, validation* and *comparison* studies. In the *application* studies, risk stratification tools were applied for purposes other than assessing their statistic diagnostic validity. Therefore, *application* studies did not present any statistical diagnostic measures of the risk stratification tools. In the *validation* studies and in most of the *comparison* studies, statistical diagnostic measures of the applied risk stratification tools were provided. Area Under the Curve (AUC) and C-statistics for models with dichotomous outcomes and R^2 values for models with continues outcomes were used to validate risk stratification tools. Models with AUC or C-statistic values between 0.5 and 0.6 were classified as performing *poorly*, values between 0.6 and 0.7 were considered *sufficient* and values above 0.7 were considered *good* (9). Ten of the reviewed papers, the *comparison* studies, compared more than one risk stratification tool in the same study population with the same record data, enabling a more appropriate comparison between risk stratification tools. Most of the comparison studies presented statistical diagnostic values, as they are mostly also validation studies.

For **performance in primary care**, we assessed the type of routinely collected data that is used as input of the model. Models using input data available in Dutch primary care health records were assumed to have a good potential performance in Dutch primary care.

Results

A total of 31 risk stratification models were identified in the literature. The three most frequently applied tools, taking into account all included studies, concern the Adjusted Clinical Groups (ACG), the Charlson Comorbidity Index (CCI) and the Hierarchical Condition Categories (HCC). These three main risk stratification tools are presented in table 1, with predicted outcomes and diagnostic values. Assessment of these tools, their diagnostic validity and applicability in primary care are described in order. The remaining 28 risk stratification tools can be found in supplementary table 2.

Table 1: Overview of the three most frequently identified risk stratification models with their characteristics and diagnostic properties for different outcomes.

	First author, year	Adjusted Clinical Group (ACG)	Charlson Comorbidity Index (CCI)	Hierarchical Condition Categories (HCC)
Categories		ACG-categories (1-93), Resource Utilization Bands (RUBs), Expended Diagnosis Clusters (EDC) count	Six categories based on chronic condition count	Score based on aggregated conditions (70 categories)
Total number of studies in which the model was applied		**n=23**	**n=19**	**n=7**
Diagnostic properties for different outcomes:				
Hospitalization	Haas, 2013,[4]	C=0.73	C=0.68	C=0.67
	Lemke, 2012,[12]	AUC=0.80	AUC=0.78	
(number of hospitalizations)	Shadmi, 2011,[16]	R^2=0.24	R^2=0.11	
(unplanned hospitalization)	Maltenfort, 2019,[11]	AUC=0.82		
	Inouye, 2008,[20]		C=0.72	
	Ou, 2011,[21]		C=0.61	
	Mosley, 2009,[25]			AUC=0.64
Emergency department visits	Haas, 2013,[4]	C=0.67	C=0.59	C=0.58
	Ou, 2011,[21]		C=0.63	
	Wallace, 2016,[22]		C=0.58	
Costs				
(Top 10% total costs)	Haas, 2013,[4]	C=0.76	C=0.70	C=0.70
(total costs)	Brilleman, 2014,[14]	R^2=0.41	R^2=0.34	
(pharmaceutical costs)	Aguado, 2008,[10]	R^2=0.39		
(total costs)	Sicras-Mainar, 2013,[13]	R^2=0.37		
(total costs)	Charlson, 2008,[18]		R^2=0.22	
(total costs)	Charlson, 2014,[19]		R^2=0.20	
(high total costs)	Ou, 2011,[21]		C=0.64	

Table 1: Continued

	First author, year	Adjusted Clinical Group (ACG)	Charlson Comorbidity Index (CCI)	Hierarchical Condition Categories (HCC)
Utilization of different healthcare services				
(GP visits)	Brilleman, 2013,[15]	R^2=0. 37	R^2=0.26	
(primary care visits)	Shadmi, 2011,[16]	R^2=0.54	R^2=0.18	
(specialist visits)	Shadmi, 2011,[16]	R^2=0.45	R^2=0.13	
(number of diagnostic imaging tests)	Shadmi, 2011,[16]	R^2=0.37	R^2=0.15	
(visits)	Sicras-Mainar, 2013,[13]	R^2=0.42		
(number of diagnoses / reasons for visit)	Sicras-Mainar, 2013,[13]	R^2=0.77		
(high outpatient visits)	Ou, 2011,[21]		C=0.63	
Input data for the model		Age, gender, diagnostic codes, pharmaceutical information, healthcare costs	Presence or absence of chronic conditions based on diagnosis codes; weighted	ICD-9 of ICD-10 diagnosis codes

AUC= Area Under the ROC Curve; C= C-statistic; R2=R square

In the grey fields diagnostic values according to the comparison studies are represented. The dark grey fields present values from comparison studies in which all three most frequently used risk stratification models are compared with each other. The studies associated with the light grey concern comparison studies comparing only two out of the three main models.

Adjusted Clinical Groups: 23 studies

The ACG is the most frequently applied risk stratification tool in our review. The ACG system is a risk stratification model designed by the Johns Hopkins University. The model was originally developed to predict and measure multimorbidity in a population. The ACG system is a measure of comorbidity and can predict utilization costs, hospitalization and emergency department visits. The model is able to use patients' data from Electronic Health Records (EHRs), insurance claims, disease registries and health status surveys(10). Minimal input data for the model are healthcare diagnoses in a specific time interval, gender and age, to which the ACG classifies people to one of 93 ACG categories. These categories represent expected healthcare utilization. In addition, different probabilities for future utilization of healthcare services are calculated. This information can be used by healthcare professionals to make informed clinical and administrative decisions (4).

Of the 23 ACG studies, eight provided statistical diagnostic values for the accuracy of the model, calculated for different outcomes. For ***prediction of hospitalization***, the model is diagnostically assessed three times with AUC and C statistic values between 0.73 and 0.82 (4,11,12). The diagnostic accuracy can be classified as *good*.

In one study a C-value of 0.67 is presented for ***prediction of emergency department visitation***, which classifies as *Sufficient*, and a C-value of 0.76 for ***prediction of high total costs***, again classifying as *good* (4). Three other studies presented R^2 values between 0.37 and 0.41 for explaining the variation of healthcare ***costs*** by the ACG model (10,13,14). ***Variations in high utilization of different healthcare services***, such as primary care visits, specialists' visits and numbers of diagnostic imaging tests, diagnoses and hospitalizations, are discussed in three studies, with R^2 values ranging from 0.24 to 0.77 (13,15,16).

ACG is highly suitable for application in primary care populations, as using *International Classification of Primary Care* (ICPC) codes as input is possible (10). ICPC codes are used to classify complaints and diagnoses of patients in many primary care settings, such as in the Netherlands. This information is stored in EHRs. The model uses other input variables such as age, gender, pharmaceutical information and previous visitation, stored in the EHR as well.

Charlson Comorbidity Index: 19 studies

The CCI is the second-most studied risk stratification model. The CCI was developed by Charlson and colleagues in 1987 and was originally an age-comorbidity index that predicted a relative risk of death within a year for hospital admitted cancer patients (17). Since that time, many adjustments have been made and in addition to mortality predictions the model is now used to predict hospitalization, emergency department visitation, future healthcare utilization and morbidity in wider populations. The system categorizes the population into six categories, based on the presence of comorbidities and chronic conditions, of which a weighted sum is provided (from zero conditions as category one to five or more conditions as category six) (18,19). The model investigates the effect of multimorbidity and predicts several outcomes. Variations of the CCI exist and the validity on predictions have been consistently investigated (4).

From the 18 studies in which the CCI or a modification was used, 10 provided statistical diagnostic values. AUC and C-values range from 0.61 to 0.78 for the prediction of future hospitalization (4,12,20,21), which correspond to an accuracy of Sufficient and Good. For emergency department visitation C-statistics between 0.58 and 0.63 are provided (4,21,22) (poor to sufficient) and for total costs, R^2 values were between 0.20 and 0.34 (14,18,19). For healthcare utilization of different healthcare services R^2 values were between 0.13 and 0.26 (15,16,23).

Input variables for the CCI include combinations of age, race, gender, mental illness, pregnancy, drug or alcohol addiction, type of health plan, type of provider, number of therapeutic classes and number of medications prescribed. The CCI is fit for use with primary care data, but focuses primarily on the absence or presence of chronic conditions, apart from other demographics. Although there is no evidence in the included studies of use of the CCI with ICPC codes, the coding system used in Dutch primary care, there is evidence for use with Read codes, a British primary care coding system (24). Possibilities to use the model with coding systems other than International Classifications of Disease (ICD) codes, are therefore very likely.

The software algorithm for CCI is published and available (4).

Hierarchical Condition Categories: 7 studies

The third most frequently studied model (n=7) is the HCC. This model was first designed and implemented by the Centers for Medicare and Medicaid Services (CMS) to adjust capitation payments for enrolees with higher risk than others. The model uses demographic data of patients as well as ICD 10th revision (ICD-10) diagnosis codes. ICD codes are used in all American healthcare service providers

(25). The ICD classification is adapted in other countries, yet these are codes most prominently used in hospital administrative registries (26). Based on this information, the model categorizes a patient into one of 70 aggregated condition categories which contributes to an individualized risk score.

For this model, four diagnostic values are provided in two studies included in this literature review. For hospitalization an AUC value of 0.64 (25), and a C-statistic of 0.67 (4), are provided. The study by Haas et al. provides a C-statistic equal to 0.58 for prediction of emergency department visitation, but a much higher C-statistic of 0.70 for prediction of high total costs (4).

A major concern regarding this model, is that it makes use of ICD codes rather than ICPC codes, making it difficult to apply in the Dutch primary care settings.

Comparison studies

A total of ten papers compared more than one risk stratification tool applied within the same study populations. However, only five articles compared more than one of the three above mentioned risk stratification tools while providing statistical diagnostic values to compare the different tools with each other.

For ***hospitalization*** the ACG performs slightly better than the CCI with AUC values of 0.80 versus 0.78 (12), and C-statistics of 0.73 versus 0.68 (4). The ACG also outperforms the CCI regarding ***emergency department visitation*** with C-statistics of 0.67 versus 0.59 and ***high total costs*** with C-statistics of 0.76 versus 0.70 (4), and R^2 values of 0.41 versus 0.34 (14). Furthermore, the study by Shadmi and colleagues showed evidence of the ACG providing better results compared to the CCI regarding other ***healthcare utilization*** outcomes, such as numbers of hospitalizations (R^2 = 0.24 versus R^2 = 0.11), primary care visits (R^2 = 0.54 versus R^2 = 0.18), specialist visits (R^2 = 0.45 versus R^2 = 0.13) and diagnostic imaging tests (R^2 = 0.37 versus R^2 = 0.15) all within a study period of 12 months (16). In addition, Brilleman and colleagues find R^2 values of 0.37 for the ACG and 0.26 for the CCI with the number of general practitioner (GP) visits as the predicted outcome (15).

Remaining risk stratification tools

In addition to the three above mentioned risk stratification tools, 28 other tools were identified within this systematic literature review. One of the 28 identified risk stratification tools is called the Elixhauser Index or Method and was mentioned in five studies. The Elixhauser Index uses a set of 30 dichotomous variables as comorbidity measures (27). Outcomes concern high utilization and pharmaceutical expenditure. One out of the five studies, mentioning the Elixhauser Index, provided C-statistics between 0.62 and 0.74 for different health utilization outcomes (21). The study by Ou and colleagues compared those C-statistics to values between 0.61 and 0.64 for the CCI (21).

A number of the identified risk stratification tools include disease or medication counts as comorbidity measures, such as the Chronic Disease Score (CDS) (n=3), which is based on dispensed drugs history. The previously mentioned study by Ou and colleagues provided C-statistic values between 0.61 and 0.72 for the CDS (21). The remainder of the identified risk stratification tools were only mentioned a few times (n=1, 2 or 3) in the articles, typically including only one validation study per risk stratification tool. The infrequent use of these tools does not make a review possible. The Clinical Risk Groups (CRG), for example, emerged three times within our systematic review. However, all studies using the CRG as a risk stratification tool were application studies and thus lacking statistical diagnostic values. Most other studies, describe a new risk stratification tool developed for a specific situation. In supplementary table 2 all of the risk stratification tools are presented, organized by included studies.

Discussion

Summary of main findings

This literature review revealed a broad range of risk stratification tools that have been assessed on accuracy and validity. The most common predicted outcomes were future hospitalization, emergency department visitation, high healthcare utilization, and total cost. The three most frequently studied risk stratification tools were the ACG, CCI and HCC.

With most AUC and C-statistic values above 0.70, the ACG performs *good* on a wide variety of outcomes. The CCI scores *sufficient* for different outcomes, with the exception of high utilization of healthcare for which a low score yielded. With most AUC and C-statistic values between 0.60 and 0.70 the HCC can also be classified as *sufficient*. Comparing the results of the ACG, the CCI and HCC, more convincing evidence for accuracy and validity is found for the ACG. Previous research also indicated the high accuracy and validation of the ACG model (12,28-30). The model is considered one of the leading models regarding the accuracy of predicting hospitalizations (12), and is widely used to gain insight in future healthcare utilization of patients (31). The study by Ou and colleagues is making a compelling case for the validity of the Elixhauser Index and the CDS, compared to the CCI (21). However, this result is not robust as it is only based on a single study. Nevertheless, the Elixhauser Index may have future potential for use in a European primary care setting.

For the applicability in primary care, evidence shows that the ACG has the possibility to make use of ICPC codes, the coding system of the (Dutch) primary care registry. The CCI has not yet been proven usable with ICPC codes. Nevertheless, evidence has shown possibilities for the CCI to be used with Read codes (24), the UK's primary care coding system, making it highly likely that the CCI can be applied using other than ICD diagnosis codes. For the HCC model on the other hand, there is no evidence to use diagnosis codes other than the ICD coding system, making it difficult to use this model in Dutch primary care.

The results of this study support the idea that risk stratification tools are suitable for primary care data in a European context. However different models emphasize various aspects within the tools. As all applications are focusing on similar utilization outcomes, such as hospitalization, ED visits and costs, the ACG has an array of other indicators developed for risk stratification. Various applications in primary care show the potential of models for example in areas of improved resource allocation (32), and care management due to better insights into vulnerable populations (33). In addition, the ACG provides possibilities to efficiently prioritize sub-populations for tailored care interventions (34).

Limitations

Although our results support risk stratification using the ACG in primary care, there are some limitations.

We only selected studies that already performed risk stratification in primary care. As a consequence we could have missed stratification tools only applied in hospital or open source data, but with a strong potential for suitability in primary care.

Selection of studies was dependant on the inter observer reliability of the two researchers. Although inclusion and exclusion criteria were clearly formulated beforehand, the possibility remains that useful tools were missed given the relatively high number of disagreements.

We assessed the identified risk stratification tools in different studies, in an attempt to compare the statistical validity of the models with each other. However, the incomparable circumstances under which different studies are performed, such as study populations and data sources, make reasonable comparisons challenging.

We based our recommendation on diagnostic values of applied risk stratification tool reported by studies published in scientific literature. Due to publication bias promising risk stratification tools may not have emerged sufficiently from our findings.

Further research

From all the articles included in this study, a small percentage explicitly defines 'risk stratification'. With the growing need for tailored care and health management approaches, a precise definition will be useful. Risk stratification and other terms such as population segmentation are now used interchangeably. Studies contributing to a generalized definition of the term *risk stratification* will be of great scientific and practical value. By using the same definition, miscommunications regarding the meaning of risk stratification will be reduced, and information on highly performing methods and implementations thereof can be shared more effectively.

With this review, we studied which risk stratification tools are best suited for the European primary care setting. However, primary care settings differ between countries. To find the best suitable tool for a specific primary care system, the performance of different tools should be investigated within the same setting, centred on desired outcomes. Based on the results of this literature review, further studies assessing the performance of desired risk stratification models, will be beneficial for Dutch primary care.

Conclusion

In conclusion, based on application frequency, statistical validity and used diagnosis coding systems, we suggest the ACG as the best model for use in European primary care settings, such as Dutch Primary Care. However, further local assessment of the ACG system is needed to ensure proper implementation.

References

1. Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. Journal of general internal medicine 2007;22 Suppl 3:391-5.
2. Luo G, Stone BL, Sakaguchi F, et al. Using Computational Approaches to Improve Risk-Stratified Patient Management: Rationale and Methods. JMIR research protocols 2015;4(4):e128.
3. Bernstein RH. New arrows in the quiver for targeting care management: high-risk versus high-opportunity case identification. J Ambul Care Manage 2007;30(1):39-51.
4. Haas LR, Takahashi PY, Shah ND, et al. Risk-stratification methods for identifying patients for care coordination. The American journal of managed care 2013;19(9):725-32.
5. Gentil ML, Cuggia M, Fiquet L, et al. Factors influencing the development of primary care data collection projects from electronic health records: a systematic review of the literature. BMC medical informatics and decision making 2017;17(1):139.
6. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS medicine 2009;6(7):e1000097.
7. Organisation for Economic Co-operation and Development. Member Countries. http://www.oecd.org/about/members-and-partners/. Accessed March, 2020.
8. Ket H. Tips & trucs voor het uitvoeren van systematische reviews met EndNote.
9. Simundic AM. Measures of Diagnostic Accuracy: Basic Definitions. Ejifcc 2009;19(4):203-11.
10. Aguado A, Guino E, Mukherjee B, et al. Variability in prescription drug expenditures explained by adjusted clinical groups (ACG) case-mix: a cross-sectional study of patient electronic records in primary care. BMC health services research 2008;8:53.
11. Maltenfort MG, Chen Y, Forrest CB. Prediction of 30-day pediatric unplanned hospitalizations using the Johns Hopkins Adjusted Clinical Groups risk adjustment system. PloS one 2019;14(8):e0221233.
12. Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. Medical care 2012;50(2):131-9
13. Sicras-Mainar A, Velasco-Velasco S, Navarro-Artieda R, et al. Obtaining the mean relative weights of the cost of care in Catalonia (Spain): retrospective application of the adjusted clinical groups case-mix system in primary health care. Journal of evaluation in clinical practice 2013;19(2):267-76.
14. Brilleman SL, Gravelle H, Hollinghurst S, et al. Keep it simple? Predicting primary health care costs with clinical morbidity measures. Journal of health economics 2014;35:109-22.
15. Brilleman SL, Salisbury C. Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study. Family practice 2013;30(2):172-8.
16. Shadmi E, Balicer RD, Kinder K, et al. Assessing socioeconomic health care utilization inequity in Israel: impact of alternative approaches to morbidity adjustment. BMC public health 2011;11:609.
17. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-83.
18. Charlson ME, Charlson RE, Peterson JC, et al. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. Journal of clinical epidemiology 2008;61(12):1234-40.

19. Charlson ME, Wells MT, Kanna B, et al. Medicaid managed care: how to target efforts to reduce costs. BMC health services research 2014;14:461.
20. Inouye SK, Zhang Y, Jones RN, et al. Risk factors for hospitalization among community-dwelling primary care older patients: development and validation of a predictive model. Medical care 2008;46(7):726-31.
21. Ou HT, Mukherjee B, Erickson SR, et al. Comparative performance of comorbidity indices in discriminating health-related behaviors and outcomes. Health Outcomes Research in Medicine 2011;2(2):e91-e104.
22. Wallace E, McDowell R, Bennett K, et al. Comparison of count-based multimorbidity measures in predicting emergency admission and functional decline in older community-dwelling adults: a prospective cohort study. BMJ open 2016;6(9):e013089.
23. Ou HT, Mukherjee B, Erickson SR, et al. Comparative performance of comorbidity indices in predicting health care-related behaviors and outcomes among Medicaid enrollees with type 2 diabetes. Population health management 2012;15(4):220-9.
24. Khan NF, Perera R, Harper S, et al. Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. BMC family practice 2010;11:1.
25. Mosley DG, Peterson E, Martin DC. Do hierarchical condition category model scores predict hospitalization risk in newly enrolled Medicare advantage participants as well as probability of repeated admission scores? Journal of the American Geriatrics Society 2009;57(12):2306-10.
26. (RIVM) RvVeM. Internationale statistische classificatie van ziekten en met gezondheid verband houdende problemen. Tiende Revisie. 2015
27. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. Medical care 1998;36(1):8-27.
28. Dominick KL, Dudley TK, Coffman CJ, et al. Comparison of three comorbidity measures for predicting health service use in patients with osteoarthritis. Arthritis and rheumatism 2005;53(5):666-72.
29. Perkins AJ, Kroenke K, Unutzer J, et al. Common comorbidity scales were similar in their ability to predict health care costs and mortality. Journal of clinical epidemiology 2004;57(10):1040-8.
30. Huntley AL, Johnson R, Purdy S, et al. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. Annals of family medicine 2012;10(2):134-41.
31. Johns Hopkins University. https://www.hopkinsacg.org/. Accessed March, 2020.
32. Kristensen T, Rose Olsen K, Sortso C, et al. Resources allocation and health care needs in diabetes care in Danish GP clinics. Health policy (Amsterdam, Netherlands) 2013;113(1-2):206-15.
33. Burton LC, Skinner EA, Uscher-Pines L, et al. Health of Medicare Advantage plan enrollees at 1 year after Hurricane Katrina. The American journal of managed care 2009;15(1):13-22.
34. Soto-Gordoa M, de Manuel E, Fullaondo A, et al. Impact of stratification on the effectiveness of a comprehensive patient-centered strategy for multimorbid patients. Health services research 2019;54(2):466-73.
35. Akazawa M, Imai H, Igarashi A, et al. Potentially inappropriate medication use in elderly Japanese patients. The American journal of geriatric pharmacotherapy 2010;8(2):146-60.

36. Beauchet O, Launay CP, Chabot J, et al. Prediction of unplanned hospital admissions in older community dwellers using the 6-item Brief Geriatric Assessment: Results from REPERAGE, an observational prospective population-based cohort study. Maturitas 2019;122:1-7.
37. Chang HY, Richards TM, Shermock KM, et al. Evaluating the Impact of Prescription Fill Rates on Risk Stratification Model Performance. Medical care 2017;55(12):1052-60.
38. Chung S, Romanelli RJ, Stults CD, et al. Preventive visit among older adults with Medicare's introduction of Annual Wellness Visit: Closing gaps in underutilization. Preventive medicine 2018;115:110-18.
39. Crane SJ, Tung EE, Hanson GJ, et al. Use of an electronic administrative database to identify older community dwelling adults at high-risk for hospitalization or emergency department visits: the elders risk assessment index. BMC health services research 2010;10:338.
40. Davis AC, Shen E, Shah NR, et al. Segmentation of High-Cost Adults in an Integrated Healthcare System Based on Empirical Clustering of Acute and Chronic Conditions. Journal of general internal medicine 2018;33(12):2171-79.
41. Dennis S, Taggart J, Yu H, et al. Linking observational data from general practice, hospital admissions and diabetes clinic databases: can it be used to predict hospital admission? BMC health services research 2019;19(1):526.
42. Duenas-Espin I, Vela E, Pauws S, et al. Proposals for enhanced health risk assessment and stratification in an integrated care scenario. BMJ open 2016;6(4):e010301.
43. Freund T, Kunz CU, Ose D, et al. Patterns of multimorbidity in primary care patients at high risk of future hospitalization. Population health management 2012;15(2):119-24.
44. Glazier RH, Agha MM, Moineddin R, et al. Universal health insurance and equity in primary care and specialist office visits: a population-based study. Annals of family medicine 2009;7(5):396-405.
45. Hamano J, Oishi A, Kizawa Y. Prevalence and Characteristics of Patients Being at Risk of Deteriorating and Dying in Primary Care. Journal of pain and symptom management 2019;57(2):266-72.e1.
46. Hewner S, Seo JY, Gothard SE, et al. Aligning population-based care management with chronic disease complexity. Nursing outlook 2014;62(4):250-8.
47. Hong CS, Atlas SJ, Ashburner JM, et al. Evaluating a Model to Predict Primary Care Physician-Defined Complexity in a Large Academic Primary Care Practice-Based Research Network. Journal of general internal medicine 2015;30(12):1741-7.
48. Hu T, Dattani ND, Cox KA, et al. Effect of comorbidities and medications on frequency of primary care visits among older patients. Canadian family physician Medecin de famille canadien 2017;63(1):45-50.
49. Hutchings HA, Evans BA, Fitzsimmons D, et al. Predictive risk stratification model: a progressive cluster-randomised trial in chronic conditions management (PRISMATIC) research protocol. Trials 2013;14:301.
50. Khanna S, Rolls DA, Boyle J, et al. A risk stratification tool for hospitalisation in Australia using primary care data. Scientific reports 2019;9(1):5011.
51. Martin S, Wagner J, Lupulescu-Mann N, et al. Comparison of EHR-based diagnosis documentation locations to a gold standard for risk stratification in patients with multiple chronic conditions. Applied clinical informatics 2017;8(3):794-809.

52. Martin Lesende I, Mendibil Crespo LI, Castano Manzanares S, et al. Functional decline and associated factors in patients with multimorbidity at 8 months of follow-up in primary care: The functionality in pluripathological patients (FUNCIPLUR) longitudinal descriptive study. BMJ open 2018;8 (7) (no pagination)(e022377)
53. Metcalfe D, Masters J, Delmestri A, et al. Coding algorithms for defining Charlson and Elixhauser co-morbidities in Read-coded databases. BMC medical research methodology 2019;19(1):115.
54. Milla-Perseguer M, Guadalajara-Olmeda N, Vivas-Consuelo D, et al. Measurement of health-related quality by multimorbidity groups in primary health care. Health and quality of life outcomes 2019;17(1):8.
55. Moran WP, Zhang J, Gebregziabher M, et al. Chaos to complexity: leveling the playing field for measuring value in primary care. Journal of evaluation in clinical practice 2017;23(2):430-38.
56. Muratov S, Lee J, Holbrook A, et al. Unplanned index hospital admissions among new older high-cost health care users in Ontario: a population-based matched cohort study. CMAJ open 2019;7(3):E537-e45.
57. Noyes K, Liu H, Temkin-Greener H. Medicare capitation model, functional status, and multiple comorbidities: model accuracy. The American journal of managed care 2008;14(10):679-90.
58. Ranstad K, Midlov P, Halling A. Active listing and more consultations in primary care are associated with shorter mean hospitalisation and interacting with psychiatric disorders when adjusting for multimorbidity, age and sex. Scandinavian journal of primary health care 2018;36(3):308-16.
59. Rohrer JE, Rasmussen N, Adamson SA. Illness severity and total visits in family medicine. Journal of evaluation in clinical practice 2008;14(1):65-69.
60. Salisbury C, Johnson L, Purdy S, et al. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. The British journal of general practice : the journal of the Royal College of General Practitioners 2011;61(582):e12-21.
61. Sibley LM, Moineddin R, Agha MM, et al. Risk adjustment using administrative data-based and survey-derived methods for explaining physician utilization. Medical care 2010;48(2):175-82.
62. Sibley LM, Glazier RH. Evaluation of the equity of age-sex adjusted primary care capitation payments in Ontario, Canada. Health policy (Amsterdam, Netherlands) 2012;104(2):186-92.
63. Sicras-Mainar A, Serrat-Tarres J, Navarro-Artieda R, et al. Adjusted Clinical Groups use as a measure of the referrals efficiency from primary care to specialized in Spain. European journal of public health 2007;17(6):657-63.
64. Sicras-Mainar A, Velasco-Velasco S, Navarro-Artieda R, et al. Adaptive capacity of the Adjusted Clinical Groups Case-Mix System to the cost of primary healthcare in Catalonia (Spain): a observational study. BMJ open 2012;2(3).
65. Sino CG, Stuffken R, Heerdink ER, et al. The association between prescription change frequency, chronic disease score and hospital admissions: a case control study. BMC pharmacology & toxicology 2013;14:39.
66. Snooks H, Bailey-Jones K, Burge-Jones D, et al. Effects and costs of implementing predictive risk stratification in primary care: a randomised stepped wedge trial. BMJ quality & safety 2019;28(9):697-705.

67. Sternberg SA, Bentur N, Abrams C, et al. Identifying frail older people using predictive modeling. The American journal of managed care 2012;18(10):e392-7.
68. Takahashi PY, Ryu E, Olson JE, et al. Hospitalizations and emergency department use in Mayo Clinic Biobank participants within the employee and community health medical home. Mayo Clinic proceedings 2013;88(9):963-9.
69. Vest JR, Menachemi N, Grannis SJ, et al. Impact of Risk Stratification on Referrals and Uptake of Wraparound Services That Address Social Determinants: A Stepped Wedged Trial. American journal of preventive medicine 2019;56(4):e125-e33.
70. Violan C, Plana-Ripoll O, Foguet-Boreu Q, et al. Relationship between efficiency and clinical effectiveness indicators in an adjusted model of resource consumption: a cross-sectional study. BMC health services research 2013;13:421.
71. Vivas-Consuelo D, Uso-Talamantes R, Trillo-Mata JL, et al. Predictability of pharmaceutical spending in primary health services using Clinical Risk Groups. Health policy (Amsterdam, Netherlands) 2014;116(2-3):188-95.
72. Vuik SI, Mayer E, Darzi A. Enhancing risk stratification for use in integrated care: a cluster analysis of high-risk patients in a retrospective cohort study. BMJ open 2016;6(12):e012903.
73. Wennberg JE, Staiger DO, Sharp SM, et al. Observational intensity bias associated with illness adjustment: cross sectional analysis of insurance claims. BMJ (Clinical research ed) 2013;346:f549.
74. Xu J, Williams-Livingston A, Gaglioti A, et al. A Practical Risk Stratification Approach for Implementing a Primary Care Chronic Disease Management Program in an Underserved Community. Journal of health care for the poor and underserved 2018;29(1):202-13.
75. Zhou YY, Wong W, Li H. Improving care for older adults: a model to segment the senior population. The Permanente journal 2014;18(3):18-21.

Supplementary Material

Supplementary table 1: Search strategy

((((”risk”[tw] OR ”risks”[tw] OR ”Risk”[Mesh]) **AND** (”stratification”[tw] OR ”stratifications”[tw] OR ”stratified”[tw] OR ”stratify”[tw] OR ”stratifies”[tw] OR ”stratifying”[tw])) **AND** (”model”[tw] OR ”models”[tw] OR ”tool”[tw] OR ”tools”[tw] OR ”method”[tw] OR ”methods”[tw] OR ”measure”[tw] OR ”measures”[tw] OR ”measured”[tw] OR ”measuring”[tw] OR ”algorithm”[tw] OR ”algorithms”[tw] OR ”metric”[tw] OR ”metrics”[tw] OR ”score”[tw] OR ”scores”[tw] OR ”scoring”[tw] OR ”index”[tw] OR ”indexes”[tw] OR ”indices”[tw] OR ”indexed”[tw] OR ”count”[tw] OR ”counts”[tw])) OR ”adjusted clinical groups”[tw] OR ”Minnesota tiering”[tw] OR ”Hierarchical Condition Categories”[tw] OR ”elder risk assessment index”[tw] OR ”chronic condition count”[tw] OR ”Charlson comorbidity index”[tw] OR ”chronic disease score”[tw]) **AND** (”General Practitioners”[Mesh] OR ”General Practitioner”[tw] OR ”General Practitioners”[tw] OR ”physician”[tw] OR ”physicians”[tw] OR ”general practice”[tw] OR ”General Practice”[Mesh] OR ”GP”[ti] OR ”GPs”[ti] OR ”GP’s”[ti] OR ”Primary Health Care”[Mesh] OR ”Primary Health Care”[tw] OR ”primary care”[tw] OR ”Family doctor”[tw] OR ”family doctors”[tw] OR ”Family practice”[tw] OR ”Family practices”[tw] OR ”Family medicine”[tw] OR ”general medicine”[tw] OR ”Accountable Care Organizations”[Mesh] OR ”ACO”[ti] OR ”ACOs”[tw] OR ”ACO’s”[tw] OR ”care organization”[tw] OR ”care organisation”[tw] OR ”care organizations”[tw] OR ”care organisations”[tw] OR ”Ambulatory Care Facilities”[Mesh:NoExp] OR ”Community Health Centers”[Mesh:NoExp] OR ”health center”[tw] OR ”health centers”[tw] OR ”health centre”[tw] OR ”health centres”[tw] OR ”Health Maintenance Organizations”[Mesh] OR ”maintenance organization”[tw] OR ”maintenance organisation”[tw] OR ”maintenance organizations”[tw] OR ”maintenance organisations”[tw] OR ”HMO”[ti] OR ”HMOs”[ti] OR ”HMO’s”[ti] OR ”MCO”[ti] OR ”MCOs”[ti] OR ”MCO’s”[ti] OR ”managed care”[tw] OR ”integrated care”[tw])

Filters: Date 2007-2019, Language English

Supplementary table 2: Overview of included articles. Study-population, Outcome measure, Risk stratification model, Study type, Country and Journal of publication are shown.

Reference	Population	Outcome	Model	Study type	Country	Journal
(Aguado et al., 2008) [10]	Patients from five primary care centres n=65,630	Explained variability in drug expenditures	Adjusted Clinical Groups	Validation study	Spain	BMC Health Services Research
(Akazawa, Imai, Igarashi, & Tsutani, 2010) [35]	Adults aged 65 years old or older n=6,628	Incidence, healthcare utilization and costs associated with Potentially Inappropriate Medication Use	Elixhauser Classification System, Modified Beers criteria	Application study	Japan	The American Journal of Geriatric Pharmacotherapy
(Beauchet et al., 2019) [36]	Community dwelling adults aged 80 years or older, who visited a participating general practice within the study period n = 668	Hospital admissions (6 month follow-up)	6-item Brief Geriatric Assessment	Validation study	France	Maturitas
(Brilleman & Salisbury, 2013) [15]	Primary care population of general practice n=95,372	Three year mortality and consultation rate	Charlson Comorbidity Index, Adjusted Clinical Groups, Quality and Outcomes Framework	Validation and comparison study	UK	Oxford University Press
(Brilleman et al., 2014) [14]	Primary care population of general practice n=86,100	Primary health costs	Charlson Comorbidity Index, Adjusted Clinical Groups, Quality and Outcomes Framework	Validation and comparison study	UK	Journal of Health Economics
(Burton et al., 2009) [33]	Non-institutionalized Peoples Health (Managed Care Organization) managed care beneficiaries n=20,612	Morbidity level	Adjusted Clinical Groups	Application study	USA	The American Journal of Managed Care

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Chang et al., 2017) [37]	Primary care population of the Health Partners Network n=43,097	Effect of prescription fill rates on risk stratification model performance	Adjusted Clinical Groups	Comparison study	USA	Medical Care
(Charlson et al., 2008) [18]	Primary care population of a general practice in an academic hospital n=5,861	Predict costs of disease in primary care patients	Adapted Charlson Comorbidity Index	Validation study	USA	Journal of Clinical Epidemiology
(Charlson et al., 2014) [19]	Medicaid managed care beneficiaries, adults and children, who received primary care in a specific medical center, not including Medicare/ Medicaid patients (In this review presented results concern adult population) n = 4,614 (2,218 adults and 2,396 children)	Identification of potential high cost beneficiaries	Charlson Comorbidity Index	Application study	USA	BMC Health Services Research
(Chung, Romanelli, Stults, & Luft, 2018) [38]	Medicare beneficiaries, aged 65 to 85 years, who were primary care patients in a large, mixed payer outpatient healthcare organization n=108,734	Preventive primary care visits	Own model including Charlson Comorbidity Index as comorbidity measure, age category, visit frequency and insurance	Application study	USA	Preventive Medicine

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Crane et al., 2010) [39]	Older community dwelling adults of a general practice n=12,650	High-Risk Hospitalization or ED admission	The Elders Risk Assessment	Validation study	USA	BMC Health Services Research
(Davis et al., 2018) [40]	Patients from an integrated healthcare delivery system and health plan, with continuous coverage; n = 2,118,343	Health care costs	Hierarchical Condition Categories	Application study	USA	Journal of General Internal Medicine
(Dennis et al., 2019) [41]	Patients with type 2 diabetes who had been treated at a general practice n = 161,575	Hospital admission	Own prediction model	Validation study	Australia	BMC Health Services Research
(Duenas-Espin et al., 2016) [42]	Primary care populations n=2,100,000 (B) n=7,500,000 (C) n=100,000) (L) n=3,400,000 (Sc)	Health risk assessment	Adjusted Clinical Groups, Adjusted Morbidity Groups known as GMA, Clinical Risk Group	Application and comparison study	Spain, Italy, Scotland	BMJ Open
(Freund, Kunz, Ose, Szecsenyi, & Peters-Klimm, 2012) [43]	Primary care population n=6,026	Prediction risk future hospitalization	Hierarchical Condition Categories	Application study	Germany	Population Health Management
(Glazier, Agha, Moineddin, & Sibley, 2009) [44]	Primary care population n=25,558	Diagnosed health status	Adjusted Clinical Groups	Application study	Canada	Annals of Family Medicine

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Haas et al., 2013) [4]	Adult patients empaneled in 2009 and 2010 in a primary care practice n=83,187	Hospitalization, emergency department visits, 30-day readmission, high expenditures	Adjusted Clinical Group, Hierarchical Condition Categories, Elder Risk Assessment, Chronic Comorbidity Count, Charlson Comorbidity Index, Minnesota Tiering	Comparison study	USA	The American Journal of Managed Care
(Hamano, Oishi, & Kizawa, 2019) [45]	Primary care patients aged 65 years or older, who visited a participating general practice within the study period; n = 382	Detoriation and dead	Supportive and Palliative Care Indicators Tool	Validation study	Japan	Journal of Pain and Symptom Management
(Hewner, Seo, Gothard, & Johnson, 2014) [46]	Primary care population of Medicare, Medicaid and privately insured n=477,407	Risk-stratified cohorts based on chronic disease and complexity	COMPLEXedex clinical algorithm	Application study	USA	Nursing Outlook
(Hong et al., 2015) [47]	Primary care adult patients in a practice-based research network n=143,372	Prediction of complexity	Outpatient Charlson Score & Commercial Risk Prediction	Validation and comparison study	USA	Journal of General Internal Medicine
(Hu et al., 2017) [48]	Primary care population n=265	Predictors of frequent visits to family physicians	Charlson Comorbidity Index, Beers Criteria	Application study	Canada	Canadian Familiy Physician
(Huntley, Johnson, Purdy, Valderas, & Salisbury, 2012) [30]	Primary care population	Care utilization, costs, mortality, quality of life	Adjusted Clinical Groups, Charlson Comorbidity Index, Chronic Disease Score, Cumulative Illness Rating Scale, Duke Severity Index	Review	-	Annals of Family Medicine

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Hutchings et al., 2013) [49]	Primary care population n=2,400	Estimate effects on the delivery of care, patient satisfaction, quality of life and resources used.	An emergency admission risk prediction tool called PRISM	Validation study	Wales	Trial Journal
(Inouye et al., 2008) [20]	Community dwelling elderly aged 70 years and older in primary care clinics of an academic medical center n=3,919	Unplanned hospitalization in one year	Deyo-Charlson,	Validation study	USA	Medical Care
(Khan et al., 2010) [24]	Primary care cancer patients and healthy controls n=146,441	Mortality	Adapted Charlson Score for use with Read/ OXMIS instead of ICD10 diagnosis codes	Validation study	UK	BMC Family Practice
(Khanna et al., 2019) [50]	Patients who attended their primary care clinics at least once n = 393,229	1 year hospitalization, ED visit	Own model	Validation study	Australia	Scientific Reports
(Kristensen et al., 2013) [32]	Primary care patients with type 2 diabetes; n = 6,706	Fee for service costs	Adjusted Clinical Groups	Application study	Denmark	Health Policy
(Lemke et al., 2012) [12]	Primary care population n=4,700,000	Predicting hospitalization	Adjusted Clinical Groups, Charlson Comorbidity Index	Validation study	USA	Medical Care
(Maltenfort et al., 2019) [11]	Children seen in a large primary and specialty care outpatient network n= 920,051 (70% for training and 30% for testing the model)	Unplanned 30-day hospitalization	Own model with predictors derived from the Adjusted Clinical Groups	Validation study	USA	PLoS One

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Martin et al., 2017) [51]	Patients from three primary care clinics participating in the Integrated Care Coordination Information System study data set, who had at least one of the selected conditions and were seen from 2008 to 2012 n = 750	ED visits, hospitalization and heathcare costs	A modified Charlson Comorbidity Index, Hierarchical Condition Categories, count of chronic conditions defined by Affordable Care Act	Application study	USA	Applied Clinical Informatics
(Martin Lesende et al., 2018) [52]	Patients aged 65 years and older n = 241	Top 5% of the 'Kaiser Permanente pyramid'	Adjusted Clinical Groups	Application study	Spain	BMJ Open
(Metcalfe et al., 2019) [53]	Patients from primary care practices: age- and sex-matched controls (n=26,860) for hip fractured patients (n = 13,974)	Mortality: 30-day and 1-year	Charlson Comorbidity Index, Elixhauser method	Validation study	UK	BMC Medical Research Methodology
(Milla-Perseguer, Guadalajara Olmeda, Vivas-Consuelo, & Uso-Talamantes, 2019) [54]	All citizens registered in a specific health district n = 32,667	Morbidity measure	Clinical Risk Groups	Application study	Spain	Health and Quality of Life Outcomes
(Moran et al., 2017) [55]	Primary care population (Medical University) n=10,408	Clustering patients with stratification on risk for hospital and ED utilization	Own model	Validation study	USA	Journal of Evaluation in Clinical Practice
(Mosley et al., 2009) [25]	Primary care population, Medicare beneficiaries; n = 4,506	Hospitalization	Hierarchical Condition Categories	Application study	USA	Journal of the American Geriatrics Society

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Muratov et al., 2019) [56]	Adults aged 66 years old or more High cost users (n=175,847) and age and sex matched non-high cost users (n=527,541)	Hospital admission	Adjusted Clinical Groups	Validation study	Canada	CMAJ Open
(Noyes, Liu, & Temkin Greener, 2008) [57]	Community dwelling (not institutionalized for 90 days at a time) Medicare beneficiaries with continuous part A and B enrollment for at least two calender years n = 46,790	Healthcare Costs	Hierarchical Condition Categories	Validation study	USA	American Journal of Managed Care
(Ou et al., 2011) [21]	Medicaid enrolees with type 2 diabetes n=9,832	Health care behaviors (physician's diabetes .adherence standard adherence, patient's medication adherence), health care utilization and expenditures	A modified version of Romano-adapted Charlson index, Elixhauser index, Chronic Disease Score, Health-related Quality of Life Comorbidity Index	Validation and Comparison study	USA	Health Outcomes Research in Medicine
(Ou et al., 2012) [23]	Medicaid enrolees with type 2 diabetes n=9,832	Health care utilization and expenditures	A modified version of Romano-adapted Charlson index, Elixhauser index, Chronic Disease Score, Health-related Quality of Life Comorbidity Index	Validation and Comparison study	USA	Population Health Management

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Ranstad, Midlov, & Halling, 2018) [58]	Primary care listed patients n = 151,731	Hospitalization (binary) and number of days hospitalized	Own model with doctor-patient relationship, contribution of complex diagnosis patterns (psychiatric disorders) and morbidity burden (Adjusted Clinical Groups based)	Application study	Sweden	Scandinavian Journal of Primary Health Care
(Rohrer, Rasmussen, & Adamson, 2008) [59]	Primary care population n=698	High utilization, Illness severity	Charlson Comorbidity Index	Application study	USA	Journal of Evaluation in Clinical Practice
(Salisbury, Johnson, Purdy, Valderas, & Montgomery, 2011) [60]	Primary care population n=99,997	Multimorbidity	Adjusted Clinical Groups, Quality of Outcome Framework	Application study	UK	British Journal of General Practice
(Shadmi et al., 2011) [16]	Adult enrolees of Clalit Health Services, Israel's largest health care organization n = 279,241	Numbers of (1) primary care encounters, (2) specialist visits, (3) diagnostic imaging tests and (4) hospitalizations	Adjusted Clinical Groups, Charlson Comorbidity Index	Application, validation and comparison study	Israel	BMC Public Health
(Sibley, Moineddin, Agha, & Glazier, 2010) [61]	Primary care patients n=25,558	Predicting physician utilization	Adjusted Clinical Groups	Application study	Canada	Medical Care
(Sibley & Glazier, 2012) [62]	Primary care data (family health networks) n=487,131	Expected healthcare utilization	Adjusted Clinical Groups	Application study	Canada	Health Policy
(Sicras-Mainar et al., 2007) [63]	Patients attending five primary care teams n = 81,335	Referral Rate	Adjusted Clinical Groups	Application study	Spain	European Journal of Public Health

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Sicras-Mainar et al., 2012) [64]	Patients from 13 primary care teams (86,5% primary care, 13,5% peadiatrics) n = 227,235	Heatlh costs	Adjusted Clinical Groups	Application study	Spain	BMJ Open
(Sicras-Mainar et al., 2013) [13]	Patients from 13 primary care teams n = 227,235	Explaining variance of: Visits, number of diagnoses, total costs	Adjusted Clinical Groups	Validation Study	Spain	Journal of Evaluation in Clinical Practice
(Sino et al., 2013) [65]	Hospitalized (cases) and matched non-hospitalized patients from pharmacy registries n = 2 x 8,681 (cases and controls)	Prediction of hospital admission	Combining Prescription Changes Frequency, Chronic Disease Score	Application study	the Netherlands	BMC Pharmacology and Toxicology
(Snooks et al., 2019) [66]	Patients from primary care practices n = 230,099	Unscheduled hospital admission	An emergency admission risk prediction tool called PRISM	Application study	UK	BMJ Quality & Safety

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Soto-Gordoa et al., 2019) [34]	Patients with multimorbidity (at least two out of three conditions: diabetes, heart failure and chronic obstructive pulmonary disease) aged 65 years or more n = 4,225	Hospital admission	Adjusted Clinical Groups	Application study	Spain	Health Services Research
(Sternberg et al., 2012) [67]	Community dwelling elderly n=221	Predict hospitalizations and emergency department visitation	Adjusted Clinical Groups, Vulnerable Elderly Survey	Application study	Israel	American Journal of Managed Care
(Takahashi et al., 2013) [68]	Primary care population ECH Mayo clinic biobank n=22,916	Predict hospitalizations and emergency department visitation	Minnesota Tiering	Application study	USA	Mayo Clinic Proceedings
(Vest et al., 2019) [69]	Patients from primary care clinics linked to a public hospital: intervention group using risk stratification modelling (n=62,254) versus control group n = 175,833	Risk scores for needing a referral to different wraparound services such as behavioural health services, dietitian councelling and social services.	A machine learning algorithm, Elixhauser score	Application study	USA	The American Journal of Preventive Medicine
(Violan et al., 2013)[70]	Primary care population of 13 Primary Healthcare Centers n=196,593	Efficiency and effectiveness indicators on resource consumption	Adjusted Clinical Groups	Application study	Spain	BMC Health Service Research

Supplementary table 2: Continued

Reference	Population	Outcome	Model	Study type	Country	Journal
(Vivas-Consuelo et al., 2014) [71]	Primary care population of general practice n=261,054	Predict pharmaceutical spending	ATC model*, Clinical Risk Groups	Application study	Spain	Health Policy
(Vuik, Mayer, & Darzi, 2016) [72]	Primary and Secondary care population; n=300,000	Identification high risk patients	Create risk scores, no tools mentioned	Validation study	UK	BMJ Open
(Wallace et al., 2016) [22]	Older community dwelling adults aged 70 years or older n = 862	Predicting emergency hospital admission	Total disease count, Selected conditions disease count, Charlson Comorbidity Index, number of dispensed medication classes, RxRisk-V	Comparison study	Ireland	BMJ Open
(Wennberg et al., 2013) [73]	Medicare beneficiaries n=5,153,877	Predicted one-year mortality	Charlson Comorbidity Index, Lezzoni chronic condition count, Hierarchical Condition Categories	Application study	USA	BMJ
(Xu, Williams-Livingston, Gaglioti, McAllister, & Rust, 2018) [74]	Patients aged 18 year and older, seen in an urban academic family medicine clinic over a two-year period; n=5,364	Utilization (high costs)	Elixhauser	Application study	USA	Journal of Health Care for the Poor and Underserved
(Zhou, Wong, & Li, 2014) [75]	Elderly aged 65 years and older n = 91,189	Mortality, hospitalization and costs of care	Senior Segmentation Algorithm	Validation study	USA	Permanente Journal

**ATC model categorizes patients into nine categories based on Anatomic Therapeutical Chemical (ATC) codes for medications.*

CHAPTER 3

CHAPTER 3

ASSESSMENT OF THE ADJUSTED CLINICAL GROUPS SYSTEM IN DUTCH PRIMARY CARE USING ELECTRONIC HEALTH RECORDS: A RETROSPECTIVE CROSS-SECTIONAL STUDY

Shelley-Ann M. Girwar, Marta Fiocco, Stephen P. Sutch, Mattijs E. Numans, Marc A. Bruijnzeels

Published in BMC Health Services Research. 2021 Mar 10;21(1):217. doi: 10.1186/s12913-021-06222-9

Abstract

Introduction

Within the Dutch health care system the focus is shifting from a disease oriented approach to a more population based approach. Since every inhabitant in the Netherlands is registered with one general practice, this offers a unique possibility to perform Population Health Management analyses based on general practitioners' (GP) registries. The Johns Hopkins Adjusted Clinical Groups (ACG) System is an internationally used method for predictive population analyses. The model categorizes individuals based on their complete health profile, taking into account age, gender, diagnoses and medication. However, the ACG system was developed with non-Dutch data. Consequently, for wider implementation in Dutch general practice, the system needs to be validated in the Dutch healthcare setting. In this paper we show the results of the first use of the ACG system on Dutch GP data. The aim of this study is to explore how well the ACG system can distinguish between different levels of GP healthcare utilization.

Methods

To reach our aim, two variables of the ACG System, the Aggregated Diagnosis Groups (ADG) and the mutually exclusive ACG categories were explored. The population for this pilot analysis consisted of 23,618 persons listed with five participating general practices within one region in the Netherlands. ACG analyses were performed based on historical Electronic Health Records data from 2014 consisting of primary care diagnoses and pharmaceutical data. Logistic regression models were estimated and AUC's were calculated to explore the diagnostic value of the models including ACGs and ADGs separately with GP healthcare utilization as the dependent variable. The dependent variable was categorized using four different cut-off points: zero, one, two and three visits per year.

Results

The ACG and ADG models performed as well as models using International Classification of Primary Care chapters, regarding the association with GP utilization. AUC values were between 0.79 and 0.85. These models performed better than the base model (age and gender only) which showed AUC values between 0.64 and 0.71.

Conclusion

The results of this study show that the ACG system is a useful tool to stratify Dutch primary care populations with GP healthcare utilization as the outcome variable.

Introduction

With rising health care utilization and costs, a shift from disease oriented to population based approaches is being advocated worldwide. With the upcoming need for improved organization and management of healthcare and the increasing possibilities of big data, strategies based on health registry analyses are becoming popular. One use of health registry data in population health management strategies is risk stratification. With risk stratification, differences in individual health risks can be screened for, and used to assign interventions to the population and individuals that will benefit the most. With rising pressure on medical services provided by general practitioners (GPs) in most European countries (1), primary care can benefit from proven advantages of risk stratification approaches, such as improved care management (2), resource allocation (3) and identification of sub-populations for tailored care interventions (4).

Despite the proven benefits of using risk stratification, especially in primary care, there is no evidence for application of internationally used risk stratification tools in Dutch primary care. Risk stratification approaches using Dutch GP registry data can be especially beneficial due to the gatekeeper's function of Dutch GPs, providing the opportunity to overview a near total population.

Different tools for risk stratification are used worldwide, amongst which the Adjusted Clinical Groups (ACG) tool developed by the Johns Hopkins University. The ACG system is an internationally used tool for risk stratification on a generic level and is one of the most frequently used risk stratification tools in primary care. Evidence has also shown stronger statistical validity for the ACG compared with other risk stratification tools, regarding predictions of different healthcare utilization outcomes (5-7).

The ACG system uses registered diagnoses over a twelve month period, to assign individuals to one of 98 ACG categories, based on their healthcare profiles and expected health utilization (8). ACG categories are based on combinations of diagnoses types. Registered diagnoses processed by the ACG system, can include the International Classification of Primary Care (ICPC) coded (9), a commonly used registration method for diagnoses in primary care (10).

In this study we explored the potential use of Johns Hopkins University ACG System in routine registration data extracted from Dutch primary care practices. The aim of this study is to explore how well the ACG system, compared to the 17 chapters of the ICPC coding system, can distinguish between different levels of GP healthcare utilization in Dutch general practice registries.

Methods

Study design and data

For this retrospective cross-sectional study, we used data from patients registered with one of the five participating GP practices during the whole of 2014 in Nijkerk, the Netherlands. Data for 30,596 patients over the year 2014extracted from the practices' electronic health records.included age, gender, and coded healthcare procedures, diagnoses and pharmaceutical data. Diagnoses were registered as ICPC-1 diagnoses codes, as used in the Netherlands (11) and converted to ICPC-2 codes. Prescribed medication was registered as Anatomical Therapeutic Chemical (ATC) codes (12),GP visits were defined as all GP encounters, including physical and telephone consults and home visits by either GPs or nurse practitioners working at the GP practices.

From the original datasets 4,289 cases were removed, due to corrupted patient identification numbers. Another 2,689 cases belonging to three specific ACG categories, were left out of the analyses: *No Diagnosis or Only Unclassified Diagnosis* (n=281), *Non-Users* (n=2,407) and *Invalid Age or Date of Birth* (n=1). The final analyses were performed with data for 23,618 persons (77% of 30,596 registered people).

Data preparation and analyses were performed with IBM SPSS Statistics 24.

ACG System software

We used the Johns Hopkins University's ACG® System software 11. The ACG® System software 11 is a risk stratification tool, assigning each patient to one of the 98 mutually exclusive ACG categories. Assignment to ACG categories is based on combinations of diagnoses types. With the ACG system the diagnoses for each patient are grouped into 32 Aggregated Diagnosis Groups (ADGs), based on type of diagnoses rather than on specific diagnoses, i.e. specific ICPC codes. Individuals' patterns of ADGs determine the assignment of patients to one of the 98 mutually exclusive ACG categories (8).

Assessment of the ACG system

To assess the applicability of the ACG system in Dutch primary care, we looked at two aspects: face validity and model performance.

Face Validity

According to Mosier (13) an important aspect of the testing of an instrument lies in the 'consumer acceptance'. The first step in effective use of a test, is the actual selection for use and acceptance of the results. Mosier describes one of the translations of face validity as the appearance of validity: the test must appear valid in addition to the statistical validity. In this study we defined face validity as this appearance of validity described by Mosier (13).

We assessed the ACG system's face validity by exploring the actual ACG categorization with regard to age. Face validity was assessed on recognition of multimorbidity in relation to age within ACG categories. The ACG categories are grouped according to the number of ADGs: one, two to three, four to five, six to nine and lastly ten plus ADGs.

Model Performance

To investigate the impact of the ACG system in Dutch primary care, four different logistic regression models were estimated.

Dependent variable

The outcome variable, number of GP visits, was transformed into binary variables according to four definitions. According to the first definition, *no GP visits* was defined as *no utilization of care*, whereas *one or more GP visits* were defined as *utilization of care*. With the second definition, a distinction between *zero or one GP visit* and *two or more GP visits* was made. With the third definition, a distinction between *zero to two GP visits* and *three or more GP visits* was made. Accordingly, for the final definition the outcome was defined as a distinction between *zero to three* and *four or more GP visits*. The performance of each of these models was investigated.

Independent variables

In the null or base model only *age* as a continuous variable and gender were included as explanatory variables.

Model 1 included age, gender and ICPC chapters as independent variables. ICPC diagnosis codes are divided into 17 different chapters including *'General and unspecified'*, '*Blood, blood forming organs, lymphatics, spleen'*, *'Digestive'*, *'Eye'*, *'Ear'*, *'Circulatory'*, *'Musculoskeletal'*, *'Neurological'*, *'Psychological'*, *'Respiratory'*, *'Skin'*, *'Endocrine metabolic and nutritional'*, *'Urology'*, *'Pregnancy, childbirth, family planning'*, *'Female genital system and breast'*, *'Male genital system'* and *'Social problems'*. Different ICPC chapters can be registered to a single person. Therefore, the ICPC chapters were added to the model as 17 different dummy variables.

Model 2 included age, gender and ADG diagnoses as independent variables. As an individual can have more than one ADG, the 32 ADGs were added to the model as 32 dummy variables.

Model 3 included age, gender and mutually exclusive ACGs. Before estimating the logistic regression, the numbers of individuals in each ACG category were checked. Aggregation of some ACG categories was necessary due to categories with small numbers of individuals. In supplementary table 1 the aggregation of the original ACG categories is presented.

To select the best model, the performance of each logistic regression with outcome variable as defined above, was investigated. The Area Under the Curve (AUC) values were calculated for each model.

Ethics approval and patients' consent

The need for ethical approval was waived by the medical ethical committee of Leiden University Medical Center (CME - LUMC), the Netherlands.

Participants were not asked for their consent because we used routinely collected de-identified data.

Results

Population characteristics

A total of 23,618 patients registered with a GP, were included in this study. 48.1% of the patients was male. The mean age of the included patients was 41.8 years old with a standard deviation of 22.2 years. 67.7% of the patients had at least one GP visit in 2014. The mean number of GP visits was 3.5 with a standard deviation of 5.0 and the maximum number of GP visits was 92. In figure 1 the distribution of the number of GP visits within the study population is presented. As expected, this is a skewed distribution, where most of the population has had zero or one GP visits.

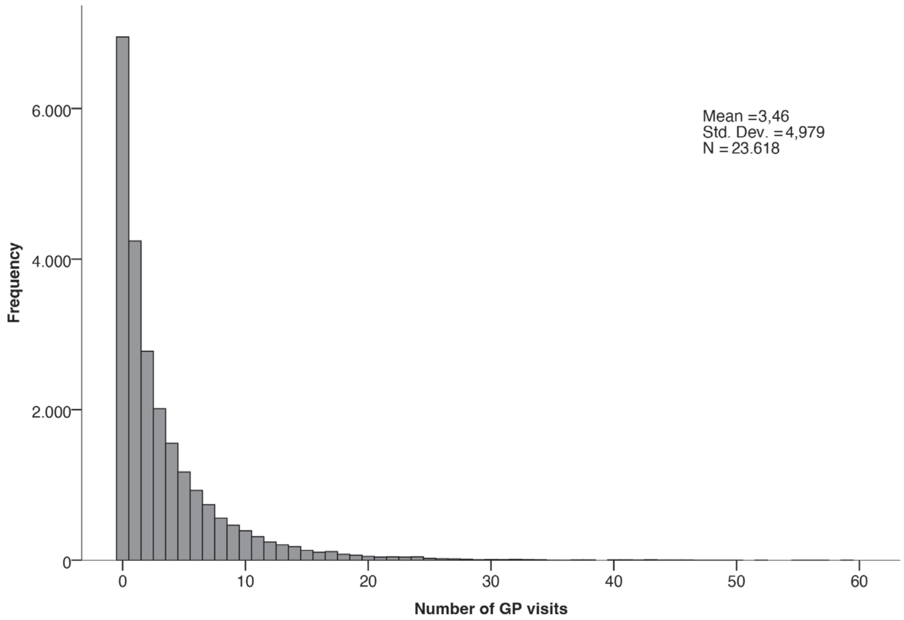


Figure 1: Distribution of the number of general practitioner (GP) visits within the study population

Figure 2 shows the health problems within the study population according to the 17 chapters of the ICPC registry system. The percentages of the study population with at least one diagnosis code corresponding to a specific ICPC chapter, are presented in the figure. ICPC chapters *Musculoskeletal* (L), *Respiratory* (R) and *Skin* (S) had the highest frequencies, with percentages between 43 and 49.

Face validity of ACG categorization

In figure 3 the distribution of age within each ACG category is presented with boxplots. Each group of ACGs corresponds with a different color, red being the highest numbers of ADGs. The figure shows that the number of ADGs gradually goes up with increasing age. Mean ages of the ACG categories with only one ADG *(green)* are mostly under 30. Exceptions are the ACG categories *Chronic medical: Stable* and *Eye/*Dental, which have mean values above 50. The mean age of ACGs with two to three ADGs *(yellow)* is mostly between 30 and 40, with the exception of ACG category *Acute Minor and Chronic Medical: Stable* (mean age of 50+). For three out of four of the ACG categories with four to five ADGs, the mean ages are between 50 and 62. However, the ACG category *Acute Minor/Acute Major/Likely Recur/Psychosocial* has a mean age of under 40. The ACG categories with six to nine ADGs have a mean age of around 63, whereas the mean age of ACG categories with ten or more ADGs is above 70. An extended overview of individuals from each ACG category, distributed over 10 year age bands, is presented in supplementary table 2.

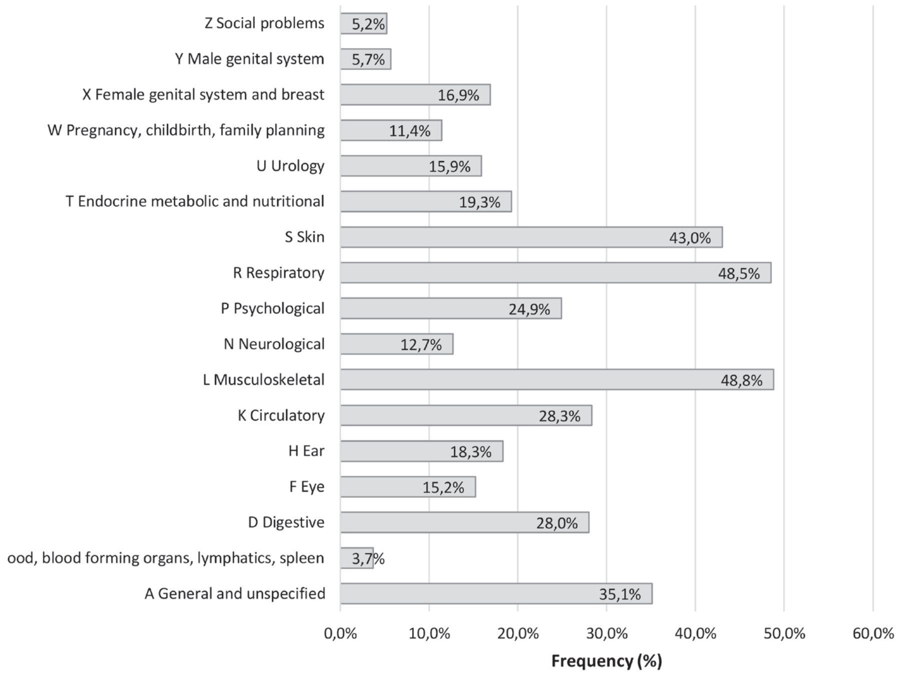


Figure 2: Overview of health problems within the study population according to the 17 main chapters of the International Classification of Primary Care (ICPC) coding system. ICPC chapters form the basis of the ICPC coding system.

Model performance

To investigate the model performances, where the outcome variable utilization of GP was defined as discussed in the methods section, AUCs along with their confidence intervals were computed.

Table 2 displays the model performances for each of the four different definitions of the outcome GP utilization. As seen in the table, model 1 and 2 perform well with AUC values between 0.79 and 0.85. They slightly perform better than model 3 with AUC values between 0.77 and 0.83. All three models outperform the null model with AUC values between 0.63 and 0.71. For all independent variables, odds ratios along with their 95% confidence intervals, are shown in supplementary tables 3 to 5.

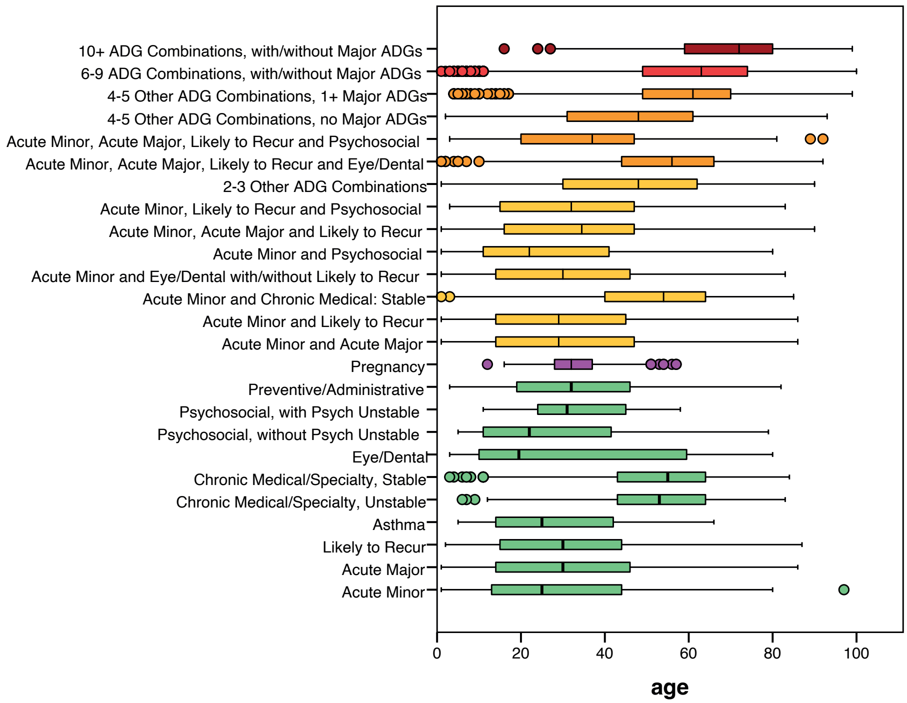


Figure 3: Age distribution per Adjusted Clinical Groups (ACG) category. ACG categories are a collapsed version of the original ACGs. Colors (with the exception of the pink one) correspond to the number of Aggregated Diagnostic Groups (ADGs): *green*=One ADG; *pink*=Pregnancy (all numbers of ADGs); *yellow*=2-3 ADGs; *orange*=4-5 ADGs; *red*=6-9 ADGs; *dark red*=10+ ADGs.

Table 3: Model performances quantified by the Area Under the ROC Curve (AUC) values along with the 95% confidence intervals (CI)

	Area Under the ROC Curve (95% Confidence Interval)			
Outcome	**Null model**	**Model 1**	**Model 2**	**Model 3**
0 vs. >=1 GPvisits	0.638 (0.630 - 0.645)	0.787 (0.781 - 0.793)	0.793 (0.787 - 0.799)	0.774 (0.768 - 0.780)
0-1 vs. >=2 GPvisits	0.675 (0.668 - 0.681)	0.816 (0.810 - 0.821)	0.818 (0.812 - 0.823)	0.799 (0.794 - 0.805)
0-2 vs. >=3 GPvisits	0.693 (0.686 - 0.700)	0.833 (0.828 - 0.838)	0.832 (0.828 - 0.837)	0.814 (0.809 - 0.819)
0-3 vs. >=4 GPvisits	0.711 (0.704 - 0.718)	0.848 (0.842-0.853)	0.848 (0.842 - 0.853)	0.829 (0.824 - 0.834)

'Outcome' is based on the four definitions of the outcome general practice (GP) healthcare utilization

Discussion

The results of this study suggest that the ACG system can be applied to Dutch primary care data, when regarding both face validity and model performance. With regard to the face validity, it can be concluded that the assignment of ACG categories is as expected: the ACG categories which indicate higher multimorbidity and thus higher expected care burden, are found amongst older patients. With respect to model performance, results showed that distinctions between the different levels of GP healthcare utilization can be made with the ACG system. The ACG and ADG categories, as well as the ICPC chapters (the commonly used primary care coding system), are highly associated with GP utilization. However, the ACG system is at patient level and provides a variety of other risk stratification variables, such as multimorbidity measures, risks of hospitalization and high costs, making the use of the ACG as risk stratification tool a good addition to the use of the ICPC coding system.

Comparison of the results of this study to previous research is challenging, as most previous studies investigating the association of the ACG system with continuous utilization outcome measures. Some previous studies were carried out on dichotomous variables however and showed C-statistics and AUC values between 0.73 and 0.82 for the ACG as predictor for hospitalization (5, 6, 14). In addition, the study by Haas et al. presented C-statistics of 0.67 for emergency department visitation and 0.76 for top 10% healthcare costs (5).

Adding to the above mentioned studies, this study suggests that the ACG system is applicable in primary care. Analyzing primary care data in such a manner is of great importance for the understanding of efficiency of healthcare systems that are under increased physical and financial pressure. A study by Sibley et al. showed that administrative data can be used to determine morbidity burden, an important indicator for future care utilization (15). Kristensen and colleagues assessed the use of the ACG system as a morbidity based casemix adjustment system amongst type 2 diabetes patients in order to allocate resources according to degree of co-morbidity (3). They stated that the Danish healthcare system, which is based on fee for service incentives, would profit from a morbidity based casemix adjustment system. The ACG has also proven to be effective for identifying inequities in healthcare utilization by Shadmi et al. (7). Identifying inequities is the first step towards minimizing unwarranted care gaps. With risk stratification tools such as the ACG, case finding for inclusion in population-level interventions can be performed in more health systems worldwide. A study by Soto-Gordoa used risk stratification to select cases for a patient-centered intervention for multimorbid patients with the goal to lower hospitalization. The approach avoided nine percent of hospitalization when cases were selected with the ACG tool (4).

With our study, a first step towards validation of the ACG system, a tool to shift from disease oriented to population based approaches, is revealed for use in the Netherlands. This is opening up a variety of opportunities to reorganize and manage Dutch primary care in an efficient way.

Although the ACG seems an excellent tool to be used in the Netherlands, local adjustment of the software is of eminent importance. A limitation of this study might be the availability of only GP data (without, for example, hospital and mental health care data), forcing us to restrict healthcare utilization outcomes to GP visits, whereas healthcare utilization may be better defined as a total overview of healthcare use. With our research we were not able to explore other types of healthcare utilization, for example defined by total healthcare costs or more costly types of healthcare utilization such as hospitalization and emergency department visitation. Consequently, a full adjustment of the ACG system for use with Dutch data was not possible yet. Further exploration of the ACG system with the use of different data sources will follow.

Moreover, the quality of data needs to be considered. For this study, routine data from GP registries were used. Risk stratification with routinely collected primary care data is an easy and practical way to perform risk stratification on a large scale. Data quality for risk stratification purposes can be improved and strengthened by linkage with different data sources such as hospital and social care registries. The exclusion of social data, such as ethnicity and underlying socio-economic variables, is another limitation of this study. Ethnicity and even more the underlying socio-economic aspect thereof, may have important aspects on patient's health profiles. The addition of social variables and thus more complete patient profiles are of added value in risk stratification approaches. However, we were unable to include these data in our models, as they were not available in the GP data.

Policy implication

Even though the use of the ACG system typically recommends the use of both primary care and hospital care data, this study shows that the ACG is very promising with the use of solely primary care data, especially in a primary care system with mandatory GP listing. With the possibility of applying risk stratification tools to such primary care based healthcare systems, without the need to link data from different sectors, the information security issues can be avoided. Patients' personal information is already available to GP's for optimal caregiving purposes.

With addition of other data sources on individual patient's level, regulations need to be considered to allow the linkage of personal data. As the value of adding hospital data is still to be explored, further research on both content-specific and regulatory aspects is desirable.

Altogether, applications such as the ACG, are very promising for healthcare systems, as their ability to predict future health utilization can be beneficial for person-tailored health intervention strategies, such as screenings for care management interventions, as well as local, regional or even nationwide healthcare management.

Further research

Before applying the ACG system in Dutch primary care, further research is required. This study showed associations between just two components of the ACG system, the ADG and ACG categories, and GP visitation. Risk scores, for example, for future hospitalization and total healthcare costs were outside the scope of this study. To justify the use of the ACG system as risk stratification tool in Dutch primary care, studies validating the ACG risk scores should be conducted. In addition, the ACG models need to be adjusted and improved for use with Dutch primary care data.

Conclusions

This study showed that the ACG is applicable as risk stratification tool in Dutch primary care using routinely registered data from general practitioners' registries. The ACG system yields good results compared to the traditional ICPC classification. Country specific adjustments in the classification and validation of specific risks are necessary.

References

1. Schäfer WL, Boerma WG, Spreeuwenberg P, Schellevis FG, Groenewegen PP. Two decades of change in European general practice service profiles: conditions associated with the developments in 28 countries between 1993 and 2012. Scandinavian journal of primary health care. 2016;34(1):97-110.
2. Burton LC, Skinner EA, Uscher-Pines L, Lieberman R, Leff B, Clark R, et al. Health of Medicare Advantage plan enrollees at 1 year after Hurricane Katrina. The American journal of managed care. 2009;15(1):13-22.
3. Kristensen T, Rose Olsen K, Sortso C, Ejersted C, Thomsen JL, Halling A. Resources allocation and health care needs in diabetes care in Danish GP clinics. Health policy (Amsterdam, Netherlands). 2013;113(1-2):206-15.
4. Soto-Gordoa M, de Manuel E, Fullaondo A, Merino M, Arrospide A, Igartua JI, et al. Impact of stratification on the effectiveness of a comprehensive patient-centered strategy for multimorbid patients. Health services research. 2019;54(2):466-73.
5. Haas LR, Takahashi PY, Shah ND, Stroebel RJ, Bernard ME, Finnie DM, et al. Risk-stratification methods for identifying patients for care coordination. The American journal of managed care. 2013;19(9):725-32.
6. Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. Medical care. 2012;50(2):131-9.
7. Shadmi E, Balicer RD, Kinder K, Abrams C, Weiner JP. Assessing socioeconomic health care utilization inequity in Israel: impact of alternative approaches to morbidity adjustment. BMC public health. 2011;11:609.
8. Johns Hopkins. The Johns Hopkins ACG System: State of the Art Technology and Tradition of Excellence In One Integrated Solution, White Paper. 2012, December.
9. Johns Hopkins University. The ACG® System: Differentiating Features [Available from: https://www.hopkinsacg.org/advantage/#features. Assessed on 29 June 2020.
10. World Health Organization. International Classification of Primary Care, Second edition (ICPC-2) 2003, March [Available from: https://www.who.int/classifications/icd/adaptations/icpc2/en/. Assessed on 22 June 2020.]
11. Nederlands Huisartsen Genootschap. ICPC [Available from: https://www.nhg.org/themas/artikelen/icpc. Assessed on 22 June 2020.]
12. World Health Organization. The Anatomical Therapeutic Chemical Classification System with Defined Daily Doses (ATC/DDD) 2003 [Available from: https://www.who.int/classifications/atcddd/en/. Assessed on 22 June 2020.]
13. Mosier CI. A critical examination of the concepts of face validity. Educ Psychol Meas. 1947;7(2):191-205.
14. Maltenfort MG, Chen Y, Forrest CB. Prediction of 30-day pediatric unplanned hospitalizations using the Johns Hopkins Adjusted Clinical Groups risk adjustment system. PloS one. 2019;14(8):e0221233.
15. Sibley LM, Moineddin R, Agha MM, Glazier RH. Risk adjustment using administrative data-based and survey-derived methods for explaining physician utilization. Medical care. 2010;48(2):175-82.

Supplementary Material

Supplementary table 1: Aggregation of the original ACG categories for logistic regression model 3.

Recoded ACG	Description recoded ACG	Original ACG	Description original ACG
300	Acute Minor (all ages)	100	Acute Minor, Age 1
		200	Acute Minor, Age 2 to 5
		300	Acute Minor, Age > 5
400	Unchanged	400	Acute Major
500	Unchanged	500	Likely to Recur, w/o Allergies
600	Unchanged	600	Likely to Recur, with Allergies
700	Unchanged	700	Asthma
800	Unchanged	800	Chronic Medical, Unstable
900	Chronic Medical/Specialty, Stable	900	Chronic Medical, Stable
		1000	Chronic Specialty, Stable
1100	Unchanged	1100	Eye/Dental
1200	Unchanged	1200	Chronic Specialty, Unstable
1300	Unchanged	1300	Psychosocial, w/o Psych Unstable
1400	Psychosocial, with Psych Unstable, with or without Psych Stable	1400	Psychosocial, with Psych Unstable, w/o Psych Stable
		1500	Psychosocial, with Psych Unstable, w/ Psych Stable
1600	Unchanged	1600	Preventive/Administrative
1711	Unchanged	1711	Pregnancy: 0-1 ADGs, delivered
1712	Unchanged	1712	Pregnancy: 0-1 ADGs, not delivered
1721	Pregnancy: 2+ ADGs, with or without Major ADGs, delivered	1721	Pregnancy: 2-3 ADGs, no Major ADGs, delivered
		1731	Pregnancy: 2-3 ADGs, 1+ Major ADGs, delivered
		1741	Pregnancy: 4-5 ADGs, no Major ADGs, delivered
		1751	Pregnancy: 4-5 ADGs, 1+ Major ADGs, delivered
		1761	Pregnancy: 6+ ADGs, no Major ADGs, delivered
		1771	Pregnancy: 6+ ADGs, 1+ Major ADGs, delivered

Supplementary table 1: Continued

Recoded ACG	Description recoded ACG	Original ACG	Description original ACG
1722	Pregnancy: 2+ ADGs, with or without Major ADGs, not delivered	1722	Pregnancy: 2-3 ADGs, no Major ADGs, not delivered
		1732	Pregnancy: 2-3 ADGs, 1+ Major ADGs, not delivered
		1742	Pregnancy: 4-5 ADGs, no Major ADGs, not delivered
		1752	Pregnancy: 4-5 ADGs, 1+ Major ADGs, not delivered
		1762	Pregnancy: 6+ ADGs, no Major ADGs, not delivered
		1772	Pregnancy: 6+ ADGs, 1+ Major ADGs, not delivered
1800	Unchanged	1800	Acute Minor and Acute Major
1900	Acute Minor and Likely to Recur, Age 1 to 5	1900	Acute Minor and Likely to Recur, Age 1
		2000	Acute Minor and Likely to Recur, Age 2 to 5
2100	Unchanged	2100	Acute Minor and Likely to Recur, Age > 5, w/o Allergy
2200	Unchanged	2200	Acute Minor and Likely to Recur, Age > 5, with Allergy
2300	Unchanged	2300	Acute Minor and Chronic Medical: Stable
2400	Unchanged	2400	Acute Minor and Eye/Dental
2500	Acute Minor and Psychosocial, with/without Psych Stable/Unstable	2500	Acute Minor and Psychosocial, w/o Psych Unstable
		2600	Acute Minor and Psychosocial, with Psych Unstable, w/o Psych Stable
		2700	Acute Minor and Psychosocial, with Psych Unstable and Psych Stable
2800	Acute Minor/Likely to Recur/Eye & Dental	2800	Acute Minor and Likely to Recur
		3400	Acute Minor/Likely to Recur/Eye & Dental
3000	Acute Minor/Acute Major/Likely to Recur, Age 1 to 5	2900	Acute Minor/Acute Major/Likely to Recur, Age 1
		3000	Acute Minor/Acute Major/Likely to Recur, Age 2 to 5

Supplementary table 1: Continued

Recoded ACG	Description recoded ACG	Original ACG	Description original ACG
3100	Unchanged	3100	Acute Minor/Acute Major/Likely to Recur, Age 6 to 11
3200	Unchanged	3200	Acute Minor/Acute Major/Likely to Recur, Age > 11, w/o Allergy
3300	Unchanged	3300	Acute Minor/Acute Major/Likely to Recur, Age > 11, with Allergy
3500	Unchanged	3500	Acute Minor/Likely to Recur/Psychosocial
3600	Unchanged	3600	Acute Minor/Acute Major/Likely Recur/Eye & Dental
3700	Unchanged	3700	Acute Minor/Acute Major/Likely Recur/Psychosocial
3800	Unchanged	3800	2-3 Other ADG Combinations, Age < 18
3900	Unchanged	3900	2-3 Other ADG Combinations, Males Age 18 to 34
4000	Unchanged	4000	2-3 Other ADG Combinations, Females Age 18 to 34
4100	Unchanged	4100	2-3 Other ADG Combinations, Age > 34
4210	Unchanged	4210	4-5 Other ADG Combinations, Age < 18, no Major ADGs
4220	Unchanged	4220	4-5 Other ADG Combinations, Age < 18, 1+ Major ADGs
4310	Unchanged	4310	4-5 Other ADG Combinations, Age 18 to 44, no Major ADGs
4320	4-5 Other ADG Combinations, Age 18 to 44, 1+ Major ADGs	4320	4-5 Other ADG Combinations, Age 18 to 44, 1+ Major ADGs
		4330	4-5 Other ADG Combinations, Age 18 to 44, 2+ Major ADGs
4410	Unchanged	4410	4-5 Other ADG Combinations, Age > 44, no Major ADGs
4420	Unchanged	4420	4-5 Other ADG Combinations, Age > 44, 1+ Major ADGs
4430	Unchanged	4430	4-5 Other ADG Combinations, Age > 44, 2+ Major ADGs

Supplementary table 1: Continued

Recoded ACG	Description recoded ACG	Original ACG	Description original ACG
4510	6-9 Other ADG Combinations, with/without Major ADGs	4510	6-9 Other ADG Combinations, Age < 6, no Major ADGs
		4520	6-9 Other ADG Combinations, Age < 6, 1+ Major ADGs
		4610	6-9 Other ADG Combinations, Age 6 to 17, no Major ADGs
		4620	6-9 Other ADG Combinations, Age 6 to 17, 1+ Major ADGs
		4710	6-9 Other ADG Combinations, Males, Age 18 to 34, no Major ADGs
		4720	6-9 Other ADG Combinations, Males, Age 18 to 34, 1+ Major ADGs
		4730	6-9 Other ADG Combinations, Males, Age 18 to 34, 2+ Major ADGs
		4810	6-9 Other ADG Combinations, Females, Age 18 to 34, no Major ADGs
		4820	6-9 Other ADG Combinations, Females, Age 18 to 34, 1+ Major ADGs
		4830	6-9 Other ADG Combinations, Females, Age 18 to 34, 2+ Major ADGs
		4910	6-9 Other ADG Combinations, Age > 34, 0-1 Major ADGs
		4920	6-9 Other ADG Combinations, Age > 34, 2 Major ADGs
		4930	6-9 Other ADG Combinations, Age > 34, 3 Major ADGs
		4940	6-9 Other ADG Combinations, Age > 34, 4+ Major ADGs
5030	10+ Other ADG Combinations, with/without Major ADGs	5030	10+ Other ADG Combinations, Age 1 to 17, 2 Major ADGs
		5040	10+ Other ADG Combinations, Age > 17, 0-1 Major ADGs
		5050	10+ Other ADG Combinations, Age > 17, 2 Major ADGs
		5060	10+ Other ADG Combinations, Age > 17, 3 Major ADGs
		5070	10+ Other ADG Combinations, Age > 17, 4+ Major ADGs

For the purpose of this study, some ACG categories were collapsed together. The recoded ACG categories are presented within the table. When the description field of the recoded ACG reads ‘unchanged’ the ACG category remained the original ACG category.

Supplementary table 2: Distribution of persons over ACG categories taking age into account

	Age categories (10yr bands)									**Total**
	0-9	**10-19**	**20-29**	**30-39**	**40-49**	**50-59**	**60-69**	**70-79**	**80+**	
One Aggregated Diagnostic Group	***724***	***1294***	***1314***	***1521***	***1243***	***697***	***399***	***117***	***18***	***7327***
Acute Minor (all ages)	361	583	345	293	378	240	104	24	5	2333
Acute Major	126	184	137	109	170	94	53	20	2	895
Likely to Recur	134	267	167	171	222	112	56	13	1	1143
Asthma	8	19	17	9	12	5	7	0	0	77
Chronic Medical or Specialty, Unstable	3	9	3	14	41	31	35	18	4	158
Chronic Medical/Specialty, Stable	7	23	22	26	54	108	105	33	2	380
Eye/Dental	8	10	5	1	0	3	5	3	1	36
Psychosocial, w/o Psych Unstable	55	90	42	38	48	25	11	3	0	312
Psychosocial, with Psych Unstable, with or without Psych Stable	0	3	11	6	4	5	0	0	0	29
Preventive/Administrative	22	94	79	70	90	50	23	3	3	434
Pregnancy	***0***	***6***	***243***	***392***	***112***	***12***	***0***	***0***	***0***	***765***
Pregnancy: delivered or not delivered	0	6	243	392	112	12	0	0	0	765
Two-Three Aggregated Diagnostic Groups	***1051***	***1490***	***1040***	***1049***	***1616***	***1326***	***1065***	***404***	***106***	***9137***
Acute Minor and Acute Major	193	270	176	166	194	166	80	18	10	1273
Acute Minor and Likely to Recur	207	332	221	191	264	162	80	16	2	1465
Acute Minor and Chronic Medical: Stable	12	23	44	30	70	99	115	42	10	445
Acute Minor and Eye/Dental with or w/o Eye & Dental	98	121	93	83	124	63	31	15	1	629
Acute Minor and Psychosocial, with/without Psych Stable/Unstable	77	98	52	36	51	36	14	0	1	365
Acute Minor/Acute Major/Likely to Recur, with or w/o Allergy	199	228	168	188	266	157	102	26	8	1342
Acute Minor/Likely to Recur/Psychosocial	49	61	37	43	58	47	13	3	2	313
2-3 Other ADG Combinations	216	357	249	312	589	596	630	284	72	3305

Supplementary table 2: Continued

	Age categories (10yr bands)									**Total**
	0-9	**10-19**	**20-29**	**30-39**	**40-49**	**50-59**	**60-69**	**70-79**	**80+**	
Four-Five Aggregated Diagnostic Groups	***200***	***304***	***274***	***384***	***727***	***894***	***967***	***577***	***213***	***4540***
Acute Minor/Acute Major/Likely Recur/Eye & Dental	11	45	48	57	136	193	209	113	32	844
Acute Minor/Acute Major/Likely Recur/Psychosocial	48	52	48	84	101	55	32	5	4	429
4-5 Other ADG Combinations, no Major ADGs	107	167	117	160	274	313	282	134	50	1604
4-5 Other ADG Combinations, 1+ Major ADGs	34	40	61	83	216	333	444	325	127	1663
Six-Nine Aggregated Diagnostic Groups	***54***	***93***	***102***	***139***	***369***	***533***	***647***	***668***	***382***	***2987***
6-9 Other ADG Combinations, with/without Major ADGs	54	93	102	139	369	533	647	668	382	2987
Over 10 Aggregated Diagnostic Groups	***0***	***1***	***7***	***15***	***32***	***48***	***76***	***111***	***102***	***392***
10+ Other ADG Combinations, with/without Major ADGs	0	1	7	15	32	48	76	111	102	392
Total	**2029**	**3176**	**2484**	**2716**	**3875**	**3486**	**3154**	**1877**	**821**	**23618**

Distribution of individuals within the ACG categories. Some ACG categories are regrouped, due to small numbers in said categories. ACG categories are grouped according to number of ADG's and disaggregated in 10-year age categories.

Supplementary table 3: Odds Ratio with 95% confidence intervals for dependent variables in model 1 with the second definition for GP utilization.

Variables	***Odds Ratio***	***95% Confidence intervals***	
		Lower	**Upper**
Gender	0,925	0,861	0,994
Age	1,003	1,001	1,005
ICPC chapter A *General and unspecified*	1,42	1,33	1,516
ICPC chapter B *Blood, blood forming organs, lymphatics, spleen*	2,017	1,658	2,454
ICPC chapter D *Digestive'*	1,836	1,706	1,975
ICPC chapter F *Eye'*	1,324	1,206	1,454
ICPC chapter H *Ear'*	1,453	1,338	1,578
ICPC chapter K *Circulatory'*	2,701	2,48	2,942
ICPC chapter L *Musculoskeletal'*	1,87	1,758	1,988
ICPC chapter N *Neurological'*	1,713	1,55	1,893
ICPC chapter P *Psychological'*	1,985	1,843	2,139
ICPC chapter R *Respiratory'*	1,736	1,63	1,85
ICPC chapter S *Skin'*	1,974	1,855	2,101
ICPC chapter T *Endocrine metabolic and nutritional*	2,466	2,245	2,709
ICPC chapter U *Urology'*	1,759	1,597	1,938
ICPC chapter W *Pregnancy, childbirth, family planning*	1,886	1,701	2,091
ICPC chapter X *Female genital system and breast*	1,509	1,374	1,657
ICPC chapter Y *Male genital system*	1,619	1,406	1,864
ICPC chapter Z *Social problems*	2,624	2,241	3,073
Constant	0,123		

Odds Ratio's and Confidence Intervals of the variables in model 1 (including ICPC chapters) are presented for outcome 2 (zero or one GP visit versus two or more GP visits).

Supplementary table 4: Odds Ratio with 95% confidence intervals for dependent variables in model 2 with the second definition for GP utilization.

Variables	***Description***	***Odds Ratio***	***95% Confidence intervals***	
			Lower	**Upper**
Gender		0,85	0,797	0,906
Age		1,008	1,006	1,01
ADG01	Time Limited: Minor	1,489	1,382	1,605
ADG02	Time Limited: Minor-Primary Infections	1,788	1,655	1,932
ADG03	Time Limited: Major	1,273	0,918	1,764
ADG04	Time Limited: Major-Primary Infections	1,817	1,468	2,249
ADG05	Allergies	1,354	1,223	1,5
ADG06	Asthma	1,677	1,479	1,902
ADG07	Likely to Recur: Discrete	1,653	1,518	1,799
ADG08	Likely to Recur: Discrete-Infections	1,566	1,426	1,72
ADG09	Likely to Recur: Progressive	1,487	1,133	1,951
ADG10	Chronic Medical: Stable	2,881	2,66	3,12
ADG11	Chronic Medical: Unstable	1,921	1,723	2,143
ADG12	Chronic Specialty: Stable-Orthopedic	1,704	1,279	2,27
ADG13	Chronic Specialty: Stable-Ear,Nose,Throat	1,335	0,983	1,813
ADG14	Chronic Specialty: Stable-Eye	1,303	1,079	1,573
ADG16	Chronic Specialty: Unstable-Orthopedic	1,636	1,333	2,008
ADG17	Chronic Specialty: Unstable-Ear,Nose,Throat	0,708	0,417	1,203
ADG18	Chronic Specialty: Unstable-Eye	1,227	0,887	1,699
ADG20	Dermatologic	1,968	1,819	2,129
ADG21	Injuries/Adverse Effects: Minor	1,828	1,646	2,031
ADG22	Injuries/Adverse Effects: Major	1,416	1,235	1,625
ADG23	Psychosocial: Time Limited, Minor	2,335	2,044	2,668
ADG24	Psychosocial:Recurrent or Persistent,Stable	1,945	1,775	2,132
ADG25	Psychosocial:Recurrent or Persistent,Unstable	1,564	1,183	2,066
ADG26	Signs/Symptoms: Minor	1,877	1,766	1,996
ADG27	Signs/Symptoms: Uncertain	2,086	1,96	2,22
ADG28	Signs/Symptoms: Major	1,588	1,382	1,825
ADG29	Discretionary	1,523	1,358	1,709
ADG30	See and Reassure	1,99	1,72	2,303
ADG31	Prevention/Administrative	1,501	1,397	1,614
ADG32	Malignancy	1,515	1,276	1,798
ADG33	Pregnancy	1,605	1,351	1,907
ADG34	Dental	1,312	0,891	1,932
Constant		0,108		

Odds Ratio's and Confidence Intervals of the variables in model 2 (including ADGs) are presented for outcome 2 (zero or one GP visit versus two or more GP visits).

Supplementary table 5: Odds Ratio with 95% confidence intervals for dependent variables in model 3 with the second definition for GP utilization.

Variables	***Description***	***Odds Ratio***	***95% Confidence intervals***	
			Lower	**Upper**
Gender		0,779	0,732	0,83
Age		1,009	1,007	1,011
ACG 300	Acute Minor (all ages)			
ACG 400	Acute Major	0,933	0,761	1,143
ACG 500	Likely to Recur, w/o Allergies	0,851	0,691	1,048
ACG 600	Likely to Recur, with Allergies	0,559	0,372	0,84
ACG 700	Asthma	1,153	0,649	2,05
ACG 800	Chronic Medical, Unstable	1,287	0,835	1,985
ACG 900	Chronic Medical/Specialty, Stable	2,206	1,738	2,798
ACG 1100	Eye/Dental	0,124	0,017	0,911
ACG 1200	Chronic Specialty, Unstable	0,395	0,14	1,112
ACG 1300	Psychosocial, w/o Psych Unstable	1,382	1,038	1,84
ACG 1400	Psychosocial, with Psych Unstable, with or without Psych Stable	1,207	0,487	2,987
ACG 1600	Preventive/Administrative	0,141	0,082	0,243
ACG 1711	Pregnancy: 0-1 ADGs, delivered	1,065	0,483	2,349
ACG 1712	Pregnancy: 0-1 ADGs, not delivered	0,323	0,099	1,057
ACG 1721	Pregnancy: 2+ ADGs, with or without Major ADGs, delivered	9,108	7,067	11,739
ACG 1722	Pregnancy: 2+ ADGs, with or without Major ADGs, not delivered	5,159	4,044	6,581
ACG 1800	Acute Minor and Acute Major	3,214	2,755	3,75
ACG 1900	Acute Minor and Likely to Recur, Age 1 to 5	4,958	3,144	7,816
ACG 2100	Acute Minor and Likely to Recur, Age > 5, w/o Allergy	3,244	2,759	3,813
ACG 2200	Acute Minor and Likely to Recur, Age > 5, with Allergy	3,249	2,535	4,165
ACG 2300	Acute Minor and Chronic Medical: Stable	4,024	3,235	5,006
ACG 2400	Acute Minor and Eye/Dental	1,835	0,907	3,714
ACG 2500	Acute Minor and Psychosocial, with/without Psych Stable/Unstable	3,144	2,482	3,981
ACG 2800	Acute Minor/Likely to Recur/Eye & Dental	2,806	2,303	3,418
ACG 3000	Acute Minor/Acute Major/Likely to Recur, Age 1 to 5	5,957	3,913	9,066
ACG 3100	Acute Minor/Acute Major/Likely to Recur, Age 6 to 11	6,406	4,48	9,16
ACG 3200	Acute Minor/Acute Major/Likely to Recur, Age > 11, w/o Allergy	7,442	6,238	8,879
ACG 3300	Acute Minor/Acute Major/Likely to Recur, Age > 11, with Allergy	6,738	5,083	8,931

3

Supplementary table 5: Continued

Variables	***Description***	***Odds Ratio***	***95% Confidence intervals***	
			Lower	**Upper**
ACG 3500	Acute Minor/Likely to Recur/Psychosocial	7,892	6,112	10,19
ACG 3600	Acute Minor/Acute Major/Likely Recur/Eye & Dental	14,665	11,933	18,024
ACG 3700	Acute Minor/Acute Major/Likely Recur/Psychosocial	12,683	9,929	16,202
ACG 3800	2-3 Other ADG Combinations, Age < 18	3,131	2,529	3,875
ACG 3900	2-3 Other ADG Combinations, Males Age 18 to 34	3,395	2,591	4,448
ACG 4000	2-3 Other ADG Combinations, Females Age 18 to 34	5,346	3,944	7,247
ACG 4100	2-3 Other ADG Combinations, Age > 34	4,34	3,762	5,007
ACG 4210	4-5 Other ADG Combinations, Age < 18, no Major ADGs	6,971	5,289	9,187
ACG 4220	4-5 Other ADG Combinations, Age < 18, 1+ Major ADGs	10,412	6,073	17,852
ACG 4310	4-5 Other ADG Combinations, Age 18 to 44, no Major ADGs	10,637	8,422	13,435
ACG 4320	4-5 Other ADG Combinations, Age 18 to 44, 1+ Major ADGs	7,881	5,954	10,432
ACG 4410	4-5 Other ADG Combinations, Age > 44, no Major ADGs	11,398	9,367	13,869
ACG 4420	4-5 Other ADG Combinations, Age > 44, 1+ Major ADGs	9,907	8,237	11,915
ACG 4430	4-5 Other ADG Combinations, Age > 44, 2+ Major ADGs	11,045	8,035	15,181
ACG 4510	6-9 Other ADG Combinations, with/without Major ADGs	29,765	25,175	35,192
ACG 5030	10+ Other ADG Combinations, with/without Major ADGs	76,667	45,043	130,493
Constant		0,195		

Odds Ratio's and Confidence Intervals of the variables in model 3 (including ACGs) are presented for outcome 2 (zero or one GP visit versus two or more GP visits).

CHAPTER 4

CHAPTER 4

VALIDATING AND IMPROVING ACG'S FUTURE HOSPITALIZATION AND HIGH-COST PREDICTION MODELS FOR DUTCH PRIMARY CARE

Shelley-Ann M Girwar, Marta Fiocco, Stephen P Sutch, Mattijs E Numans, Marc A Bruijnzeels

Submitted

Abstract

Introduction

The rise in healthcare costs, caused by older and more complex patient populations, requires Population Health Management approaches including risk stratification. With risk stratification, patients are assigned individual risk scores based on medical records. These patient stratifications focus on future high costs and expensive care utilization such as hospitalization, for which different models exist. With this study, we validated the accuracy of risk prediction scores for future hospitalization and high healthcare costs, calculated by the ACG's risk stratification models, using Dutch primary care health data registries. In addition, we aimed to adjust the US-based predictive models for Dutch primary care.

Methods

The statistical validity of the existing models was assessed. In addition, the underlying prediction models were trained on 95,262 patients' data from de Zoetermeer region and externally validated on data of 48,780 patients from Zeist, Nijkerk and Urk. Information on age, sex, number of GP visits, International Classification of Primary Care coded information on the diagnosis and Anatomical Therapeutic Chemical Classification coded information on the prescribed medications, were incorporated in the model. C-statistics were used to validate the discriminatory ability of the models. Calibrating ability was assessed by visual inspection of calibration plots.

Results

Adjustment of the hospitalization model based on Dutch data improved C-statistics from 0.69 to 0.75, whereas adjustment of the high-cost model improved C-statistics from 0.78 to 0.85, indicating good discrimination of the models. The models also showed good calibration.

Conclusion

In conclusion, the local adjustment of the ACG prediction models, show great potential for use in Dutch primary care, in terms of prediction of future hospitalization and high costs.

Introduction

Multimorbidity is increasingly being recognized as the norm rather than the exception, since patient populations are becoming older and more complex, and patient information is becoming more complete. The increased complexity leads to increased healthcare utilization. In particular, there is a rise in expensive care such as hospital and emergency care, which has a major effect on healthcare costs. To manage these rising health costs, approaches that focus on complete patients' health profiles are needed. Population Health Management is focusing on the coordination of care delivery across specified sub-populations to improve the population's health and care utilization (1, 2). In addition, analysis of routine health registry data is increasingly being used to provide a basis for proactive care interventions, attempting to lower healthcare costs by reducing expensive and avoidable care such as hospital admissions. Risk stratification, the assignment of individual risk scores to patients based on registered health profiles, has proven to be an effective tool in the provision of proactive care. A study by Freund *et al* successfully selected high risk patients for care management programs, using risk stratification (3). Another study has shown that efficient care management approaches using risk stratification have led to reduced hospitalization rates (4).

A wide variety of risk stratification tools exist, with risk predictions for various health and health utilization outcomes, such as the risk for future hospitalization, high healthcare costs, emergency care utilization and even mortality. One of the most frequently used risk stratification tools in primary care is the Johns Hopkins Adjusted Clinical Groups (ACG) system (5), with proven efficacy for prediction of not only health outcomes such as morbidity, but also of different types of future healthcare utilization, such as hospitalization and emergency department visits, and future health costs (6).

Most risk stratification tools predominantly use hospital data, with and without primary care data. However, patient's privacy protection in Europe is complicating the linkage of different health data sources. Performing risk stratification based on primary care routine registry data, extracted from only one source in which essential information from most other relevant sources is present, is a way to overcome the privacy challenge. Evidence of the efficacy and accuracy of risk stratification approaches based on primary care data, is still insufficient in the Netherlands. With this study, we aimed to validate the accuracy of the ACG's risk prediction scores for future hospitalization and high healthcare costs, using Dutch primary care health data registries as input data, and to adjust the US based predictive models for Dutch primary care.

Methods

Design

This study had three aims: 1) assessment of two existing prediction models within the ACG tool, which are based on US data, 2) adjustment of the prediction algorithms towards Dutch primary care data and 3) assessment of the adjusted prediction models.

Assessment of ACG's fixed prediction models (based on US data)

The ACG system, developed by the Johns Hopkins University, includes many different risk prediction models. With this study, two of those existing prediction models were assessed: 1) the 'hospitalization model', estimating probabilities for becoming hospitalized at least once in the following 12 months and 2) the 'high cost model', estimating probabilities for being in the top 5% of the population with the highest healthcare costs in the following 12 months.

The ACG models are existing models, based on years of research with US data. We applied these two ACG models to retrospective Dutch primary care data, available in general practitioners' (GPs) electronic medical records. Subsequently, we assessed model performances, using observed outcomes extracted from historic medical specialty data.

Adjustment of the models, based on Dutch primary care data

In addition to the application and assessment of the fixed prediction models of the ACG tool, we aimed to adjust the two prediction models to the Dutch situation. Therefore, we produced logistic regression models for hospitalization and high-cost with the same predictors used by the ACG, using retrospective data from a Dutch primary care population, and adjusted the coefficients of those predictors.

Assessment of the adjusted models

To assess the performance of both the hospitalization and the high-cost model, we investigated the discriminating and calibrating ability. The discriminating ability relates to how well a prediction model can distinguish those with the outcome from those without, while the calibrating ability relates to the agreement between observed and predicted values (7). The assessment was performed by externally validating the models with retrospective data from a second Dutch primary care population.

Data and study population

For this study, we used data from GP enlisted patient populations in the Netherlands. We used extractions of the GPs' electronic medical records as input data for the applied prediction models, and secondary care (hospital) data for the observed outcomes of the models.

Assessment of ACG's existing prediction models

To assess the ACG's existing prediction models, the models were applied to historic primary care data from 95,262 primary care patients within the Zoetermeer region in the Netherlands. Data from January to December 2014, were extracted from participating GPs' electronic health records and were used as input data for the prediction model. Information on age, sex, number of GP visits, International Classification of Primary Care version 1 (ICPC-1) coded information on the diagnosis and Anatomical Therapeutic Chemical Classification (ATC) coded information on the prescribed medications, were incorporated in the model. We translated ICPC-1 codes, used in Dutch primary care, to the international ICPC-2 codes, required as input for the ACG System. As ICPC-1 codes are sometimes more specific than ICPC-2 codes, we have translated some ICPC-1 codes to International Classification of Diseases 10th revision (ICD-10) codes, a coding system that can also be recognized by the ACG System, rather than to ICPC-2 codes. Translation was based on ICPC-1 and ICPC-2 differences described by Wonca International Classification Committee (8) with additional expert opinions. (supplementary table 1)

The outcome variables for the prediction models were extracted from medical specialty care records, available as microdata from Statistics Netherlands, the Dutch Central Bureau for Statistics. Outcomes extracted from Statistics Netherlands' microdata included information on hospitalization and reimbursed healthcare costs from January to December 2015. In the Netherlands, healthcare costs are reimbursed by health insurers based on mandatory basic health insurance law and only the costs covered by the basic health insurance are included as healthcare costs for this study.

As GP data from 2014 were used, patients were included when registered with one of the participating GP practices for the complete year of 2014, but only when linkage with the Statistics Netherlands database was possible (91.7% of the patients).

Data from the GP's electronic health registries was linked to medical specialty data by encryption of both datasets. To each individual a unique Record Identification Number (RIN) was assigned, based on birth date, gender, and complete postal code. The RINs were the used to link the GP data to the Statistics Netherlands' microdata.

Adjustment of the models, based on Dutch primary care data

For the adjustment of the two US based prediction models, the same data and study population were used as described in previous paragraphs.

Assessment of the adjusted models

To assess the adjusted prediction models, a second study population was used. The study population of 48,780 patients from Zeist, Nijkerk and Urk was used to externally validate the prediction models. Similar retrospective primary and secondary care data were used as described previously.

Statistical Analysis

Firstly, the similarity between the two study populations was assessed. Continuous variables have been tested with t-tests. In case of violation of the normality assumption, a non-parametric test was used. For the categorical variables, chi-squared test was used.

Assessment of ACG's existing prediction models

The ACG System US based hospitalization and high costs models were assessed on model performance. Predicted values, generated by the ACG were compared to the observed outcomes (described in the next section 'Adjustment of the models') by calculation of C-statistics. C-statistics below 0.6 were taken to indicate poor model performance, C-statistics between 0.6 and 0.7 to indicate sufficient model performance and C-statistics above 0.7 indicate good model performance (9).

Adjustment of the models

To adjust the two prediction models to the Dutch primary care data, we used the underlying logistic regressions for hospitalization and high healthcare costs. We estimated the logistic regressions using the first primary care population to find new coefficients for the predictors, resulting in adjusted prediction models.

Dependent variables

The dependent variable for the first model was hospitalization in the second year. Hospitalization in the second year was defined as being at least on hospital admission in the period between January and December 2015 based on Statistic Netherlands microdata.

The dependent variable for the second prediction model was high healthcare costs in the next year. High healthcare costs in the next year was defined as being in the top 5% of highest healthcare costs within the population in the period between January and December 2015, again based on Statistic Netherlands microdata.

Independent variables

Independent variables concern all variables of the ACG's hospitalization and high costs models, which were available in our Dutch data. Next to patients' characteristics such as age, sex and GP care utilization, independent variables included specific diagnoses, types of diagnosis, burden of care categories and mutually exclusive multimorbidity categories, which are based on complete diagnosis and medication profiles of individual patients.

Assessment of the adjusted models

C-statistics were calculated to assess model performance regarding discrimination. First, C-statistics were calculated for the prediction model estimated in the first population, resulting in coefficients adjustments of the fixed AGC models. C-statistics for those adjusted prediction models were compared to those of the existing US-based ACG models. The adjusted models show improvement when C-statistics are higher than those of the US-based ACG models.

Second, the adjusted prediction models were externally validated in a second study population. Both discrimination and calibration were estimated. Calibration was assessed by dividing the validation dataset population into deciles based on ascending predicted values for the different outcomes. For each group the mean observed and expected values were plotted in a calibration plot. Models with a 45-degree angle plot (mean observed value equals mean expected value) are considered perfectly calibrated. Models below this reference line are overestimating, whereas models above it are underestimating.

Privacy

Primary care patients were informed about the use of their data for research purposes. Patients were given the opportunity to opt out.

Patient data were encrypted by Statistics Netherlands under strict rules to secure individuals´ privacy. Linkage and analyses of the data was performed within the secured environment of Statistics Netherlands.

Results

Population Characteristics

In table 1 the differences for various characteristics between the two study populations are shown. The two populations are comparable with respect to the percentage of females within the population, the mean number of GP visits in 2014 and the percentage of people hospitalized in 2015. The mean age shows a difference of one year between the two populations.

Table 1: Population characteristics; differences between the populations (Zoetermeer versus Nijkerk + Urk + Zeist).

	Total population	**Zoetermeer (n= 95, 262)**	**Nijkerk, Urk, Zeist (n= 48, 780)**
Mean age in years (SD)	39.5 (22.2)	39.9 (22.0)	38.8 (22.8)
Sex (% females)	50.9	51.4	50.0
Number of GP visits < 2 (n; %) >= 2 (n; %)	 66, 643; 46,3% 77, 399; 53,7%	 41, 480; 43.5% 53, 782; 56.5%	 16, 834; 34.5% 31, 946; 65.5%
Hospitalized in 2015 (%)	9.9	10.1	9.4
Median costs in 2015	€ 479,99	€ 531,18	€ 393,35

Statistical Assessment

Assessment of ACG's existing prediction models (based on US data)

To assess the performance of the existing ACG models, which are based on US data, we calculated C-statistics. The C-statistic for the ACG hospitalization model was

0.69, suggesting a modest performance of the model. With a C-statistic of 0.78, the discriminating ability of the high-cost model can be classified as good.

Adjustment of the models, based on Dutch primary care data

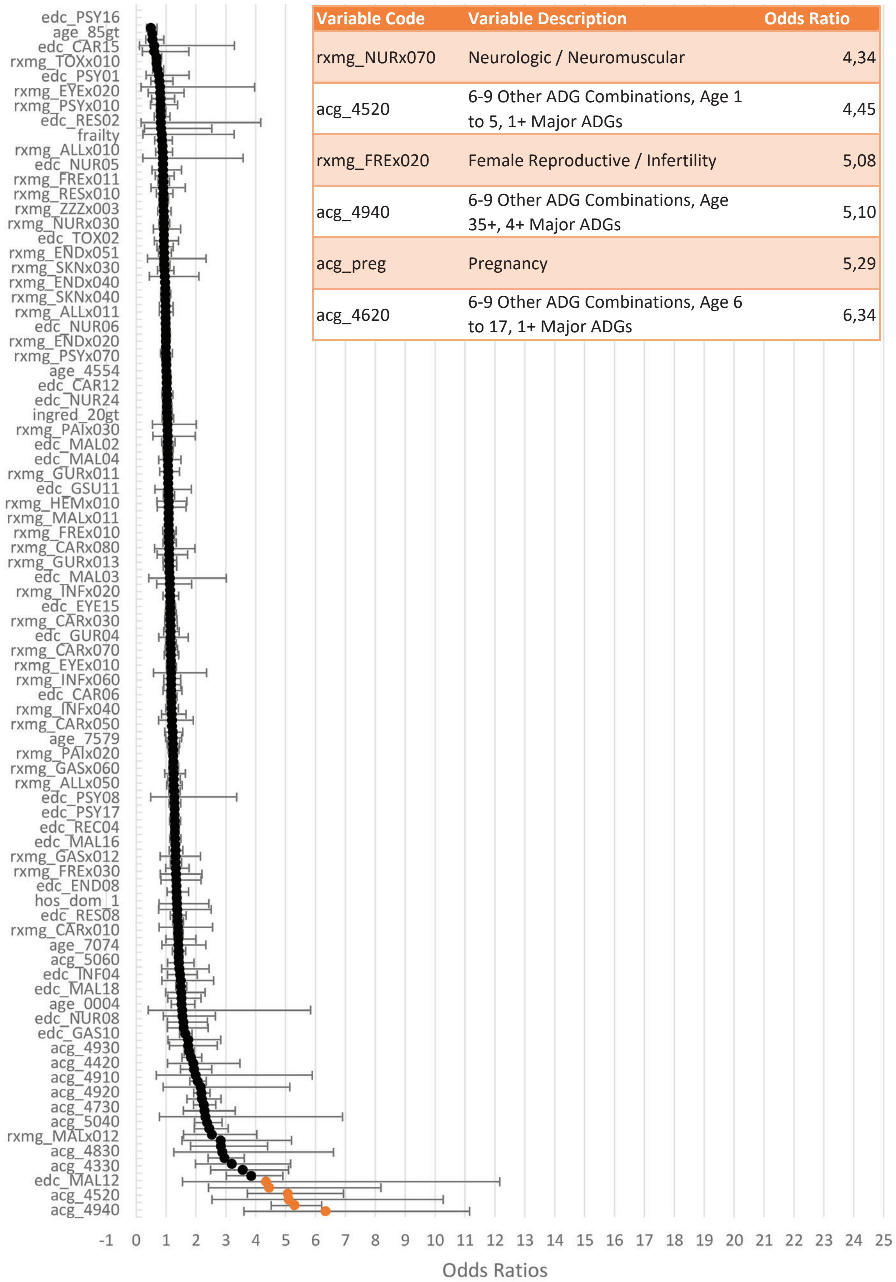
To adjust the models to the Dutch primary care setting, we first estimated and then validated the logistic regression models, producing new prediction models. Figures 1 and 2 show the odds ratios along with the 95% confidence interval for the variables included in respectively the hospitalization and high cost prediction models, arranged from lowest (top) to highest (bottom). For the hospitalization model the ACG categories for children under 18 years old with six to nine diagnoses types, amongst which at least one was assigned as a major diagnosis and the ACG category for adults above 34, years old, with six to nine diagnoses types, amongst which at least four were assigned as major diagnoses, along with neurological/neuromuscular problems, female infertility and pregnancy are the variables with the highest odds ratios (>4).

With odds ratios above five, ACG categories for children under 18 years old with four to nine diagnoses, with or without major diagnoses, along with conditions such as female infertility, acute major viral infections, malignancies with high impact and Multiple Sclerosis, all contribute highly to the high-cost model. In addition, pregnancy, systemic inflammation with high impact, muscle spasms, chromosomal anomalies, chronic kidney disease and the ACG category for males between 18 and 34 years old with six to nine diagnoses, amongst which two or more major ones, also contribute highly to the model with odds ratios above four.

Assessment of the adjusted models

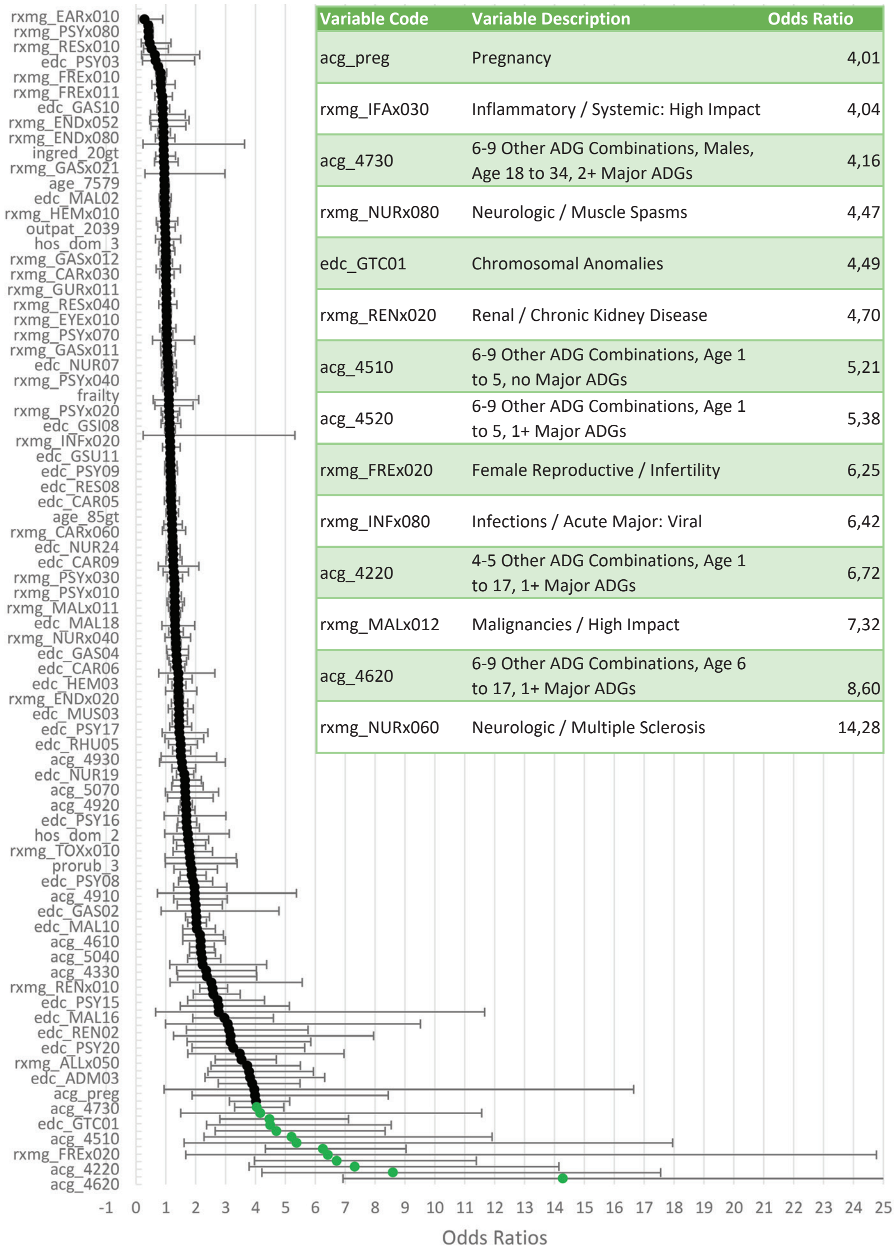
Discriminatory ability

To assess the discriminatory ability of the adjusted models, we compared the C-statistics estimated for the model based on the Dutch data, to those for the US-based models. Table 2 shows fairly high C-statistics for the US hospitalization and high-cost models (0.69 (CI 0.68, 0.70) and 0.78 (CI 0.77, 0.79)). In addition, the Dutch models both show improvements of discriminatory ability with C-statistics raising to 0.75 (CI 0.74, 0.75) for the hospitalization model and 0.85 (CI 0.84, 0.85) for the high-cost one. C-statistics for both Dutch adjusted models were similar for training and validating datasets, suggesting a similar discriminating performance of the adjusted models in an external study population.



Variable Code	Variable Description	Odds Ratio
rxmg_NURx070	Neurologic / Neuromuscular	4,34
acg_4520	6-9 Other ADG Combinations, Age 1 to 5, 1+ Major ADGs	4,45
rxmg_FREx020	Female Reproductive / Infertility	5,08
acg_4940	6-9 Other ADG Combinations, Age 35+, 4+ Major ADGs	5,10
acg_preg	Pregnancy	5,29
acg_4620	6-9 Other ADG Combinations, Age 6 to 17, 1+ Major ADGs	6,34

Figure 1: On the left: S-curve of the odds ratios including confidence intervals for the variables included in the hospitalization model, arranged from highest (bottom) to lowest (top); on the right: table zoomed in on the odds ratio's above four.



Variable Code	Variable Description	Odds Ratio
acg_preg	Pregnancy	4,01
rxmg_IFAx030	Inflammatory / Systemic: High Impact	4,04
acg_4730	6-9 Other ADG Combinations, Males, Age 18 to 34, 2+ Major ADGs	4,16
rxmg_NURx080	Neurologic / Muscle Spasms	4,47
edc_GTC01	Chromosomal Anomalies	4,49
rxmg_RENx020	Renal / Chronic Kidney Disease	4,70
acg_4510	6-9 Other ADG Combinations, Age 1 to 5, no Major ADGs	5,21
acg_4520	6-9 Other ADG Combinations, Age 1 to 5, 1+ Major ADGs	5,38
rxmg_FREx020	Female Reproductive / Infertility	6,25
rxmg_INFx080	Infections / Acute Major: Viral	6,42
acg_4220	4-5 Other ADG Combinations, Age 1 to 17, 1+ Major ADGs	6,72
rxmg_MALx012	Malignancies / High Impact	7,32
acg_4620	6-9 Other ADG Combinations, Age 6 to 17, 1+ Major ADGs	8,60
rxmg_NURx060	Neurologic / Multiple Sclerosis	14,28

Figure 2: On the left: S-curve of the odds ratios including confidence intervals for the variables included in the high cost model, arranged from highest (bottom) to lowest (top); on the right: table zoomed in on the odds ratio's above four.

Table 2: C-statistics for hospitalization and high healthcare costs (top 5% highest healthcare costs) models: fixed US model versus adjusted model based on Dutch data.

	Hospitalization Model *C-statistics (95% CI interval)*	**High Cost Model** *C-statistics (95% CI interval)*
US-based model (training dataset)	0.689 (0.683, 0.695)	0.779 (0.772, 0.786)
US-based model (validation dataset)	0.704 (0.695, 0.712)	0.793 (0.784, 0.803)
Dutch Model (training dataset)	0.748 (0.743, 0.753)	0.844 (0.838, 0.850)
Dutch Model (validation dataset)	0.756 (0.748, 0.763)	0.857 (0.849, 0.865)

Calibrating ability

The calibration plots of both the adjusted hospitalization model (figure 3) and the high-cost model (figure 4) are located near the 45 degree reference line, indicating that the calibrating ability of both models is good: the persons with higher predicted values indeed have a higher chance of being hospitalized or generating higher healthcare costs.

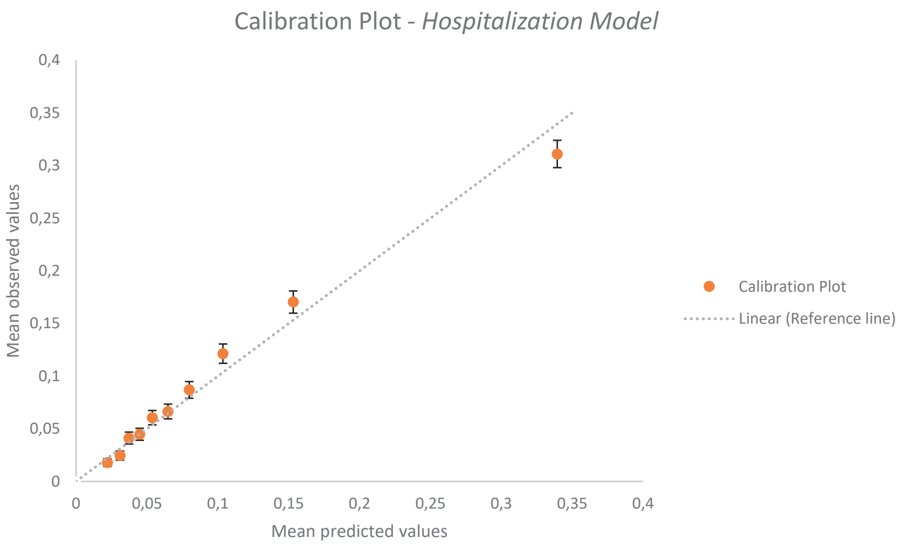


Figure 3: Calibration plot hospitalization model (external validation)

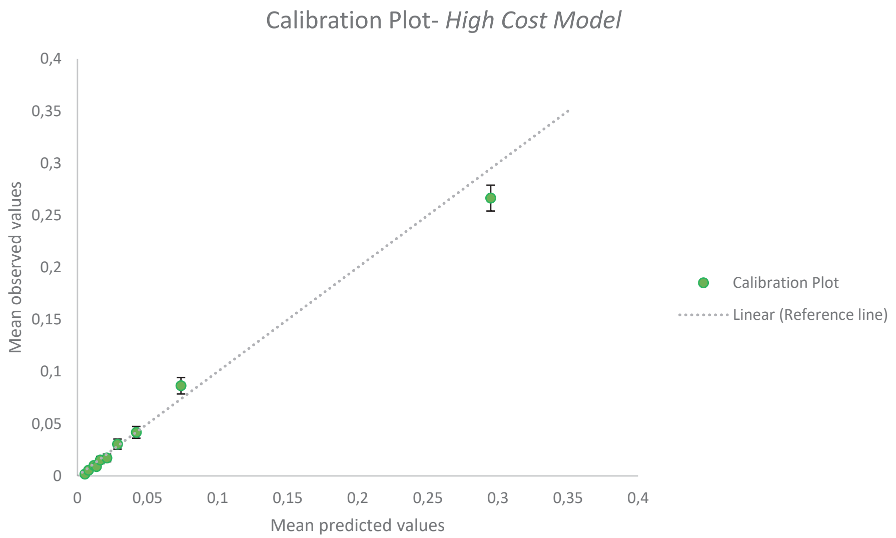


Figure 4: Calibration plot high healthcare costs / top 5% highest healthcare costs model (external validation)

Discussion

With this study we have identified promising risk stratification tools to be used in Dutch primary care. With the ACG tool applied on Dutch primary care data, model performances for the US based models are 0.69 for the hospitalization model and 0.78 for the high-cost model. The ACG has already been proven to be an efficient risk stratification tool in different countries with C-statistics between 0.73 and 0.82 for hospitalization risk and C-statistics of 0.76 for prediction of high healthcare costs (10, 11, 12). This study suggests that the ACG's can also be used properly in the Netherlands, especially after adjustment of the model towards Dutch data. Adjustment of the hospitalization model based on Dutch data improved C-statistics to 0.75, upgrading the model's performance. The high-cost model produced C-statistics of 0.85 after adjustment, which is regarded as 'very good'. Next to good discriminatory ability, the models also showed good calibrating ability: the models can discriminate well between low- and high-risk individuals and the predicted values are in line with the observed ones. The models show excellent potential for predicting high risk individuals within a Dutch primary care population.

Good prediction models to identify future risk of hospitalization or high costs, can be of great value for planning and organizing effective healthcare provision. Applying such models in primary care, enables identification of high-risk patients at an early stage, potentially resulting in pro-active care and proper allocation of resources. As resources are getting scarce in most European countries, including the Netherlands, approaches focusing on effective and efficient resource allocation are highly valuable.

Different studies have already shown the success of selecting appropriate patients for specific interventions such as care management programs with the use of efficient risk stratification tools (3). Subsequently, the effect of tailor-made approaches based on patients' individual risks has proven its value in reducing hospitalization and high healthcare costs (4). Population Health Management approaches like those have the capacity to keep healthcare costs under control.

This study has shown the high potential of the ACG's adjusted risk models. However, this study only focused on the ACG's hospitalization and high costs model. The many other risk models that are included in the ACG and other similar tools, all need to be validated in the Netherlands before being used in practice. However, with the validation of the hospitalization and high costs models, we expect that the other ACG models will also perform well.

Secondly, to strengthen the models even more, the clinical validity of the predictors in the models, needs to be reassessed for a Dutch setting. A strong statistical association with a predictor and the outcome does not necessarily establish the clinical meaning of the predictor. Focus should be put on the association of the model predictors with avoidable hospitalization and high costs. Involvement of health professionals in this process is important.

In addition, as promising as the application of a risk stratification tool is, the strength of a prediction model only reaches as far as the quality of the health registries. The more primary care physicians realize the strengths of a registry of good quality, the better routinely collected data can be used for risk stratification approaches. Creating awareness amongst physicians is the first step in successful application of risk stratification tools. Not only will awareness amongst healthcare professionals lead to better registration, but it is also important for an efficient practical use of risk stratification approaches in healthcare. To create awareness amongst professionals, more evidence is needed of the effectiveness of risk stratification models. Intervention studies in which patients are selected for specific interventions with the use of risk stratification models, will contribute to this.

In conclusion, the Dutch healthcare system might truly benefit from the use of risk stratification models, especially when applied in an early stage of care provision such as primary care. The ACG system provides a solid basis to measure multimorbidity and local adjustments of the ACG's models improve results.

References

1. Hill M. (2002). McGraw Hill Concise Medical Dictionary of Modern Medicine.
2. Hillman M. Testimony before the Subcommittee on Health of the House of Committee on Ways and Means, hearing on promoting disease management in Medicare.
3. Freund T, Kunz CU, Ose D, et al. Patterns of multimorbidity in primary care patients at high risk of future hospitalization. Popul Health Manag. 2012; 15(2): 119-124.
4. Soto-Gordoa M, de Manuel E, Fullaondo A, et al. Impact of stratification on the effectiveness of a comprehensive patient-centered strategy for multimorbid patients. Health Serv Res. 2019;54(2):466-473.
5. Girwar SM., Jabroer R, Fiocco M, Sutch SP, Numans, M. E., & Bruijnzeels, M. A. (2021). A systematic review of risk stratification tools internationally used in primary care settings. Health science reports, 4(3), e329.
6. Johns Hopkins University. The ACG® system: differentiating features [Available from: https://www.hopkinsacg.org/advantage/#features. Assessed 29 Jun 2020.]
7. Steyerberg, Ewout. (2009). Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. 10.1007/978-0-387-77244-8.
8. Oxford University Press 1998. ICPC-2: Classification of Primary Care. ISBN 0 19 262802 X, page 46-8
9. Simundic AM. Measures of diagnostic accuracy: basic definitions. eJIFCC. 2009; 19(4): 203-211.
10. Haas LR, Takahashi PY, Shah ND, et al. Risk-stratification methods for identifying patients for care coordination. Am J Manag Care. 2013; 19(9): 725-732.
11. Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. Med Care. 2012; 50(2): 131-139.
12. Maltenfort MG, Chen Y, Forrest CB. Prediction of 30-day pediatric unplanned hospitalizations using the Johns Hopkins Adjusted Clinical Groups risk adjustment system. PLoS One. 2019; 14(8):e0221233.

Supplementary material

Supplementary table 1: ICPC-1 translation to codes recognized by the ACG

ICPC-1 (Dutch version)	Description ICPC-1 code	ICPC-2 (international version)	Code used in models	Code Type used in models
A04.01	Chronisch vermoeidheidssyndroom	A04	F480	ICD-10
A09.01	Nachtzweten	A09	R619	ICD-10
A09.02	Gelokaliseerd overmatig zweten	A09	R619	ICD-10
A17.00		A17	R688	ICD-10
A29.01	Hart- en vaatziekten in familie-anamnese	A29	Z824	ICD-10
A29.02	Mammacarcinoom in familie-anamnese	A21	Z803	ICD-10
A29.03	Ovariumcarcinoom in familie-anamnese	-	Z804	ICD-10
A29.04	Coloncarcinoom in familie-anamnese	A21	Z800	ICD-10
A29.05	Diabetes in familie-anamnese	-	Z833	ICD-10
A29.06	Hypercholesterolemie in familie-anamnese	-	Z834	ICD-10
A76.01	Exanthema subitum/zesde ziekte	A76	B082	ICD-10
A76.02	Erythema infectiosum/vijfde ziekte	A76	B083	ICD-10
A76.03	Hand-voet-mondziekte	A76	B084	ICD-10
A78.05	Borreliose/Lyme	A78	A692	ICD-10
A87.01	Leven met stoma	A87	Z934	ICD-10
A87.02	Status na transplantatie	A87	Z949	ICD-10
A88.01	Perniones	A88	T68	ICD-10
A88.02	Zonnesteek	A88	T670	ICD-10
A88.03	Reisziekte	A88	T753	ICD-10
A89.01	Aanwezigheid pacemaker/interne defibrillator	A89	Z950	ICD-10
A90.01	Syndroom van Down	A90	Q909	ICD-10
A91.05	Gestoorde glucosetolerantie	-	R730	ICD-10
A91.06	Subklinische hypothyreoïdie	-	E02	ICD-10
A91.07	Subklinische hyperthyreoïdie	-	E059	ICD-10
A96.01	Natuurlijke dood	A96	R99	ICD-10
A96.02	Onnatuurlijke dood	A96	R99	ICD-10
A97.02	Kinderwens	-	A97	ICPC-2
A99.01	Dragerschap met risico voor eigen persoon	-	Z229	ICD-10
A99.02	Dragerschap met risico voor nageslacht/omgeving	-	Z229	ICD-10
B72.02	Non-Hodgkin lymfoom	B72	C819	ICD-10
B81.01	Foliumzuurdeficiëntie-anemie 4	B81	D529	ICD-10
B81.02	Vitamine B12-deficiëntie-anemie	B81	D519	ICD-10
K49.01	Cardiovasculair risicomanagement (CVRM)	K49	Z13.6	ICD-11
L49.01	Valpreventie/ fractuurpreventie	L49	R268	ICD-10
T90.01	Diabetes mellitus type 1	T89	E109	ICD-10
T90.02	Diabetes mellitus type 2	T90	E119	ICD-10

ICPC-1=International Classification of Primary Care version 1, used in Dutch primary care; ICPC-2= International Classification of Primary Care version 2, internationally used; ICD-10 = International Classification of Diseases 10th revision

CHAPTER 5

CHAPTER 5

IDENTIFYING COMPLEX PATIENTS USING ADJUSTED CLINICAL GROUPS RISK STRATIFICATION TOOL

Shelley-Ann M. Girwar, Jozefine C. Verloop, Marta Fiocco, Stephen P. Sutch, Mattijs E. Numans, Marc A. Bruijnzeels

Published in the American Journal of Managed Care. 2022 Apr 1;28(4):e140-e145. doi: 10.37765/ajmc.2022.88867

Abstract

Objectives

To produce an efficient and practically implementable method, based on primary care data exclusively, to identify patients with complex care needs who have problems in several health domains and are experiencing a mismatch of care. The Johns Hopkins ACG System was explored as a tool for identification, using its Aggregated Diagnosis Group (ADG) categories.

Study Design

Retrospective cross-sectional study, using general practitioners' electronic health records, combined with hospital data.

Methods

A prediction model for patients with complex care needs was developed using a primary care population of 105,345 individuals. Dependent variables in the model included age, sex, and the 32 ADGs. The prediction model was externally validated on 30,793 primary care patients. Discrimination and calibrations were assessed by computing C statistics and by visual inspection of the calibration plot, respectively.

Results

Our model was able to discriminate very well between complex and noncomplex patients (C statistic = 0.9; 95% CI, 0.88-0.92), whereas the calibration plot suggests that the model provides overestimates of complex patients.

Conclusions

With this study, the ACG System has proven to be a useful tool in the identification of patients with complex care needs in primary care, opening up possibilities for tailored interventions of care management for this complex group of patients. Utilizing ADGs, the prediction model that we developed had a very good discriminatory ability to identify those complex patients. However, the calibrating ability of the model still needs improvement.

Introduction

As populations age and the presence of multimorbid and complex patients becomes the norm, the pressure on health systems, regarding workload as well as costs, is immense (1). Single-disease management approaches are no longer sufficient to meet the needs of an increasing number of complex patient groups, who need care oriented toward their overall health (2). In addition, strategies distinguishing different levels of complexity within a population are desirable. Population health management (PHM) approaches aim to allocate available health resources to the appropriate patient groups within the population. Risk stratification tools, such as the widely used Johns Hopkins Adjusted Clinical Groups (ACG) system, play an important role in the identification of specific patient groups for PHM, aiming to identify subgroups in whom avoidable adverse health events could be prevented. With predictive modeling, high-risk patients can be successfully selected for extensive and proactive care management programs (3).

One group for whom it seems beneficial to set up a PHM approach, including risk stratification, is that of patients with complex care needs, who have problems involving multiple health domains and experience a mismatch of care offerings with their needs. Often the consequence of this mismatch is high care utilization—in particular, of expensive and undesirable care, such as emergency or unplanned care. This group of patients was first described by Atul Gawande, MD, MPH, in his 2011 article in *The New Yorker*, "The Hot Spotters"(4). For this group of patients, it would seem that a multidisciplinary and personalized approach would be advantageous, but evidence for the effectiveness of this kind of approach is still ambiguous(5). One of the reasons is the incorrect assignment of patients to this intense but effective individualistic approach, leading to a greater mismatch in care. Hence, the first step in providing those complex patients with the appropriate necessary care is a practical and efficient identification of the population who is most likely to benefit from the intervention.

Different methods to identify patients with complex care needs have been utilized by different organizations, using various criteria and types of data sources. The heterogeneity of high-cost patients, including patients with unpreventable costs, makes a general identification based on claims data inefficient (6,7). A focus on hospitalizations or emergency department (ED) use has proven to be more effective (8). However, a complete profile of complex patients requires complete health profile data. In most health care systems, patient data are registered in fragmented silos with significant data linkage challenges; a wide application of risk stratification tools, based on complete patient profiles, is therefore limited. Clinically based predictive models, using medication and diagnostic data, are especially efficient for prospectively identifying candidates for care management programs (9). Especially in primary care led health systems, in which primary care physicians function as gatekeepers, we have the opportunity to look at more complete health profiles of patients, including medication and diagnostic data, without linkage of data between silos. Unfortunately, despite this great opportunity, evidence for validated risk stratification models to identify patients with complex care needs in primary care is lacking.

Various validated models to map different levels of multimorbidity are available. As multimorbidity plays an important role in complex patients, the level of multimorbidity is a useful tool, in conjunction with hospitalization and ED visits, in identifying these patients. Strong evidence exists for the ACG System as a tool to determine care burden or multimorbidity, as well as risks for hospitalization and ED utilization (10-12). With this research, we explored the possibilities of the ACG System as a potential tool to identify patients with complex care needs in primary care.

This study's aim is to produce an identification method for patients with complex care needs using primary care data in order to perform a PHM approach. Our goal was to answer this research question: "Is a prediction model using the ACG risk stratification tool to identify patients with complex care needs, defined as patients with problems in multiple health domains and high acute hospital care, statistically valid for use in primary care?"

Methods

Study Design and Study Population

This work was designed as a retrospective cross-sectional study. To identify patients with complex care needs, a prediction model was developed and externally validated, using 2 study populations from 2 different regions in the Netherlands. Population 1 (n = 105,345) and population 2 (n = 30,793) were used as training and validation sets, respectively. To ensure completeness of primary care data, patients were included only if they were registered with a participating general practice for the complete period from January 2016 until December 2016. As dates for registration and deregistration are unreliable within general practitioners' electronic health records (EHRs), the period of registration was established using reimbursed registration fees. For each registered patient, a registration fee is reimbursed by the general practice each quarter. As registration fees are recorded very well within the EHRs, patients were included only when 4 registration fees for 2016 were reimbursed, indicating that the person was registered with the general practice for the complete year. In supplementary table 1, all registration fee codes nationally used in the Netherlands are presented. In addition, patients were included only when linkage with the database of Statistics Netherlands, the Dutch central bureau for statistics, was possible.

For ethical reasons, deceased patients were excluded from this study.

Data and Linkage

For this study, data from different data sources were linked anonymously at the individual patient level. Data extracted from EHRs of participating general practices included individuals' general information and information on diagnoses and medications from January to December 2016. Encryption of the EHR data was performed by Statistics Netherlands under strict rules to secure individuals' privacy. A record identification number (RIN) was assigned to each individual based on birth date, gender, and complete postal code. With the RINs, linkage of EHR data with Statistics Netherlands' microdata was made possible within the secured environment of Statistics Netherlands. Microdata comprise different types of non–publicly available data on an individual level. Under strict conditions, these microdata are accessible for statistical and scientific research. For this study, we linked the encrypted EHR data to the number of acute care visits extracted from the Dutch medical specialty information system, available as microdata within the secured environment of Statistics Netherlands. Acute care visits were defined as acute hospital care visits to the ED or to another hospital department in which an emergency care practitioner was involved.

Definition of Patients With Complex Care Needs

Adapted from Gawande's "The Hot Spotters" definition, 2 prerequisites were set to define patients with complex care needs. The first prerequisite concerns having health problems in at least 2 of 3 different health domains registered within primary care. Problems relating to the chronic physical, the mental, and the social domains were identified by selecting corresponding International Classification of Primary Care version 1 (ICPC-1) diagnosis codes within the EHR data. Supplementary table 1 gives an overview of the ICPC-1 codes for all 3 domains. The second prerequisite concerns having at least 2 acute care visits in a 12-month period (January-December 2016), considered high acute care utilization. Acute care visits were identified from health care activity codes within the medical specialty data set available as microdata. Health care activity codes are used nationwide by health care insurers and providers. In supplementary table 1, the specific codes corresponding with acute health care use are presented.

Statistical Analysis

To assess the ACG System as an identification method for patients with complex care needs, we developed a prediction model by using logistic regression analysis in the first study population, then externally validated the model in the second study population. To investigate the similarity of both study populations, individuals' characteristics were compared by using χ^2 and independent *t* tests for categorical and continuous variables, respectively.

Dependent variables

The dependent variable, or the outcome, in the logistic regression analysis was whether or not a person was identified as a complex patient (as described above).

Independent variables

Aggregated Diagnosis Groups (ADGs), the ACG System's categorization of diagnosis types, were used as independent variables. ICPC-1 diagnosis codes are clustered into 32 ADGs by the ACG software. All 32 ADGs, as well as sex and age categories, were included as independent variables within the logistic regression model. Age categories were based on ACG age categorization and clustered into 6 categories: aged 0 to 11 years, aged 12 to 34 years, aged 35 to 54 years, aged 55 to 69 years, aged 70 to 79 years, and aged 80 years or older.

Assessment of the model

The prediction model was assessed on both discriminatory and calibrating ability. To assess the discriminatory ability of the model, C statistic values were calculated in both the training and validation data sets. Values below 0.6 were considered poor; between 0.6 and 0.7, sufficient; between 0.7 and 0.8, good; between 0.8 and

0.9, very good; and above 0.9, excellent (13). The calibrating ability of the model was assessed by dividing the validation data set population into deciles based on ascending predicted values for being a complex patient. For each group, the mean observed and expected values were plotted against each other in the calibration plot. Models with a 45-degree angle plot (mean observed value equals mean expected value) were considered perfectly calibrated. Models below this reference line are overestimating, whereas models above it are underestimating.

Results

Population Characteristics

Population characteristics were compared for both the training and validation data sets. Table 1 shows the differences in population characteristics between the 2 data sets. The 2 study populations had comparable mean ages of 40.8 and 40.6 years, respectively. The percentage of women was 51.4% and 50.1%, respectively, and the percentage of complex patients was comparable in both data sets at 0.9% and 0.8%.

Table 1: Patient Characteristics Compared Between the 2 Populations Used for the Prediction Models

	Population 1 (n = 105,345)	**Population 2 (n = 30,793)**	***P***
Age in years, mean	40.8	40.6	<.001
Sex, % women	51.4	50.1	<.001
Complex patients, %	0.9	0.8	.072
Acute care visits, mean number	0.16	0.10	<.001
≥2 acute care visits, %	2.7	1.6	<.001
Health domains, %			
Chronic physical	39.8	46.8	<.001
Social	10.9	20.2	<.001
Mental	26.6	40.6	<.001
Common conditions, %			
Depression	8.5	10.3	<.001
Diabetes	6.7	5.9	<.001
Hypertension	20.1	20.3	.422
Ischemic heart disease	2.5	3.6	<.001
Asthma	11.5	13.5	<.001
Chronic obstructive pulmonary disease	3.1	2.8	.003

Population 1 is the training data set; population 2 is the validation data set; for comparison, χ^2 tests for categorical and independent *t* tests for continuous variables were performed.

The percentage of individuals with at least 2 acute care visits in 1 year was lower in the validation data set than in the training data set: 1.6% vs 2.7%. There were more health problems observed in the validation data set. Somatic chronic diseases were more prevalent in the validation data set, at 46.8% compared with 39.8%. In addition, a higher prevalence was found of psychiatric health problems (40.6% vs 26.6%) and of social health problems (20.2% vs 10.9%). Common conditions such as depression, diabetes, hypertension, ischemic heart disease, asthma, and chronic obstructive pulmonary disease were comparably prevalent in both populations.

Model Performance

Odds ratios

Figure 1 provides an overview of the odds ratios (ORs) for each variable in the model. Supplementary table 2 gives additional information about the ORs along with their 95% CIs.

Age and sex were not statistically significantly associated with the outcome. The following main predictors were identified: ADG categories *chronic medical, stable* (OR, 2.78; 95% CI, 2.35-3.29) and *chronic medical, unstable* (OR, 2.89; 95% CI, 2.49-3.35); *stable psychosocial* ADGs (OR, 3.36; 95% CI, 2.92-3.86); and *unstable psychosocial* ADGs (OR, 2.95; 95% CI, 2.33-3.73). In addition to the chronic medical and the psychosocial ADGs, the 3 ADGs *time limited: major* (OR, 2.37; 95% CI, 1.73-3.25), *injuries/adverse effects: major* (OR, 2.30; 95% CI, 1.89-2.79), and *signs/ symptoms: uncertain* (OR, 2.95; 95% CI, 2.50-3.49) were also highly associated with the outcome.

Discriminatory ability

The discriminatory ability of the prediction model for identifying complex patients was assessed by calculation of C statistics. The C statistics, estimated for the training and validation data sets, are presented in table 2.

Calibrating ability

Figure 2 shows the calibrating ability of the prediction model. The model is overestimating, meaning that it is identifying more complex patients than were observed in the study population. In supplementary table 3, the observed number of complex patients, mean expected values, and standard errors are presented for each decile of population 2.

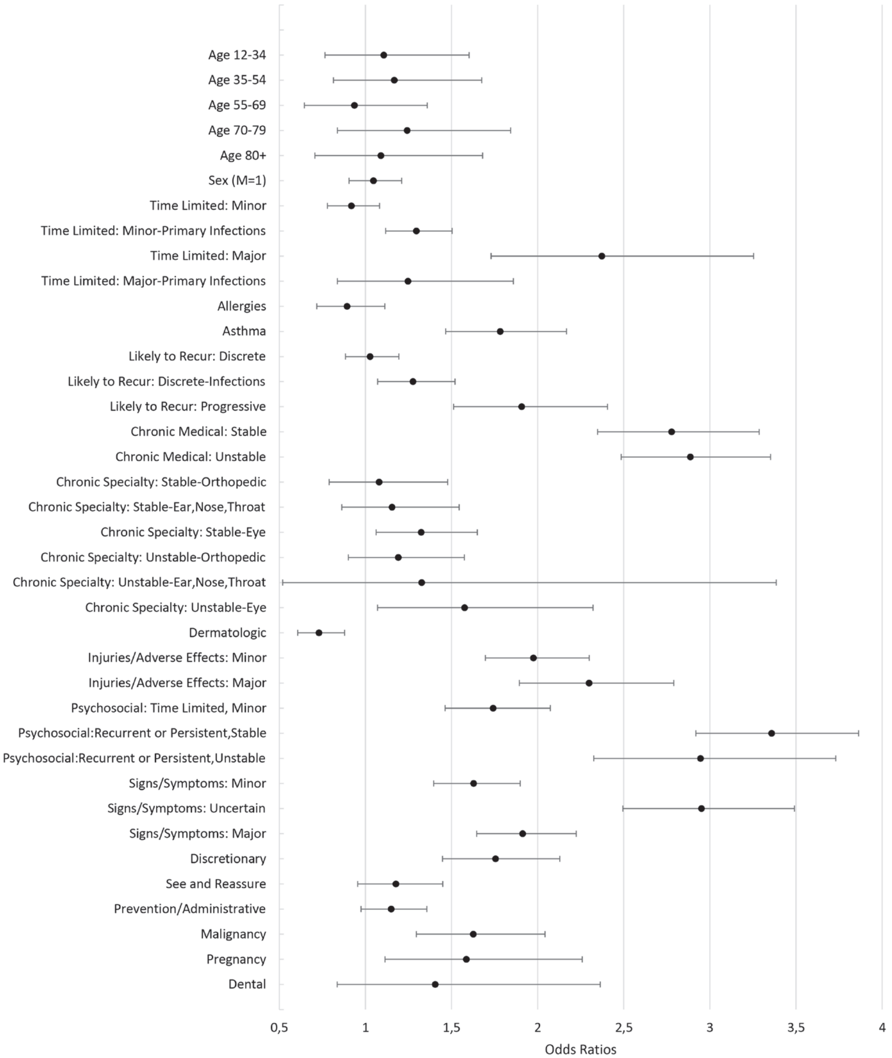


Figure 1: Estimated Odds Ratios and Their 95% Confidence Intervals. Variables include Aggregated Diagnosis Groups, a categorization of diagnosis types by the Johns Hopkins ACG System.

Table 2: Performance of Prediction Model for Being a Complex Patient

	C statistic	95% CI
Training data set	0.913	0.905 - 0.920
Validation data set	0.899	0.882 - 0.915

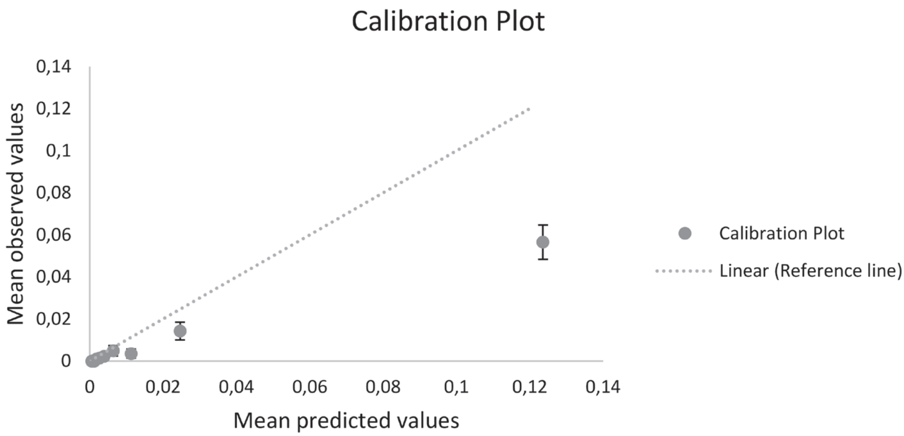


Figure 2: Calibration Plot: Observed vs Predicted Values (estimated by the prediction model for being a complex patient)

Discussion

With this study, we have developed a predictive model for patients with complex care needs. Although a good discriminatory performance is shown, the calibrating ability is modest because more complex patients were identified based on the predicted values than were observed. This study confirms results of previous studies, showing that the ACG System ADGs form a good tool to describe and determine the level of multimorbidity (14,15), which plays an important role in describing complex patient groups. With the highest weightings in the model for stable psychosocial problems, the proven prevalence of psychosocial conditions among our group of complex patients (16) is taken into account. Stable and unstable chronic medical problems and unstable psychosocial problems are strongly associated with the outcome. Following chronic medical and psychosocial problems, uncertain signs or symptoms also had a high predictive effect on being a complex patient.

In our study, a statistics-based quantitative tool was found to identify patients with complex care needs. As our model is overestimating complex patients, this tool based on routinely collected data should be used as a first screening method. Additional qualitative screening of identified complex patients is of great importance to select the right subpopulation for interventions. Additional qualitative screening also allows for distinction between patients with high care needs that can be avoided with proactive care, and patients in whom high care needs cannot be avoided. As our model was not built to distinguish between avoidable and unavoidable emergency care, this addition of qualitative screening is of great importance. It is therefore recommended, to investigate the practical use of our identification method, intervention studies with complex patients identified with our screening method should follow.

Further, high-quality registration is a necessary condition to be able to use complete health records for the identification of complex patients. Results of this study show that the prevalence of registered somatic chronic problems, as well as psychiatric and social health problems, differed between both study populations. These differences may have been due to dissimilar registration policies rather than differences in prevalence. For efficient and practical PHM approaches using data-driven identification methods to designate resources to the patient groups in which they will be most beneficial, good quality of registration is important. Once practitioners realize that registry data can be used for producing prediction models that are helpful in practice, improvement of registration habits is expected. In addition to the differences between the two study populations, the prevalence of social problems seemed lower than expected. As studies have shown that complex patients are more likely to have underlying social problems, such as low income, living alone, living in a deprived area, and being less likely to own a home (17-19), alongside chronic and psychosocial problems, the registration of social problems in primary care should be emphasized and stimulated. Alternatively, the identification of vulnerable and complex patient groups may be improved by creating access to social data sources. However, most risk stratification tools such as the ACG System do not currently include social data, as it is less routinely collected. In addition, linkage of social data to primary care data still entails significant information security and legal issues.

An accurate identification of complex patients in primary care can be of great value to health care systems. Not only can early identification and intervention prevent deterioration of patients' health, but health resources can also be put towards groups of patients who will benefit most from them. As general practitioners' workload is increasing, resulting in rising pressures on them in most European countries (20), allocation of available resources in primary care is of utmost importance. In situations in which different types of health care data can easily be combined, such as in integrated and managed care organizations, more complete profiles of patients can be used to allocate resources to the patient groups for whom they will be most beneficial. Risk stratification approaches to identify subpopulations have proven valuable in forms of both tailored care interventions and improved care management (21,22). Effective PHM interventions following identification of complex groups of patients may lead to improved health outcomes, not only for these complex groups, but for the whole population.

Limitations

This research is based on a definition of complex patients that is principally composed of biomedical characteristics. As social determinants are underrepresented in primary care registration, defining complex patients was mostly limited to the somatic physical and psychiatric diagnoses of patients. Although we believe that we have identified a group of complex patients with unmet health needs, the social health burden is most likely to be under recorded. This may have caused our selected population of complex patients (outcome) to ignore other important groups of patients. As our PHM approach using primary care data is a holistic approach, the biomedical focus of registry systems may still cause important gaps in patients' complete care profiles.

Further, the age categories used in this study are broad. We explored the use of smaller age categories (eg, 5-year age bands) and the use of age as a continuous variable, but age did not contribute statistically significantly to the identification of patients with complex care needs adjusting for other factors.

Lastly, by selecting people registered with one of the participating general practices for a period of 12 months, people born within this 12-month period were excluded from this study. However, there are very few infants who would meet our requirements to be classified as a patient with complex care needs according to our definition.

Conclusions

Using broad morbidity groups of diagnoses or comorbidities (such as ADGs) seems to be an effective approach to identify complex patients in primary care. Risk stratification tools are key methods in putting available registry data to good use to identify complex patient groups. In addition to biomedical health determinants, social determinants play an important role in the identification of complex patients. Effective identification of complex patients in primary care can result in appropriate and proactive care management for this group of patients, benefiting the total primary care population.

References

1. Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med. 2007;22(suppl 3):391-395.
2. Starfield B, Lemke KW, Bernhardt T, Foldes SS, Forrest CB, Weiner JP. Comorbidity: implications for the importance of primary care in 'case' management. Ann Fam Med. 2003;1(1):8-14.
3. Freund T, Mahler C, Erler A, et al. Identification of patients likely to benefit from care management programs. Am J Manag Care. 2011;17(5):345-352.
4. Gawande A. The hot spotters. The New Yorker. January 24, 2011:40-51.
5. Roblin DW, Segel JE, McCarthy RJ, Mendiratta N. Comparative effectiveness of a complex care program for high-cost/high-need patients: a retrospective cohort study. J Gen Intern Med. 2021;36(7):2021-2029.
6. Joynt KE, Gawande AA, Orav EJ, Jha AK. Contribution of preventable acute care spending to total spending for high-cost Medicare patients. JAMA. 2013;309(24):2572-2578.
7. Lee NS, Whitman N, Vakharia N, Taksler GB, Rothberg MB. High-cost patients: hot-spotters don't explain the half of it. J Gen Intern Med. 2017;32(1):28-34.
8. Iovan S, Lantz PM, Allan K, Abir M. Interventions to decrease use in prehospital and emergency care settings among super-utilizers in the United States: a systematic review. Med Care Res Rev. 2020;77(2):99-111.
9. Forrest CB, Lemke KW, Bodycombe DP, Weiner JP. Medication, diagnostic, and cost information as predictors of high-risk patients in need of care management. Am J Manag Care. 2009;15(1):41-48.
10. Haas LR, Takahashi PY, Shah ND, et al. Risk-stratification methods for identifying patients for care coordination. Am J Manag Care. 2013;19(9):725-732.
11. Huntley AL, Johnson R, Purdy S, Valderas JM, Salisbury C. Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. Ann Fam Med. 2012;10(2):134-141.
12. Lemke KW, Weiner JP, Clark JM. Development and validation of a model for predicting inpatient hospitalization. Med Care. 2012;50(2):131-139.
13. Šimundić AM. Measures of diagnostic accuracy: basic definitions. EJIFCC. 2009;19(4):203-211.
14. Sibley LM, Moineddin R, Agha MM, Glazier RH. Risk adjustment using administrative data-based and survey-derived methods for explaining physician utilization. Med Care. 2010;48(2):175-182.
15. Starfield B, Weiner J, Mumford L, Steinwachs D. Ambulatory care groups: a categorization of diagnoses for research and management. Health Serv Res. 1991;26(1):53-74.
16. Mautner DB, Pang H, Brenner JC, et al. Generating hypotheses about care needs of high utilizers: lessons from patient interviews. Popul Health Manag. 2013;16(suppl 1):S26-S33.
17. Colligan EM, Pines JM, Colantuoni E, Wolff JL. Factors associated with frequent emergency department use in the Medicare population. Med Care Res Rev. 2017;74(3):311-327.
18. Fitzpatrick T, Rosella LC, Calzavara A, et al. Looking beyond income and education: socioeconomic status gradients among future high-cost users of health care. Am J Prev Med. 2015;49(2):161-171.

19. Krieg C, Hudon C, Chouinard MC, Dufour I. Individual predictors of frequent emergency department use: a scoping review. BMC Health Serv Res. 2016;16(1):594.
20. Schäfer WLA, Boerma WGW, Spreeuwenberg P, Schellevis FG, Groenewegen PP. Two decades of change in European general practice service profiles: conditions associated with the developments in 28 countries between 1993 and 2012. Scand J Prim Health Care. 2016;34(1):97-110.
21. Burton LC, Skinner EA, Uscher-Pines L, et al. Health of Medicare Advantage plan enrollees at 1 year after Hurricane Katrina. Am J Manag Care. 2009;15(1):13-22.
22. Soto-Gordoa M, de Manuel E, Fullaondo A, et al. Impact of stratification on the effectiveness of a comprehensive patient-centered strategy for multimorbid patients. Health Serv Res. 2019;54(2):466-473.

Supplementary Material

Supplementary table 1: Overview of specific codes corresponding with different health domains and acute healthcare use:

Characteristic	Coding system	Specific codes
Diagnoses in the chronic physical domain	International Classification for Primary Care, version 1 (ICPC-1)	A28, A79, A90, B28, B72, B73, B74, B78, B79, B83, B90, D28, D74, D75, D76, D77, D81, D92, D94, D97, F28, F81, F83, F84, F91, F93, F94, H28, H80, H83, H84, H85, H86, K28, K73, K74, K76, K77, K82, K86, K87, K90, K91, K92, L28, L82, L84, L85, L88, L89, L90, L91, L95, L98, N28, N70, N74, N85, N86, N87, N88, R28, R84, R85, R89, R91, R95, R96, S28, S77, S81, S83, S87, S91, T28, T71, T78, T80, T81, T86, T90, T92, T93, U28, U75, U76, U77, U85, U88, W28, W72, W76, X28, X75, X76, X77, X83, X88, Y28, Y77, Y78, Y82, Y84
Diagnoses in the mental domain	ICPC-1	P01, P02, P02.01, P03 to P06, P09, P10, P10.01, P10.02, P11, P15 to P19.02 (not P17), P20, P21, P22, P23, P24, P25, P27, P28,P29, P70 to P77.02, P78, P79 to P80.01, P80.02, P 85, P98, P99, P99.01, T06 to T06.02.
Diagnoses in the social domain	ICPC-1	Z01 to Z29
Acute healthcare utilization	Dutch healthcare activities coding system	190060/190013 *in combination with emergency physician*, 190015, 190016

For the three health domains, ICPC-1 codes corresponding to the three health domains including the chronic physical, the mental and the social health domain, are presented. For acute healthcare utilization, specific healthcare activity codes are presented

Supplementary table 2: Odds ratios, including the 95% confidence intervals, for all in the model included independent variables

	Odds Ratio	95% Confidence interval	
		Lower bound	Upper bound
Age 12-34 year	1.107	0.765	1.602
Age 35-54 year	1.168	0.814	1.676
Age 55-69 year	0.936	0.645	1.359
Age 70-79 year	1.242	0.837	1.844
Age 80+ year	1.090	0.707	1.680
Sex (M=1)	1.047	0.905	1.211
1 Time Limited: Minor	0.918	0.779	1.082
2 Time Limited: Minor-Primary Infections	1.296	1.117	1.503
3 Time Limited: Major	2.372	1.730	3.254
4 Time Limited: Major-Primary Infections	1.247	0.837	1.859
5 Allergies	0.894	0.718	1.113
6 Asthma	1.783	1.467	2.167
7 Likely to Recur: Discrete	1.028	0.884	1.194
8 Likely to Recur: Discrete-Infections	1.276	1.071	1.520
9 Likely to Recur: Progressive	1.907	1.512	2.406
10 Chronic Medical: Stable	2.778	2.348	3.286
11 Chronic Medical: Unstable	2.886	2.485	3.352
12 Chronic Specialty: Stable-Orthopedic	1.080	0.789	1.477
13 Chronic Specialty: Stable-Ear, Nose, Throat	1.154	0.863	1.544
14 Chronic Specialty: Stable-Eye	1.324	1.062	1.650
16 Chronic Specialty: Unstable-Orthopedic	1.191	0.900	1.574
17 Chronic Specialty: Unstable-Ear, Nose, Throat	1.327	0.520	3.385
18 Chronic Specialty: Unstable-Eye	1.576	1.070	2.321
20 Dermatologic	0.731	0.607	0.879
21 Injuries/Adverse Effects: Minor	1.975	1.696	2.300
22 Injuries/Adverse Effects: Major	2.299	1.894	2.790
23 Psychosocial: Time Limited, Minor	1.741	1.462	2.073
24 Psychosocial: Recurrent or Persistent, Stable	3.358	2.919	3.863
25 Psychosocial: Recurrent or Persistent, Unstable	2.946	2.326	3.730
26 Signs/Symptoms: Minor	1.628	1.396	1.899
27 Signs/Symptoms: Uncertain	2.951	2.495	3.490
28 Signs/Symptoms: Major	1.913	1.646	2.224
29 Discretionary	1.755	1.447	2.128
30 See and Reassure	1.177	0.955	1.449
31 Prevention/Administrative	1.150	0.974	1.357
32 Malignancy	1.627	1.295	2.043
33 Pregnancy	1.586	1.114	2.259
34 Dental	1.406	0.836	2.364

Dependent variables include age categories, sex and 32 Aggregated Diagnosis Groups from the Johns Hopkins Adjusted Clinical Groups system. (Aggregated Diagnosis Groups are an aggregation of the International Classification of Primary Care diagnosis codes.)

Supplementary table 3: Summary statistics of the ten deciles of the population on which the calibration plot is based

Total n per group	Number of complex patients	Mean observed values	Mean predicted values	Standard Error
3291	0	0,000	0,001	0,000
2879	0	0,000	0,001	0,000
3062	0	0,000	0,001	0,000
3086	3	0,001	0,002	0,001
3073	4	0,001	0,003	0,001
3086	7	0,002	0,004	0,001
3080	15	0,005	0,006	0,001
3079	11	0,004	0,011	0,001
3081	44	0,014	0,025	0,002
3077	174	0,057	0,124	0,004

CHAPTER 6

CHAPTER 6

GENERAL DISCUSSION

The objective of this thesis was to find a risk stratification tool that can be used in Dutch primary care covering total primary care populations and answering the research question *'What risk stratification tool is most suitable for Dutch primary care and how can this tool be used appropriately using Dutch routine primary care data?'* In our studies we identified a risk stratification tool to be used in Dutch primary care. In addition, we added recommendations to improve the tool according to the Dutch primary care setting.

From our systematic literature review, the ACG had proven to be not only the most frequently used risk stratification model in primary care, but also the most accurate model to predict different types of both current and future healthcare utilization. Our findings have confirmed these results in the Dutch primary care context. With our pilot study, the ACG has shown great promise to identify different levels of multimorbidity and burden of healthcare in the Dutch primary care (general practice) setting. Models using the ACG's diagnosis and multimorbidity categories, the ADGs and ACGs respectively, perform well in predicting the amount of GP visits. In addition, the ACG models for predicting future hospitalization and high costs showed good performances in terms of discrimination and calibration properties. When adapting the models to fit the Dutch primary care data by adjusting the coefficients of the underlying predictors, model performances even improved. The ACG is also an appropriate tool to identify complex patients with problems in multiple health domains. With these results, the ACG has proven to be a well performing risk stratification tool that can be effectively used in Dutch primary care, predicting healthcare utilization in the form of GP visits, future hospitalization and future high healthcare costs.

Methodological improvements of the models

Focusing on specific aspects to improve the models for use in Dutch primary care, we can enhance to the already highly performing models. Two important methodological aspects need to be considered for further optimizing the models for risk stratification purposes. Firstly, the clinical validity of the predictors underlying the models, needs to be investigated. Although the statistical validity of the models has been assessed with this research, the predictors are based on US data. The predictor 'pregnancy' for example is a contributing factor for both hospitalization and the high costs risk, with ORs greater than four. However, the inclusion of pregnancy within the models is based on the US context where pregnancy may pose much greater risk than in the Netherlands. Therefor the clinical significance of the underlying predictors needs to be validated for the Dutch setting, as clinical characteristics may have different importance in various countries. We recommend the use of focus groups and Delphi studies which include local clinicians to dive deeper into the clinical validity of the underlying predictors of the risk stratification models.

Secondly, enriching the data by adding diagnostic measurements such as laboratory values, may be interesting to include in the models and further research to investigate the added value is needed. The ACG software has recently been updated to include diagnostic measurements, which were not included in this study. This could be considered for further optimization of the Dutch version of the models. However, a balance should be struck between putting more effort into predicting more precisely based on more biomedical data and focusing on improving predictions based on contextual data, broadening the holistic profile of patients.

Why risk stratification in primary care?

The focus of this study is on risk stratification specifically performed in the primary care setting. My assertation is that performing risk stratification in primary care is beneficial to the whole healthcare continuum and can improve the health of the total population, because it can be used as the basis for further targeting interventions. By using primary care data to identify patients at risk for high healthcare utilization, such as high numbers of GP visits, hospitalization and high healthcare costs, high risk patients can be detected at an early stage and appropriate interventions can be provided proactively. Using risk stratification in primary care is therefore beneficial for the total population, especially in countries where primary care functions as the gatekeeper for specialty care. The fact that in such countries, the majority of the population is registered with a GP, makes it even more promising to use risk stratification in primary care. With almost complete profiles of patients available in primary care registries, which cover nearly the total population, these data are very valuable for PHM purposes. When moving risk stratification to secondary care, the chance exists of providing care too late in the care continuum with the risk of providing care in a reactive rather than proactive manner.

Another reason why risk stratification in primary care gives excellent opportunities for improvement of healthcare, is the enormous pressure currently put on primary care especially in Western countries. Due to the diminishing workforce and increasing complexity of patients, GPs find themselves overloaded with work, overwhelmed and even burned out, resulting in degradation of healthcare quality (1-4). Broadening the tasks of GPs (4), for example with increased policies to move more care from specialists towards primary care, is increasing this burden. With efficient use of available resources and time, approaches to ease the rising pressure on GPs are possible. The ACG algorithm has a large numbers of predictors, that can be applied for management of care provision and planning. In addition, with the use of risk stratification, referrals to specialty care can be coordinated and managed in a more efficient and effective manner. A clear distinction between patients that can be treated in primary care and patients that need referrals to specialty care can be made using patients' profiles based on registered data.

First steps to be taken: creating a 'fertile' environment

However, to effectively take advantage of the benefits of risk stratification in primary care, healthcare professionals need to collaborate and be aware of the direct and indirect benefits of risk stratification. In the Netherlands, the typical GP is not yet been convinced of these benefits. This has also been encountered within this study: participating GPs were not ready to continue with the risk stratification-based project that was part of this dissertation, since they were not yet convinced of how this form of risk stratification would benefit their personal caregiving.

Awareness and trust in the models need to be created amongst healthcare professionals. A study by Wagner et al. showed that for healthcare professionals to feel at ease using risk stratification algorithms, considerations such as trust in the algorithms play a strong role (5). Healthcare professionals need to be convinced that the already established algorithm is specific enough to the practice's patient population in order to feel at ease using such a model (5). During our study, we have tried to create trust amongst healthcare professionals, in particular GPs, by statistically validating specific risk stratification algorithms. However, during the project, we have observed that the participating GPs are not convinced of the added value of the algorithms to benefit their specific patient populations. Adjusting the algorithms with input of healthcare professionals can create more trust. Wagner and colleagues showed that healthcare professionals that didn't trust existing risk stratification algorithms enough to benefit their patient population, were still positive towards performing risk stratification by using their own criteria for risk stratification (5). Another way to create trust amongst healthcare professionals towards already existing risk stratification algorithms is to prove the added value of the algorithms for their patient populations and show the benefits of interventions using risk stratification tools in Dutch primary care. Selecting the right patients for specific interventions by using risk stratification strategies in primary care and proving the effects on GPs' workload and the health of their patient populations, can show the benefits for healthcare professionals and their practices as a whole, creating a solid support base for risk stratification based-approaches. An example of the benefits of using risk stratification is shown by a study by Hewner et al., where risk stratification was used to appoint cohorts of patients with high risk of hospitalization (6). With targeted case management these patients received appropriate care in order to avoid unnecessary hospitalizations (6), resulting in increased health of the population. A first attempt to investigate the benefits of risk stratification by selecting specific patients for interventions, has already been set into action in Dutch primary care. One of the algorithms resulting from this study, the algorithm to identify complex patients with the use of the ACG software, will be used to select patients for an intervention study. Complex patients, the so called 'hotspotters', will be selected for a proactive integrated care approach in different general practices in the Netherlands. The effects of the intervention on the health and care needs of the participating patients as well as the cost-effectiveness of the approach will be measured.

Next to trust in the algorithms, Wagner and colleagues found that practical implementation of risk stratification models is another key consideration to determine practices' attitudes towards risk stratification. Practical implementation requires the availability of technical skills and consistent information systems, but also the assistance to generate user friendly reports that healthcare professionals can use during their consultations. (5) Within our study, we have successfully implemented risk stratification algorithms using the available information systems. However, we observed differences in available technical skills between practices. We also encountered the disadvantages of practices not getting the right assistance in generating user friendly information and the algorithm software not then aligning with their workflow. These are aspects that result in hesitation and skepticism of GPs towards the use of these risk stratification algorithms. Attention needs to be paid to these considerations regarding practical implementation alongside the creation of trust among healthcare professionals.

As it is important to generate holistic patient profiles, integration of information systems would be beneficial. This requires the addition of stakeholders other than GPs, either within or outside the health sector. However, persuasion of other stakeholders is of immense importance for other reasons too. Once stakeholders such as insurance companies, health policy makers and municipal authorities, realize the benefits of risk stratification in primary care, support for implementation can be reached on national, regional, and local governance level, promoting integration between these different parties. In the UK where the National Health Service functions as the overarching government-funded institute that organizes healthcare nationally, this integration of care is a key policy at national level. In the United States, different stakeholders are bundled into Accountable Care Organizations (ACOs) in order to coordinate care for patients in a more efficient way, and in Germany the integrated care provision in the Gesundes Kinzigtal approach yields great results (7). The use of data for risk stratification purposes resulting in efficient provision of care is made much easier when care is more integrated for example in situations such as the UK, Gesundes Kinzigtal and the ACOs in the US. It goes without saying that GPs need to stay included. With the responsibility of the health of their own population, trust and awareness amongst GPs is needed in order for them to engage with other stakeholders towards more integration of care.

As financial benefits are most appealing for the different stakeholders, the beneficial financial implications of risk stratification should also be considered in above mentioned intervention studies in order to convince payers of a return on investment. Although decreases in expensive healthcare utilization, such as hospitalizations, as a result of risk stratification approaches have already been proven internationally (6, 8), evidence within Dutch primary care is still needed. A great example of a successful effort that has been made in the Netherlands is the Lung care program that has been set up in Nijkerk (9). Using the available data, patients with Chronic pulmonary obstructive disease (COPD) are risk stratified in order to provide efficient care. This was done in a collaborative setting, including healthcare professionals such as GPs and respiratory specialists, social care and voluntary organizations and one of the leading insurance companies in the region of Nijkerk, the Netherlands. Special financial arrangements between healthcare professionals and the insurance company were made and reported in the Health Affairs article, a million dollars saving was accomplished within three years. With the savings, expansion to diabetes and cardiac care is being implemented. With more evidence of programs such as the Lung care program in Nijkerk, the right stakeholders can be attracted to participate in risk stratification approaches used in integrated programs on regional and even national level.

Future improvements

Although primary care seems an appropriate place within the care continuum to implement risk stratification approaches, benefits can be strengthened by linking primary care data with data from the social domain. Including social determinants in risk stratification strategies enriches the data, posing the opportunity to make predictions even more accurate, and enables to provide care following the specific risks in various subpopulations. Literature has shown associations of social determinants such as educational level (10) and income (11, 12) with level of healthcare utilization. Despite current privacy challenges regarding linkage of different data sources, cooperation between different healthcare and wellbeing sectors should still be pursued. Collaboration and integration on different levels, including information systems, can maximize the benefits of risk stratification approaches in primary care. Including the social sector in risk stratification approaches provides opportunities for more holistic approaches, where patient profiles extend beyond biomedical characteristics. Including cultural and ethnical components for risk stratification can also be of added value as ethnicity has been proven to be associated with healthcare utilization (13) and health outcomes (14). Agyemang *et al.* and Uitewaal *et al.* found that although there is no difference in treatment between indigenous Dutch and ethnic minorities in the Netherlands, health outcomes are worse for ethnic minorities (14, 15). This indicates that ethnicity has great potential to be recognized in the implementation of risk stratification algorithms. However, using ethnicity in risk stratification algorithms can be controversial, as this can be seen as evidence of institutional racism. With our current healthcare system we often strive for equality in the care provided, which means that every person receives the same care and is treated the same no matter the situation. Within healthcare systems fairness is often translated as providing everyone with equal care, excluding every form of discrimination. However, with the realization that fairness is not merely the provision of equal care, but providing equity in care, discriminating between different risk groups, which may include ethnic minorities, should be adopted in the provision of improved care. However, these discussions are being complicated in the Netherlands when institutions such as the Dutch tax authorities are suspected of institutional racism in a series of cases where multiple beneficiaries for child allowance were disadvantaged based on ethnical profiling algorithms. Despite these controversies, it is my belief that including ethnic background amongst other social determinants, can be of great added value to risk stratification approaches. There is a need for more investigation of this added value as well as the correct use of those determinants without harming privacy and dignity of individuals within the population. Using determinants underlying to ethnicity, such as cultural behaviour and socio-economic status, should also be taken into account, which might soften the controversy.

On the other side, a majority of research has been undertaken with data originating from white populations, not taking into account the determinants for ethnic minorities. This study, has not included the multicultural setting of the Netherlands and thus the risk stratification algorithms are mostly based on native Dutch inhabitants. The exclusion of ethnic minorities in research can also be considered as institutional racism, as cultural determinants are not taken into account. An important recommendation is therefore to validate the risk stratification algorithms resulting from this thesis in more multicultural settings of the Netherlands, such as Amsterdam and the Hague.

Conclusion

In my opinion, to gain the best effects for healthcare, risk stratification approaches can best be performed in primary care, since this is the setting closest to real life in which most health issues of most citizens, at least administratively, come together. In our research, several risk stratification models using the ACG tool, have shown good performance in the Dutch primary care setting. Methodological improvements can be achieved by optimizing the models, especially with regards to the clinical validity of the predictors. However, next to the methodological improvements, practical implementation needs to be taken into account. Both an increase in confidence in reliability and applicability among healthcare professionals, as well as improvement of technical implementation tools for risk stratification, need to be considered. Inclusion of sectors outside healthcare, for example the enrichment of data as input for the models with social domain data, can add great value to risk stratification approaches in primary care.

References

1. Hall LH, Johnson J, Watt I, O'Connor DB. (2019). Association of GP wellbeing and burnout with patient safety in UK primary care: a cross-sectional survey. British Journal of General Practice, 69(684), e507-e514.
2. Pedersen AF, Nørøxe KB, Vedsted P. (2020). Influence of patient multimorbidity on GP burnout: a survey and register-based study in Danish general practice. British Journal of General Practice, 70(691), e95-e101.
3. Picquendar G, Guedon A, Moulinet F, Schuers M. (2019). Influence of medical shortage on GP burnout: a cross-sectional study. Family Practice, 36(3), 291-296.
4. Schäfer WL, Boerma WG, Spreeuwenberg P, Schellevis FG, Groenewegen PP. Two decades of change in European general practice service profiles: conditions associated with the developments in 28 countries between 1993 and 2012. Scand J Prim Health Care. 2016;34(1):97–110.
5. Wagner J, Hall JD, Ross RL, Cameron D, Sachdeva B, Kansagara D, et al. (2019). Implementing risk stratification in primary care: challenges and strategies. The Journal of the American Board of Family Medicine, 32(4), 585-595.
6. Hewner S, Seo JY, Gothard SE, Johnson BJ. (2014). Aligning population-based care management with chronic disease complexity. Nursing outlook, 62(4), 250-258.
7. Hildebrandt H, Pimperl A, Schulte T, Hermann C, Riedel H, Schubert I, et al. (2015). Triple Aim–Evaluation in der Integrierten Versorgung Gesundes Kinzigtal–Gesundheitszustand, Versorgungserleben und Wirtschaftlichkeit. Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz, 58(4), 383-392.
8. Soto-Gordoa M, de Manuel E, Fullaondo A, et al. Impact of stratification on the effectiveness of a comprehensive patient-centered strategy for multimorbid patients. Health Serv Res. 2019;54(2):466-473. doi:10.1111/1475-6773.13094
9. Marill MC. (2020). From Rural Germany, Integrated Care Grows Into A Global Model: Article describes the challenges and successes of an accountable care organization in rural Germany created several years before the shared savings model emerged from the Affordable Care Act.
10. Glazier RH, Agha MM, Moineddin R, Sibley LM. (2009). Universal health insurance and equity in primary care and specialist office visits: a population-based study. The Annals of Family Medicine, 7(5), 396-405.
11. Shadmi E, Balicer RD, Kinder K, Abrams C, Weiner JP. (2011). Assessing socioeconomic health care utilization inequity in Israel: impact of alternative approaches to morbidity adjustment. BMC public health, 11(1), 1-8.
12. Sibley LM, & Glazier RH. (2012). Evaluation of the equity of age–sex adjusted primary care capitation payments in Ontario, Canada. Health Policy, 104(2), 186-192.
13. Stronks K, Ravelli ACJ, Reijneveld SA. Immigrants in the Netherlands: Equal access for equal needs? J Epidemiol Community Health 2001;55:701-707.
14. Agyemang C, Bindraban N, Mairuhu G, Montfrans van G, Koopmans R, Stronks, K. Prevalence, awareness, treatment, and control of hypertension among Black Surinamese, South Asian Surinamese and White Dutch in Amsterdam, The Netherlands: the SUNSET study. J Hypertens 2005;23:1971-1977.
15. Uitewaal PJ, Bruijnzeels MA, Bernsen RM, Voorham AJ, Hoes AW, Thomas S. Diabetes care in Dutch general practice: differences between Turkish immigrants and Dutch patients. Eur J Public Health 2004;14(1):15-8.

CHAPTER 7

CHAPTER 7

SUMMARY

As healthcare resources are running scarce, efficient use of the available resources is of eminent importance. With lots of routinely collected data available in primary care, opportunities open up to analyze and use these data in order to make efficient use of healthcare resources. With data-driven approaches, inequities in healthcare needs and supply can be made visible within the population. Risk stratification, the systematic assessment of patients' profiles in order to assign individual risk scores, can be used to provide personalized and population care according to that risk. In different countries, risk stratification is used to identify the right subpopulations for specific care interventions and lower expensive care utilization such as emergency care and hospitalization.

In the Netherlands, risk stratification is not yet notably used in primary care. Despite the great possibilities due to the widespread catchment area of primary care in the Netherlands and the gatekeepers function of general practitioners, systematic risk stratification approaches in Dutch primary care are minimal. The aim of this dissertation was therefor to identify and asses a suitable risk stratification tool to be used in Dutch primary care. After the introduction of the concept of risk stratification and the research question to be answered with this dissertation, the performed studies aiming at answering this question, are described in the chapters 2 till 5, with each chapter answering a sub-question derived from the central research question.

In chapter 2 a systematic literature review was described answering the sub-question *'What risk stratification tool is most appropriate for use in primary care?'* Studies were screened and systematically reviewed to identify risk stratification tools most suitable for primary care. Risk stratification tools were assessed on statistical validity. 61 articles were systematically reviewed and assessed, resulting in the identification of three mainly used risk stratification tools in primary care: 1) the Adjusted Clinical Groups (ACG), 2) the Charlson Comorbidity Index (CCI) and 3) the Hierarchical Condition Categories (HCC). Of these three identified risk stratification tools, the ACG was most frequently used in primary care and had the best statistical validity.

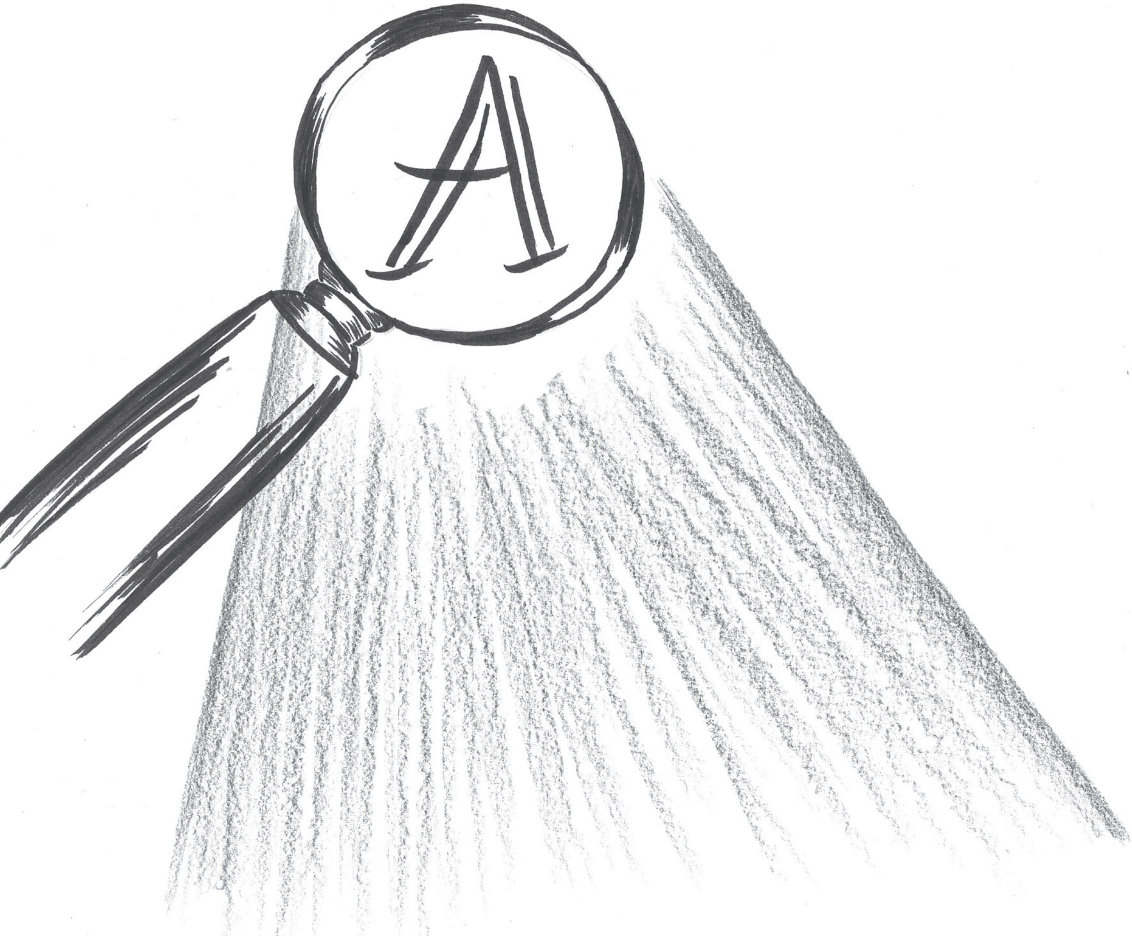
Chapter 3 described a pilot study in which the ACG tool was first applied in a Dutch primary care setting. The results showed the potential of the ACG in identifying different levels of morbidity and care burden using routinely collected Dutch general practitioners' data. Prediction models with ACG's diagnosis and multimorbidity categories as predictors have proven to be accurate for the prediction of GP visits per year.

The main question of this thesis was answered with chapter 4, describing a study assessing the model performance of ACG's hospitalization and high costs models applied in the Dutch primary care setting. Both models showed good performance regarding discrimination and calibration properties. In addition, the models were adjusted to best fit the Dutch primary care situation. Coefficients of the underlying predictors of the models were adjusted, improving the good model performances for both the hospitalization and the high costs model.

In chapter 5 an illustration is given of how risk stratification can be used to identify a specific subpopulation. With the use of the ACG, a group of complex patients with problems on multiple health domains (somatic chronic, psychological or social domain) could successfully be identified. Being able to identify such subpopulations can result in better care management and coordination of the right patients.

In the discussion of this dissertation, recommendations were done to methodologically improve the risk stratification models assessed in the different studies of this thesis. Suggested methodological improvements next to the statistical validation that has been performed in this thesis, included clinical validation of the underlying predictors of the models, taking the Dutch contexts into consideration. Social determinants are a valuable addition to biomedical determinants when it comes to the health of the population and the potential of adding them to risk stratification approaches in Dutch primary care should surely be investigated.

In addition to the methodological improvements of the models, first practical steps to be taken were discussed, including creating awareness of the benefits of risk stratification and trust in the tools, amongst health care professionals in and outside primary care as well as other stakeholders such as insurance companies, policy makers and municipalities. Lastly, although I believe that primary care is the best place in the care continuum to start off with risk stratification approaches, it is discussed that collaboration with especially the social domain would be beneficial to the improvement of the healthcare of the population. Social determinants are a valuable addition to biomedical determinants when it comes to the health of the population and the potential of adding them to risk stratification approaches in Dutch primary care should surely be investigated.



APPENDICES

APPENDICES

SAMENVATTING
LIST OF PUBLICATIONS
CURRICULUM VITAE
DANKWOORD

Samenvatting

Door het steeds schaarser worden van de middelen in de gezondheidszorg, wordt de noodzaak efficiënt om te gaan met de beschikbare middelen hoger. Met de beschikbare routinematig verzamelde data in de eerstelijns gezondheidszorg worden mogelijkheden gecreëerd om de schaarse middelen efficiënt in te zetten. Middels datagerichte benaderingen worden onevenredige gezondheidsbehoeften binnen populaties zichtbaar gemaakt. Risicostratificatie, de systematische analyse van patiënt profielen om zo risicoscores toe te kennen aan individuen, biedt de mogelijkheid persoons- en populatiegerichte zorg te verlenen. Wereldwijd wordt risicostratificatie gebruikt voor het identificeren van de juiste subpopulaties voor specifieke zorginterventies om op die manier het gebruik van dure zorg zoals spoedzorg en hospitalisatie te verminderen.

In Nederland wordt risicostratificatie nog niet op grote schaal toegepast in de eerstelijns gezondheidszorg. Hoewel de mogelijkheden groot zijn door de poortwachtersfunctie van Nederlandse huisartsen en het feit dat het grootste deel van de populatie geregistreerd staat bij een huisarts, wordt er minimaal gebruik gemaakt van datagerichte risicostratificatie. Het doel van dit proefschrift is dan ook het identificeren en beoordelen van een risicostratificatie model, dat het best gebruikt kan worden in de Nederlandse eerste lijn. Na de introductie van zowel het concept risicostratificatie als de onderzoeksvraag die middels dit proefschrift wordt beantwoord, worden de verschillende studies waarmee deze onderzoeksvraag wordt beantwoord, beschreven in de hoofdstukken 2 tot en met 5. Ieder hoofdstuk beantwoordt een van de sub-vragen, die zijn afgeleid van de centrale onderzoeksvraag.

Hoofdstuk 2 beschrijft een systematische literatuur review, waarmee de sub-vraag *'Welke risicostratificatie modellen zijn het best geschikt voor gebruik in de eerstelijns gezondheidszorg?'* wordt beantwoord. Studies zijn op systematische wijze gescreend en beoordeeld ter identificatie van risicostratificatie modellen geschikt voor de eerstelijns gezondheidszorg. De statistische validiteit van de modellen is beoordeeld. 61 artikelen zijn beoordeeld met als resultaat de identificatie van drie meest gebruikte risicostratificatie modellen in de eerstelijns gezondheidszorg: 1) de Adjusted Clinical Groups (ACG), 2) de Charlson Comorbidity Index (CCI) en 3) de Hierarchical Condition Categories (HCC). De ACG bleek het meest gebruikte risicostratificatie model in de eerstelijns gezondheidszorg, met de beste statistische validiteit.

Hoofdstuk 3 beschrijft een pilot studie, waarin een eerste toepassing van de ACG in de Nederlandse huisartsenpraktijk wordt gedaan. De resultaten laten de potentie van de ACG zien in het identificeren van verschillende niveaus van morbiditeit en zorgzwaarte, waarbij gebruik gemaakt wordt van routinematig verzamelde data uit een Nederlandse huisartsenpraktijk. Predictiemodellen met ACG's, diagnose- en multimorbiditeitscategorieën, als voorspellers bleken accuraat voor de predictie van de jaarlijkse hoeveelheid huisartsbezoeken.

De centrale onderzoeksvraag van dit proefschrift wordt beantwoord in hoofdstuk 4, waarin twee predictiemodellen van de ACG, predictie van hospitalisatie en van hoge zorgkosten, worden beoordeeld en aangepast aan de Nederlandse situatie. De coëfficiënten van de onderliggende predictiemodellen zijn hierbij afgesteld op de data uit Nederlandse huisartsenpraktijken, waarmee de statische validiteit van beide modellen verbeterd is.

In hoofdstuk 5 wordt een voorbeeld gegeven van hoe risicostratificatie gebruikt kan worden voor het identificeren van een specifieke subpopulatie. Met behulp van de ACG wordt een groep complexe patiënten met problemen op tenminste twee verschillende gezondheidsdomeinen (het somatisch chronisch, mentaal en sociaal gezondheidsdomein) succesvol geïdentificeerd. De mogelijkheid om dergelijke groepen subpopulaties te identificeren, kan resulteren in betere zorgmanagement en zorgcoördinatie van de juiste groepen patiënten.

In de discussie van dit proefschrift worden onder andere aanbevelingen gedaan om de risicostratificatie modellen methodologisch te verbeteren. Een methodologische verbetering betreft het klinisch valideren van de voorspellers van de predictiemodellen in navolging op de statistische validering, die middels dit proefschrift is gedaan. Sociale determinanten zijn een waardevolle toevoeging aan de biomedische determinanten voor het in kaart brengen van de gezondheid van de populatie en de potentie van het toevoegen van deze sociale determinanten aan risicostratificatie benaderingen in de Nederlandse eerste lijn is zeker verder onderzoek waard. In aanvulling op de methodologische verbeteringen van de risicostratificatie modellen wordt het belang van eerste praktische stappen zoals het creëren van bewustwording en draagvlak onder zowel eerstelijns zorgverleners als andere partijen zoals zorgverzekeraars, gemeenten en beleidsmakers, benadrukt. Ten slotte wordt, ondanks mijn overtuiging dat risicostratificatie in de eerstelijns gezondheidszorg het best op zijn plek is, bediscussieerd dat samenwerking met het sociale domein zeer bevorderlijk kan zijn voor de verbetering van de gezondheid van de populatie.

List of Publications

2022 ***Identifying complex patients using adjusted clinical groups risk stratification tool***
Girwar, S. M., Verloop, J. C., Fiocco, M., Sutch, S. P., Numans, M. E., & Bruijnzeels, M. A.
American Journal of Managed Care. April 2022
doi: 10.37765/ajmc.2022.88867

2021 ***Assessment of the Adjusted Clinical Groups system in Dutch primary care using electronic health records: a retrospective cross-sectional study***
Girwar, S. M., Fiocco, M., Sutch, S. P., Numans, M. E., & Bruijnzeels, M. A.
BMC Health Services Research. March 2021
doi: 10.1186/s12913-021-06222-9

2021 ***A systematic review of risk stratification tools internationally used in primary care settings***
Girwar, S. M., Jabroer, R., Fiocco, M., Sutch, S. P., Numans, M. E., & Bruijnzeels, M. A.
Health Science Reports. July 2021
doi: 10.1002/hsr2.329

Curriculum Vitae

Shelley-Ann Girwar was born on the 18th of December 1987 in Paramaribo, Suriname. In 2005, she finished her high school education and obtained her VWO diploma at the Ewald P. Meyer Lyceum in Paramaribo. In 2005 she moved to Amsterdam, the Netherlands, for her bachelor's in Biomedical Sciences at the Vrije Universiteit. As her interest shifted towards public health, she continued her path with a master's program in Health Sciences, again at the Vrije Universiteit. During her master's, she did her research internship on maternal health and antenatal care in rural South Africa, in cooperation with Busfare Babies Birth Centre and local NGO the Keiskamma Trust.

After finishing her studies, Shelley-Ann started to work as a junior researcher at the Jan van Es Institute, an independent expertise center, connecting research and practice. During her employment at the Jan van Es Institute, she worked on a project orienting on risk stratification possibilities with Dutch primary care data, which resulted in her PhD project. In April 2017, she started as an external PhD candidate at the Leiden University Medical Center (LUMC) - Campus the Hague, while still being employed by the Jan van Es Institute. With her PhD research, which was supervised by Marc Bruijnzeels and Mattijs Numans, she validated an internationally used risk stratification model for the Dutch primary care setting.

During her PhD project, she supervised various students during their scientific internships and gave lectures and workgroups to students and professionals.

From 2021 on, she coordinated the Population Health Management (PHM) Summer School, which is organized on a yearly base, as well as the *Health Systems and Management* course at Leiden University College (LUC). She was also involved in the preparations for the new LUMC Master's program Population Health Management, as coordinator of the project case, a real-life case running through the first semester, in which students need to apply the knowledge and insights gained during their courses. When the master's program officially started in 2022, her role within the program expanded with the coordination of the course *Fundamentals in Population Health Management* and the tutoring of students regarding their personal and professional development.

Shelley-Ann currently still works as tutor, teacher and course coordinator within the PHM Master's program and at LUC.

Dankwoord

Het is eindelijk zo ver! Met het behalen van deze mijlpaal, blik ik vergenoegend terug op dit bijzonder leerrijke traject en degenen, die een belangrijke rol hebben vervuld in de succesvolle afronding ervan.

Allereerst zou ik niet zijn waar ik ben, zonder mijn Heer. Mijn geloof in de Allerhoogste God geeft mij kracht en biedt mij de houvast die ik nodig heb. Ook de afgelopen jaren kon ik op stressvolle en uitdagende momenten gedurende dit traject terugvallen op mijn geloof, tot rust komen en vol vertrouwen er weer tegenaan.

Mijn enorme dank gaat uit naar Marc, mijn grote voorbeeld, mentor en inspirator, aan wie ik een groot deel van mijn professionele ontwikkeling te danken heb. Jij gaf mij kansen, motiveerde mij om een promotie onderzoek te doen en stimuleerde mij meer op mijn eigen kunnen te vertrouwen.

Marta, naast jouw statistische inbreng, was jij zorgzaam en waar nodig mijn 'stok achter de deur'. Toen ik toch eindelijk moest gaan beginnen aan mijn eerste artikel, gaf jij net die extra push en wanneer jij het idee had dat het mij allemaal iets te veel werd, stond je klaar met je zorgzaamheid.

Steve, thank you for your nice collaboration during the project, in which no call, message or meeting was too much for you. Thank you for your support with my analyses, your accessibility and your kindheartedness. In addition, your language checks of all my papers, were precious.

Mattijs, dank voor jou vertrouwen in mij, ook al liet de wetenschappelijke output even op zich wachten. Dank voor de mogelijkheden die jij binnen de Health Campus voor mij hebt gecreëerd.

Alle campus - (ex)collega's, dank voor de fijne werkomgeving. Ik kijk met plezier terug naar die geweldige allereerste campus atmosfeer die er heerste op de 5e. Aisha, jouw ontvangst en goede zorgen maakten de campus echt tot een thuis. Jouw prachtige persoonlijkheid, positieve energie en dankbare instelling, zijn inspirerend. Marlieke, wat waren wij een goed team. Samen zwommen we in die diepe zee, maar we wisten elkaar drijvende te houden, met onze brainstorm sessies, koffiemomentjes en dansjes. Laura, dank voor jouw steeds terugkerende behulpzaamheid: Je wist altijd wel de juiste weg richting benodigde documenten of relevante informatie aan

te wijzen. Naomi, niet alleen maakte jij het leven op de campus levendig, maar ook fysiek verantwoord met jouw 1-minuut-per-uur-planking. Thom, jouw droge humor voegde de nodige gezelligheids-ingrediënten toe aan het werken op de campus. Robert, met jouw gestructureerde werkwijze en geweldig schrijfniveau, maakte je ons het gezamenlijk schrijven van een systematische review makkelijk. Het was je ook niet teveel om, naast je nieuwe baan nog mee te helpen met het screenen van artikelen. Jorieke, ik heb jou, ondanks het vele thuiswerken wegens corona, fijn mogen begeleiden voor je wetenschapsstage en jij hebt veel analysewerk verzet voor het artikel over de identificatie van complexe patiënten.

Dank aan de zorggroepen en organisaties die onderdeel waren van het project; Sam, Harry, Robin, Philip, Hans Erik, Priscilla, Laurens, Nienke, Frens, Els, Sanne, Jacob en alle betrokken huisartsen, dank voor het beschikbaar stellen van data voor werkstroom 1 en jullie inzet in de praktische uitwerking binnen werkstroom 2. Thomas, Alex, Peter, Johnny, Carla B en Maartje, dank voor jullie hulp bij de privacy borging betreffende de data. Carla van der B, dank voor het meedenken over de juiste praktische verwerking van de data.

Dank aan mijn familie en vrienden, teveel om op te noemen; jullie gaven precies de juiste balans in mijn leven om dit traject aan te kunnen.

Ma, dank voor alle offers die je hebt gebracht om mij mooie kansen in het leven te bieden, waaronder de mogelijkheid om in Nederland te komen studeren. Dank voor alle financiële ondersteuning tijdens mijn studies, maar ook je mentale en geestelijke ondersteuning tijdens dit traject. En bovenal, dank voor al je gebeden.

Dank aan mijn schoonmoeder voor alle lekkere maaltijden en de goede zorgen voor Micah.

Shellice, mijn *best friend*, dank voor alle gesprekken, over van alles en nog wat, maar vooral ook over onze beide promotietrajecten. De fysieke afstand tussen ons kan niet voorkomen dat wij elkaar ook hierin tot steun zijn, hetzij middels het doorspreken van mogelijke analysemethoden, het doorlezen van elkaars stukken, motivatiegesprekken of gewoon het bieden van een luisterend oor.

Kim en Marlieke, de paranimfen aan mijn zij, dank voor jullie geweldige hulp in het regelen van van alles rondom deze dag. Mede door jullie hulp, kan ik ten volle genieten van deze bijzondere dag.

En *last but definitely not least*: Jurgen, mijn steun en toeverlaat! Jij geeft mij altijd de ruimte te doen wat belangrijk is voor mij. Ook tijdens dit traject gaf je mij steeds weer die nodige ruimte. En als mijn analyse-of schrijfavonden weer eens tot in de diepe nachtelijke uurtjes overliepen, dan zorgde jij ervoor dat de huishouding gedaan werd en dat het mij aan niets ontbrak. Ik ben dankbaar voor jou aan mijn zij.


10101
00010