



Universiteit
Leiden
The Netherlands

Clinical proteomics in oncology : a passionate dance between science and clinic

Noo, M.E. de

Citation

Noo, M. E. de. (2007, October 9). *Clinical proteomics in oncology : a passionate dance between science and clinic*. Retrieved from <https://hdl.handle.net/1887/12371>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/12371>

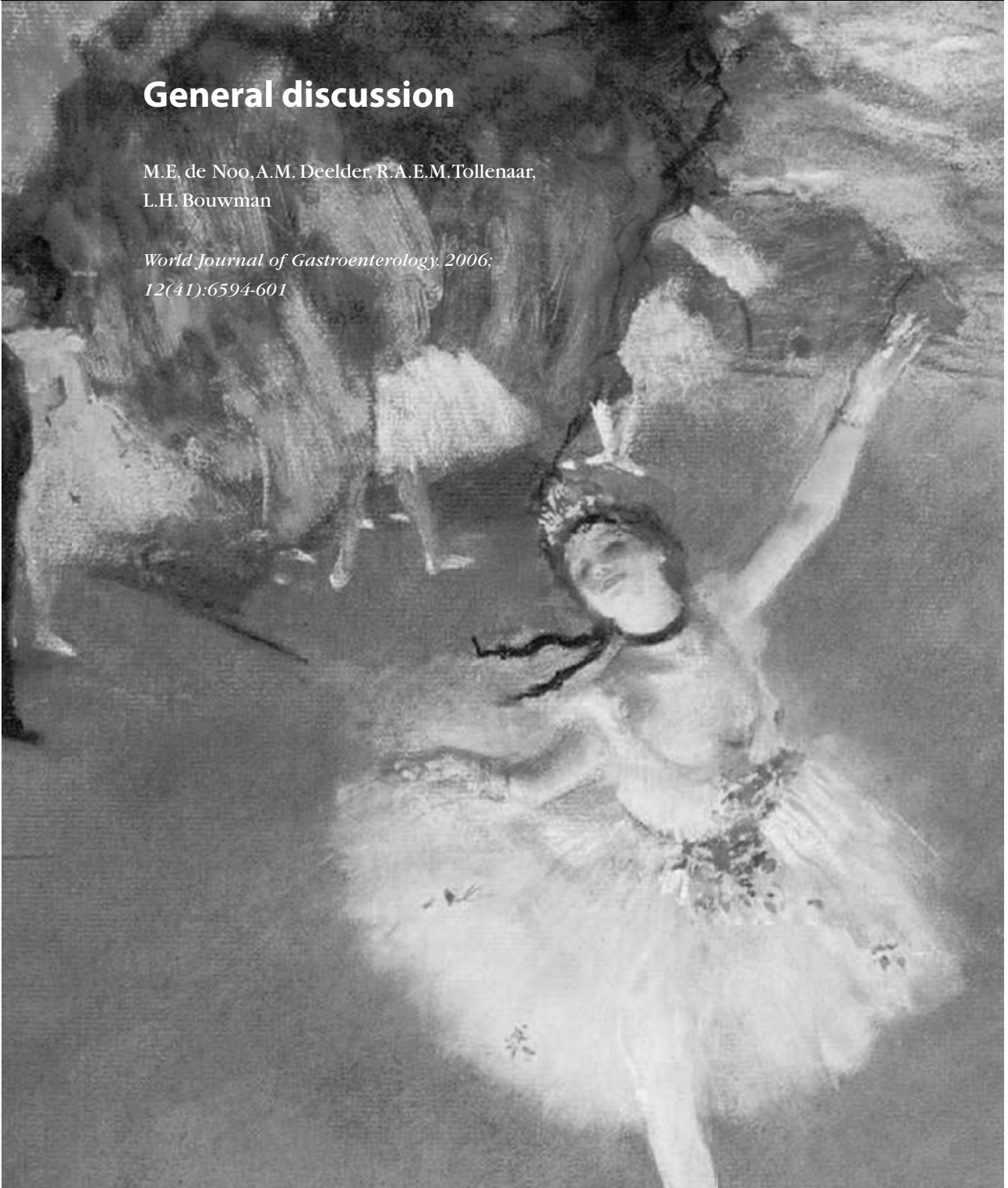
Note: To cite this publication please use the final published version (if applicable).

Chapter 8

General discussion

M.E. de Noo, A.M. Deelder, R.A.E.M. Tollenaar,
L.H. Bouwman

World Journal of Gastroenterology. 2006;
12(41):6594-601



BACKGROUND

There is an urgent need for new biomarkers in oncology to improve early detection, monitor disease outcome and find targets for more individualised therapy. A field of recent interest is clinical proteomics, which was reported to lead to high sensitivity and specificities for early detection of several solid tumours.[1;2] This emerging field uses mass spectrometry based protein profiles/patterns of easy accessible body fluids to distinguish cancer from none-cancer patients. This would be a solution to the problem that cancer is often diagnosed in late stages, when curative resection of the diseased organ is not possible anymore and the disease has already metastasised, dropping survival rates dramatically. However, after the initial hype in early 2002 critical noise has been heard on several aspects of serum proteomics. In this paper we describe the hopes and fears for the introduction of clinical proteomics for (early) detection of CRC.

COLORECTAL CANCER

Colorectal adenocarcinoma is the third most common cancer and the fourth most frequent cause of death due to cancer worldwide. Worldwide almost one million new cases occur yearly, with 492,000 related deaths.[3] In developed countries it is the second most common tumour, with a lifetime risk of 5%, but its incidence and mortality are now decreasing.[4;5] Surgery is the cornerstone of therapy when the disease is confined to the bowel wall. This results in 70 to 80% of patients who have tumours that, at diagnosis, can be resected with curative intent.[6] After curative surgery the five-year survival rate for patients with localised disease is 90%, decreasing to 65% in case of metastasised disease in the lymph nodes. Adjuvant radiation therapy, chemotherapy, or both are useful in selected patients. Classification of tumours into pathogenetical subtypes with distinct clinical courses enables clinicians to target therapy. For CRC TNM staging system remains the golden standard and relies entirely on morphological appearance of the tumour. However, tumours with similar histopathological characteristics may have different clinical outcome and responsiveness to therapy.[7] Therefore, more individualised treatment would benefit the individual patient and avoid unnecessary morbidity. Nonetheless, early detection of CRC will increase survival most, in view of the fact that it is well recognised that CRC arises from a multistep sequence of genetic alterations that result in the transformation of normal mucosa to a precursor adenoma and ultimately to carcinoma. Given the natural history of CRC, early diagnosis appears to be the most appropriate tool to reduce disease-related mortality.[8-10]

BIOMARKERS

In cancer research biomarkers are molecules that indicate the presence of cancer in the body. Most biomarkers are based on abnormal changes or mutation in genes, RNA, proteins and metabolites. Since the molecular changes that occur during tumour development can take place over a number of years, some biomarkers can potentially be used to detect colorectal cancer early. Furthermore, they might be used to predict prognosis, monitor disease progression and therapeutic response. Gion et al. classified different circulating biomarkers according to their clinical application.[11] These candidate biomarkers however, are frequently found in relatively low concentrations amid a sea of other biomolecules, so biomarker research and possible diagnostic tests depend critically on the ability to make high sensitive and accurate biochemical measurements. Ideally, such biomarkers should be specific to the disease and easy accessible, such as serum, plasma or urine, increasing their clinical applicability.

Carcinoembryonic antigen (CEA) is the best-characterised serologic tumour marker for CRC. However, its use as a population based screening tool for early detection and diagnosis of CRC is hindered by its low sensitivity and specificity. Fletcher showed that for screening purposes in a normal population, a cut-off concentration of 2.5 µg/L CEA would yield a sensitivity of 30-40%. Based on these data he calculated that there would be 250 false positive tests for every true positive test, i.e. a patient with cancer. Furthermore, 60% of the cancers would not be detected. The same poor sensitivity applies for diagnosis of CRC. In addition, as CEA can be elevated in the absence of malignancy, specificity is also impaired.[12-15]

Faecal occult-blood testing (FOBT) is another biomarker for which clinical trials have shown evidence of a decreased risk of death. This approach is a non-invasive option that limits the need for follow-up colonoscopy to patients with evidence of bleeding. Neoplasms bleed intermittently, however, allowing many to escape detection with faecal occult-blood testing. Annual retesting is therefore necessary but is still insufficient, detecting only 25 to 50% of colorectal cancers and 10% of adenomas. The specificity of FOBT is also limited by frequent false positive reactions to dietary compounds, medications, and gastrointestinal bleeding from causes other than colorectal cancer.[16-18]

A NEW DIAGNOSTIC PARADIGM: CLINICAL PROTEOMICS

In 2002 several studies discriminated patients with various cancers from healthy subjects on the basis of presence/absence of multiple low-molecular-weight serum

proteins using SELDI-TOF mass spectrometry technologies.[19-22] The authors hypothesised that proteomic patterns are correlated to biological events occurring in the entire organism and are likely to change in the presence of disease. New types of bioinformatic pattern recognition algorithms were used to identify patterns of protein changes in order to discriminate cancer patients from healthy individuals with promising results.

Petricoin and his co-workers stated that finding a single disease-related biomarker is like searching for a needle in a haystack; each entity has to be separated and identified individually.[23;24] Moreover, they postulated that the blood proteome constantly changes as a consequence of the perfusion of the diseased organ adding, subtracting, or modifying the circulating proteome. These differences might be the result of proteins being abnormally produced or shed and added to the serum proteome, clipped or modified as a consequence of the disease process, or subtracted from the proteome owing to disease-related proteolytic degradation pathways. Therefore, protein pattern diagnostics would provide easier and more reliable tools for detection of cancer. The advantages of the SELDI proteomic pattern approach were stressed in several papers. In addition to the high sensitivity and specificity, cost-effectiveness, easy accessibility of body fluid and especially the high-throughput, ultimately allowing application in future screening studies, were mentioned.[20;25] Next to these hopeful voices, soon critical notes were made on analytical reproducibility and the use of the so-called black box approach, lacking identification of discriminating proteins.

In the next paragraphs this paper will focus on the current status of clinical proteomics research in oncology and will reflect on pitfalls and fears in this relatively new area in clinical medicine: reproducibility issues and pre-analytical factors; statistical issues; and identification and nature of discriminating proteins/peptides.

REPRODUCIBILITY ISSUES AND PRE-ANALYTICAL FACTORS

Boguski and McIntosh were among the first to argue that serum proteomics may be susceptible to observational biases. They stated that any confounding factor could conceivably cause a phenotypic response that might be confused with a specific characteristic of the disease process under study.[26] Confounding factors such as smoking, diet and preoperative stress, but also sample collection and quality, trouble a reliable and clear differentiation of a normal or malignant status. Another cause for concern mentioned in this study, is the sample quality and number. The authors favoured use of homogeneous groups with sufficient sample size and stringent standard procedures for serum collection, an aspect which is also advocated in other

studies.[27;28] Another critical study questioned the reliability of the presence of statistically significant signals at M/Z values less than 500, as used in one of the first studies. Sorace et al. claimed that the presence of statistically significant bands of low M/Z includes degradation products of higher molecular weight macromolecules or a matrix effect. Furthermore, this study cautioned for poor reproducibility of experimental conditions of chip based mass spectrometry.[29] This is also reported by another group, which showed the poor reproducibility of the SELDI-TOF ovarian cancer data. Baggerly and colleagues postulated that this could partly be contributed to baseline correction, poor sample features in noise regions and even a change of protocol mid-experiment.[30] Most importantly, the promising results that were reported earlier could not be reproduced and therefore stressed the importance of standardised approaches, stringent experimental design. Furthermore, their study pointed out that strong pre-processing of the protein spectra is required in order to obtain reliable classification results in the search for new biomarkers.

Possible confounding factors can be categorised into three sources of variation and bias: biological variation, pre-analytical variation and analytical reproducibility. Biological variation, consist of both environmental and individual factors, such as race, age, diet, smoking, stress, general physical condition, and use of drugs, and may also influence serum protein profiles. However, at the present no data have been published on this source of variation. Nevertheless, in a previous study our group analysed pre-analytical and reproducibility issues of our MALDI-TOF approach.[31] The pre-analytical variations corresponded to the logistical conditions in the routine clinical setting; the effects of sample handling and storage. So far, only few other studies have reported on the effects of different serum sample preparations and the use of a magnetic-beads-based approach to capture and concentrate serum proteins for MALDI-TOF mass spectrometry.[32-34] Where Villanueva et al. mostly focused on influences of different magnetic beads capturing and its automation on the reproducibility of serum protein profiles, Baumann and co-workers mainly studied pre-analytical variation of sample handling. In table 1, different results of sample handling experiments of the above mentioned studies are summarised. For clinical studies the use of two freeze/thaw cycles is recommended by 3 out of 4 manuscripts. This is mainly due to logistical reasons, such as the 'standard' for centralised sample collection in large hospitals. The point all authors agreed on is the influence of sample handling, i.e. the time venous blood is left to stand before serum centrifugation. This aspect appears to account for the largest effect on serum or plasma protein profiles. Consequently, standardised sample collection and a well documented population are recommended in all performed studies. Standardised protocols should be used from the point of sample collection, sample handling, storage and freezing of the samples. Although the importance of homogeneity and uniformity within sample groups must

Table 1. Recommendations of various pre-analytical variations from three MALDI-TOF based reproducibility studies.

	Blood component	Peptide isolation	Temp before sample handling	Time before centrifugation	Storage of serum	Freeze/thaw cycles	Circadian rhythm effect
Baumann et al.	Serum/Plasma	C3, C8, C18 beads	21° C	< 30 min	-80° C	1	N.A.
de Noo et al.	Serum	C8 beads	21° C	Ideally < 30 min, practically < 2-4 hrs	N.A.	2	No effect
West-Nielsen et al.	Serum/Plasma	C8 beads	21° C	< 8 hrs	-20/ -80° C	1	N.A.

once again be stressed, variation of such factors can not totally be excluded in a clinical setting. In all, when these recommendations are strictly followed and both clinical and analytical factors are controlled, we think that the methodology can be standardised to a level which allows application as a tool in biomarker discovery.

STATISTICAL ISSUES

As in all research with high dimensional data, two practical realities constrain the analysis of mass spectra in proteomics. The first is the ‘curse of dimensionality’: the number of features characterizing these data is in the thousands or tens of thousands. The second is the ‘curse of dataset sparsity’: the number of samples is limited. Somorjai et al. showed the influences of these two curses on classification outcomes. Both the sample per feature ratio, which should be 5 to 10 ideally, and feature selection are pivotal importance for reliable classification and biological optimal relevance.[35;36]

Previous to any feature selection or classification, raw mass spectra have to be submitted to so-called pre-processing. During this process noise of protein/peptide mass spectra is reduced and spectra are normalised. Furthermore, smoothing, binning and baseline correction are also performed during pre-processing of the data. Currently, there is a lot of discussion between several groups on how to establish the best method, because data pre-processing is extremely important. There are complex interactions between baseline subtraction, normalization, noise estimation, and peak identification, and therefore these steps should not be considered in isolation.[31;37-40]

Another recurrent topic for debate is the bioinformatic approach and statistical analysis of protein spectra. Clinically most relevant is the issue of an independent validation set for the classification of diseased versus healthy individuals. This is pri-

marily based on a specific problem in the discovery-based research field of clinical proteomics, namely overfitting. Overfitting may occur in the analysis of large datasets when multivariate models show apparent discrimination that is actually caused by data over-interpretation, and hence give rise to results that are not reproducible. [30;41;42] The chance of overfitting, however, can be reduced by appropriate application of validity estimation and assessment, such as through application of double cross-validation, when properly implemented.[43] Although we have shown this in a previous study, the general opinion is in favor performing a classification study with independent validation. Furthermore, feature selection is also given a lot of attention by statisticians in the field. Several experimental investigations have been made with different peak feature selection methods. A common approach thus far is analysing the data in two phases. First, the peaks in the spectra are extracted and quantified. Secondly, a resulting matrix of peak quantifications is created. For more detailed information on this statistical matter, we refer to the literature.[37;44-46]

IDENTIFICATION AND NATURE OF DISCRIMINATING PROTEINS

The controversy about the use of protein profiles as a pattern diagnostic without identification of the individual diagnostic biomarkers remains to be solved before its clinical application. Whereas the first clinical proteomics studies published their classification method mainly as a black box study, nowadays identification of the most discriminating proteins or peptides is required for publication in most scientific journals. Identification and functional analysis of these discriminating proteins/peptides might render new insights on tumour development and environmental responsiveness, which could eventually be translated into new diagnostic and prognostic insights for the clinician. Unfortunately, little success has been booked so far in assigning reproducible discriminating biomarkers.[35;42]

Furthermore, several studies have identified their discriminating peaks as components of the coagulation cascade or complement system.[47-51] So, in contrast to the original reflection that discriminating proteomic patterns would identify cancer-specific proteins, it appears that these potential markers belong to the normal serum and plasma proteome. Consequently, some investigators have argued that low molecular weight proteins in serum, the serum peptidome, is nothing but aspecific biological trash and therefore does not yield any reliable biomarkers in the currently technically available mass range.[29;52] Others have proposed that the discriminatory protein peaks represent acute phase reactants that are present in serum in extremely high concentrations.[49;53] Conversely, recently a study reported that although discriminating peptides do indeed belong to the well known coagulation and complement

pathways, their patterns or signatures can nevertheless indicate the presence of cancer. Villanueva et al. showed that most of the cancer-type specific biomarker fragments were generated in patient serum by enzymatic cleavage at previously known endoprotease cleavage sites after the blood sample was collected.[54;55] They postulated that the discriminating peptides originated after *ex vivo* proteolysis by tumour specific proteases of high abundance protein fragments primarily generated by the coagulation and complement enzymatic cascades. In this view, they consider these cancer-specific low molecular weight proteins in the serum peptidome an indirect snapshot of the enzyme activity in tumour cells. We support their hypothesis that proteolytic process profiles in the serum peptidome hold important information that may have direct clinical utility as a surrogate marker for the detection and classification of certain types of tumours. Unique proteases may be shed by tumour cells or reflect activity of the host immune response, which may contribute to new proteins such as chemokines and lymphokines. These processes result in subtle changes in low molecular proteomic signatures, which may ultimately be used for classification methods in various cancers and disease in the future.[54] Proteases have been extensively implicated in the development and progression of cancer.[56;57] Song et al. recently stated that proteolytic processing of high abundance host-response proteins actually amplifies the signal of potentially low-abundance biologically active disease markers such as proteases. Therefore, it might be expected that more convenient and reliable blood proteins and peptides simply serve as an endogenous substrate pool for proteases as surrogate markers for the detection and classification of cancer.[58]

Another recurrent topic of debate is which blood component is best used for protein profiling and peptidome analysis. Some investigators favour the use of plasma because they presume that, in serum, ongoing enzymatic activity, occurring during clotting, is likely to cleave even proteins that are not involved in biological relevant pathways.[53;59] Others, however, advocate the use of serum. We support the hypothesis that since the kidneys rapidly clear peptides smaller than 4 kDa which are *in vivo* generated in the circulation, the majority of peptides in blood samples exist from *ex vivo* proteolysis. This explains that low abundance proteins, including possible tumour markers, may be totally obscured and not retraceable during direct mass spectrometry. However, it has recently been shown that exogenous proteases are functionally measurable in serum, yet in higher concentrations than in plasma. [54]

Functional proteomics studies allow the investigation of environmental factors over time, rendering the monitoring of metabolic responses to various stimuli. Hence, post translational modifications can be studied, whereas they can not be detected by genomic studies. Posttranslational modifications changes like glycosylation of proteins and lipids are a common feature in colorectal cancer and influence cancer

cell behaviour and can be detected using mass spectrometry due to characteristic mass shifts.[60] We expect that both phosphoproteomics and/or glycoproteomics, enabling study of crucial post translational modifications of proteins in the cancer pathway, will revolutionize our understanding of the function of these proteins, and hence render new insights for monitoring and therapy.

CLINICAL PROTEOMICS IN CRC

So far, few protein profiling studies have been published on the detection of CRC, of two were based on SELDI/TOF and one on MALDI-TOF mass spectrometry. The first SELDI/TOF study showed seven potential biomarkers that could differentiate CRC patients from patients with colorectal adenoma with a sensitivity of 89% and specificity of 83%. The seven potential biomarkers have a large range in mass values, differing from 4654 till 21,742 Da.[61] A more recent published study found 5 possible biomarkers to differentiate between healthy control subjects and CRC patients. For three of these potential markers they found a sensitivity and specificity between 65% and 90%. They reported that m/z 3100, 3300, 4500, 6600 and 28,000 were the most important biomarkers.[62] Our group used MALDI-TOF mass spectrometry to differentiate CRC patients from healthy controls. In a randomised block design pre-operative serum samples obtained from 66 colorectal cancer patients and 50 controls were used to generate high-resolution MALDI-TOF protein profiles.[43] After pre-processing of the spectra, linear discriminant analysis with double cross-validation, based on principal component analysis was used to classify the protein profiles. A total recognition rate of 92.6%, a sensitivity of 95.2% and a specificity of 90.0% for the detection of CRC were shown. In our study two first principal components accounted for most of the between-group separation, both with a m/z between 1000 and 2000 Da.

Although a lot of research has been done using 2D gel electrophoresis to detect possible biomarkers and targets for CRC, this falls outside the scope of this paper since this technique can not be scaled up to a directly applicable diagnostic test. On the other hand, recently a screening assay based on APC protein truncation test has been proposed and other studies mention the potential use of protein microarrays. [2;63-65] However, studies linking large protein expression patterns with clinical outcome in colorectal cancer are still in their infancy. To be able to predict occurrence of disease, and treatment outcome, more studies on genotype-phenotype correlations are needed both in sporadic and in hereditary colorectal cancer.

FUTURE PERSPECTIVES

The best anticancer strategies still rely on early detection followed by close monitoring for early relapse so that therapies can be appropriately adjusted.[66] In addition, new targets for therapy are a constant subject of study in oncology. In fact, increased understanding of the molecular mechanisms of cancer progression may refine treatment and management of patients. Advances in genomics and proteomics may lead to earlier detection of cancer and may enable a more precise classification of (smaller subsets of) patients based on their predicted response to individual therapies. Conceptually, proteomics is more suitable than genomics for novel targeted therapies, since most of protein biomarkers are based on aberrant protein signalling circuits represented by post translational modifications. The dynamic range of the proteome allows more insight in the functional state of a cell, tissue or organ over time. Besides, protein profiling and classification of several components of multiple aberrant cell signalling cascades would be expected to predict disease behaviour better than just single pathways in isolation.[64] Therefore, proteomics could be expected to render better insight in pathogenetic mechanisms, disease progression and treatment response. This is of paramount importance as cancer advances dynamically and affects heterogeneous cell populations, either as a part of cancer or as a part of a tumour-host reaction.[49;67]

Further refinement of serum protein profiles is needed before these mass spectrometry based techniques become part of clinical routine. Nowadays, several studies have carefully evaluated reproducibility, automation, sample throughput and sensitivity of serum proteomic techniques. The first problems related to these factors seem to have been overcome due to stringent standardised approaches as described earlier. However, proteomics studies still have several drawbacks: 1) current tools only allow narrow-range analyses, 2) identification of proteins of interest remains cumbersome, 3) protein studies address mixtures of high complexity. Hence, due to the dynamic ranges of the human proteome and the lack of amplification methods in protein studies, targeted proteomics techniques for (quantitative) identification of low-abundant proteins have to be further investigated.[68] Another approach to study proteins at a functional level might be the use of array-based proteomics platforms. This technique offers the potential for highly multiplex and sensitive analysis of serum or tumour proteins.[64] Using this direct approach of studying the proteomic circuitry would theoretically allow for the creation of functional signalling maps of cancers, even at the level of the individual patient. Regarding identification of potential biomarkers, limitations of direct MS/MS have been stressed before as well as the fact that antibody-approaches may yield higher sensitivity.[53;54]

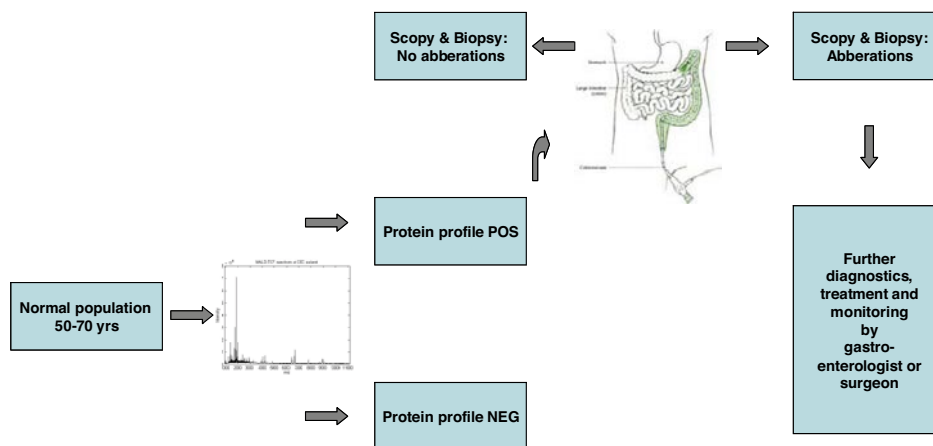


Figure 1. Flow chart of possible clinical application of MALDI-TOF

In the next era research in oncology will drift to more individualised medicine. In this view, molecular profiling forms a welcome addition to the pathology report of cancer. Until now, histopathological staging and demographics have been used to predict disease outcome. However, we believe that protein profiling and other proteomics techniques may lead to more individualised medicine and tailor made therapy.[69;70] At first, both approaches should be used complementary instead of competitively.

It is unlikely that in the next decade, serum protein profiles will certainly replace the current gold standard colonoscopy for the diagnosis of CRC. Nevertheless, we hypothesise that MALDI-TOF based serum protein profiles, once validated in independent studies, could be used as selection criteria for the more invasive and time consuming diagnostic colonoscopy (Figure 1). Eventually, with the present debate on screenings programs for colorectal cancer in several countries, clinical proteomics may replace and surpass the use of faecal occult-blood testing (FOBT). When in independent validation studies sensitivity and specificity remain about 90% protein profiling might even replace FOBT, since this approach has a lower specificity and a number of disadvantages. Non-bleeding tumours and more relevant, polyps and adenomas can not be detected using FOBT, whereas we expect to realise this with serum protein profiling within the next decade.[17;18]

So, although the current reality may not have kept pace with previous expectations and the translation from bench to bedside is more laborious than initially thought, there is supporting evidence for the potential great use of clinical proteomics in oncology. Particularly, when efforts for technical innovations to further increase sensitivity and specificity of proteomic techniques will be implemented and more sensi-

tive methods for protein identification on alternations are developed. In combination with the use and set-up of well-defined cases with well documented serum banks, including not only CRC samples, but also inflammatory disease and polyps, serum protein profiling may propel diagnostic research in CRC in the right direction.

REFERENCES

1. Chambers,G., Lawrie,L., Cash,P., and Murray,G.I. (2000) Proteomics: a new approach to the study of disease. *J.Pathol.*, 192, 280-288.
2. Posadas,E.M., Simpkins,F., Liotta,L.A., Macdonald,C., and Kohn,E.C. (2005) Proteomic analysis for the early detection and rational treatment of cancer--realistic hope? *Ann.Oncol.*, 16, 16-22.
3. Weitz,J., Koch,M., Debus,J., Hohler,T., Galle,P.R., and Buchler,M.W. (2005) Colorectal cancer. *Lancet*, 365, 153-165.
4. Russo,M.W., Wei,J.T., Thiny,M.T., Gangarosa,L.M., Brown,A., Ringel,Y., Shaheen,N.J., and Sandler,R.S. (2004) Digestive and liver diseases statistics, 2004. *Gastroenterology*, 126, 1448-1453.
5. Jemal,A., Tiwari,R.C., Murray,T., Ghafoor,A., Samuels,A., Ward,E., Feuer,E.J., and Thun,M.J. (2004) Cancer statistics, 2004. *CA Cancer J Clin*, 54, 8-29.
6. Pfister,D.G., Benson,A.B., III, and Somerfield,M.R. (2004) Clinical practice. Surveillance strategies after curative treatment of colorectal cancer. *N.Engl.J Med.*, 350, 2375-2382.
7. Liefers,G.J. and Tollenaar,R.A. (2002) Cancer genetics and their application to individualised medicine. *Eur.J.Cancer*, 38, 872-879.
8. Ruo,L., Gougoutas,C., Paty,P.B., Guillem,J.G., Cohen,A.M., and Wong,W.D. (2003) Elective bowel resection for incurable stage IV colorectal cancer: prognostic variables for asymptomatic patients. *J.Am.Coll.Surg.*, 196, 722-728.
9. Gill,S. and Sinicrope,F.A. (2005) Colorectal cancer prevention: is an ounce of prevention worth a pound of cure? *Semin.Oncol.*, 32, 24-34.
10. Hawk,E.T. and Levin,B. (2005) Colorectal cancer prevention. *J Clin Oncol*, 23, 378-391.
11. Gion,M. and Daidone,M.G. (2004) Circulating biomarkers from tumour bulk to tumour machinery: promises and pitfalls. *Eur.J Cancer*, 40, 2613-2622.
12. Duffy,M.J., van Dalen,A., Haglund,C., Hansson,L., Klapdor,R., Lamerz,R., Nilsson,O., Sturgeon,C., and Topolcan,O. (2003) Clinical utility of biochemical markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines. *Eur.J.Cancer*, 39, 718-727.
13. Fletcher,R.H. (2002) Rationale for combining different screening strategies. *Gastrointest.Endosc.Clin.N.Am.*, 12, 53-63.
14. Winawer,S., Fletcher,R., Rex,D., Bond,J., Burt,R., Ferrucci,J., Ganiats,T., Levin,T., Woolf,S., Johnson,D., Kirk,L., Litin,S., and Simmang,C. (2003) Colorectal cancer screening and surveillance: clinical guidelines and rationale-Update based on new evidence. *Gastroenterology*, 124, 544-560.
15. Ouyang,D.L., Chen,J.J., Getzenberg,R.H., and Schoen,R.E. (2005) Noninvasive testing for colorectal cancer: a review. *Am.J.Gastroenterol.*, 100, 1393-1403.
16. Pignone,M., Rich,M., Teutsch,S.M., Berg,A.O., and Lohr,K.N. (2002) Screening for colorectal cancer in adults at average risk: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann.Intern.Med.*, 137, 132-141.
17. Ransohoff,D.F. and Lang,C.A. (1997) Screening for colorectal cancer with the fecal occult blood test: a background paper. American College of Physicians. *Ann.Intern.Med.*, 126, 811-822.
18. Ransohoff,D.F. (2005) Colon cancer screening in 2005: status and challenges. *Gastroenterology*, 128, 1685-1695.
19. Adam,B.L., Qu,Y., Davis,J.W., Ward,M.D., Clements,M.A., Cazares,L.H., Semmes,O.J., Schellhammer,P.F., Yasui,Y., Feng,Z., and Wright,G.L., Jr. (2002) Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Res.*, 62, 3609-3614.

20. Petricoin,E.F., Ardekani,A.M., Hitt,B.A., Levine,P.J., Fusaro,V.A., Steinberg,S.M., Mills,G.B., Simone,C., Fishman,D.A., Kohn,E.C., and Liotta,L.A. (2002) Use of proteomic patterns in serum to identify ovarian cancer. *Lancet*, 359, 572-577.
21. Rai,A.J., Zhang,Z., Rosenzweig,J., Shih,I., Pham,T., Fung,E.T., Sokoll,L.J., and Chan,D.W. (2002) Proteomic approaches to tumor marker discovery. *Arch.Patbol.Lab Med.*, 126, 1518-1526.
22. Yanagisawa,K., Shyr,Y., Xu,B.J., Massion,P.P., Larsen,P.H., White,B.C., Roberts,J.R., Edgerton,M., Gonzalez,A., Nadaf,S., Moore,J.H., Caprioli,R.M., and Carbone,D.P. (2003) Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet*, 362, 433-439.
23. Petricoin,E.F. and Liotta,L.A. (2002) Proteomic analysis at the bedside: early detection of cancer. *Trends Biotechnol.*, 20, S30-S34.
24. Wulfkuhle,J.D., Liotta,L.A., and Petricoin,E.F. (2003) Proteomic applications for the early detection of cancer. *Nat.Rev.Cancer*, 3, 267-275.
25. Petricoin,E.E., Paweletz,C.P., and Liotta,L.A. (2002) Clinical applications of proteomics: proteomic pattern diagnostics. *J.Mammary.Gland.Biol.Neoplasia.*, 7, 433-440.
26. Boguski,M.S. and McIntosh,M.W. (2003) Biomedical informatics for proteomics. *Nature*, 422, 233-237.
27. Diamandis,E.P. (2003) Re: Serum proteomic patterns for detection of prostate cancer. *J.Natl. Cancer Inst.*, 95, 489-490.
28. Diamandis,E.P. (2002) Proteomic patterns in serum and identification of ovarian cancer. *Lancet*, 360, 170-171.
29. Sorace,J.M. and Zhan,M. (2003) A data review and re-assessment of ovarian cancer serum proteomic profiling. *BMC.Bioinformatics.*, 4, 24.
30. Baggerly,K.A., Morris,J.S., and Coombes,K.R. (2004) Reproducibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments. *Bioinformatics.*, 20, 777-785.
31. de Noo,M.E., Tollenaar,R.A.E.M., Ozalp,A., Kuppen,P.J.K., Bladergroen,M.R., and Deelder A.M. (2005) Reliability of human serum protein profiles generated with C8 magnetic beads assisted MALDI-TOF mass spectrometry. *Anal.Chem.*, 77, 7232-7241.
32. Villanueva,J., Philip,J., Entenberg,D., Chaparro,C.A., Tanwar,M.K., Holland,E.C., and Tempst,P. (2004) Serum Peptide profiling by magnetic particle-assisted, automated sample processing and maldi-tof mass spectrometry. *Anal.Chem.*, 76, 1560-1570.
33. Baumann,S., Ceglarek,U., Fiedler,G.M., Lembcke,J., Leichtle,A., and Thiery,J. (2005) Standardized Approach to Proteome Profiling of Human Serum Based on Magnetic Bead Separation and Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Clin. Chem.*, 51, 973-980.
34. West-Nielsen,M., Hogdall,E.V., Marchiori,E., Hogdall,C.K., Schou,C., and Heegaard,N.H. (2005) Sample handling for mass spectrometric proteomic investigations of human sera. *Anal. Chem.*, 77, 5114-5123.
35. Somorjai,R.L., Dolenko,B., and Baumgartner,R. (2003) Class prediction and discovery using gene microarray and proteomics mass spectroscopy data: curses, caveats, cautions. *Bioinformatics.*, 19, 1484-1491.
36. Morris,J.S., Coombes,K.R., Koomen,J., Baggerly,K.A., and Kobayashi,R. (2005) Feature extraction and quantification for mass spectrometry in biomedical applications using the mean spectrum. *Bioinformatics.*
37. Baggerly,K.A., Morris,J.S., Wang,J., Gold,D., Xiao,L.C., and Coombes,K.R. (2003) A comprehensive approach to the analysis of matrix-assisted laser desorption/ionization-time of flight proteomics spectra from serum samples. *Proteomics.*, 3, 1667-1672.

38. Yasui, Y., Pepe, M., Thompson, M.L., Adam, B.L., Wright, G.L., Jr., Qu, Y., Potter, J.D., Winget, M., Thornquist, M., and Feng, Z. (2003) A data-analytic strategy for protein biomarker discovery: profiling of high-dimensional proteomic data for cancer detection. *Biostatistics.*, 4, 449-463.
39. Eilers, P.H. (2003) A perfect smoother. *Anal. Chem.*, 75, 3631-3636.
40. Coombes, K.R., Morris, J.S., Hu, J., Edmonson, S.R., and Baggerly, K.A. (2005) Serum proteomics profiling—a young technology begins to mature. *Nat. Biotechnol.*, 23, 291-292.
41. Ransohoff, D.F. (2004) Rules of evidence for cancer molecular-marker discovery and validation. *Nat. Rev. Cancer*, 4, 309-314.
42. Diamandis, E.P. (2004) Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems. *J. Natl. Cancer Inst.*, 96, 353-356.
43. de Noo, M.E., Mertens, B.J., Ozalp, A., Bladergroen, M.R., van der Werff, M.P., van de Velde, C.J., Deelder, A.M., and Tollenaar, R.A. (2006) Detection of colorectal cancer using MALDI-TOF serum protein profiling. *Eur. J. Cancer*, 42, 1068-1076.
44. Hu, J., Coombes, K.R., Morris, J.S., and Baggerly, K.A. (2005) The importance of experimental design in proteomic mass spectrometry experiments: some cautionary tales. *Brief. Funct. Genomic. Proteomic.*, 3, 322-331.
45. Wang, X., Zhu, W., Pradhan, K., Ji, C., Ma, Y., Semmes, O.J., Glimm, J., and Mitchell, J. (2006) Feature extraction in the analysis of proteomic mass spectra. *Proteomics.*, 6, 2095-2100.
46. Levner, I. (2005) Feature selection and nearest centroid classification for protein mass spectrometry. *BMC. Bioinformatics.*, 6, 68.
47. Koomen, J.M., Zhao, H., Li, D., Nasser, W., Hawke, D.H., Abbruzzese, J.L., Baggerly, K.A., and Kobayashi, R. (2005) Diagnostic protein discovery using liquid chromatography/mass spectrometry for proteolytic peptide targeting. *Rapid Commun. Mass Spectrom.*, 19, 1624-1636.
48. Ye, B., Cramer, D.W., Skates, S.J., Gygi, S.P., Pratomo, V., Fu, L., Horick, N.K., Licklider, L.J., Schorge, J.O., Berkowitz, R.S., and Mok, S.C. (2003) Haptoglobin-alpha subunit as potential serum biomarker in ovarian cancer: identification and characterization using proteomic profiling and mass spectrometry. *Clin. Cancer Res.*, 9, 2904-2911.
49. Koomen, J.M., Shih, L.N., Coombes, K.R., Li, D., Xiao, L.C., Fidler, I.J., Abbruzzese, J.L., and Kobayashi, R. (2005) Plasma protein profiling for diagnosis of pancreatic cancer reveals the presence of host response proteins. *Clin. Cancer Res.*, 11, 1110-1118.
50. Wang, X., Wang, E., Kavanagh, J.J., and Freedman, R.S. (2005) Ovarian cancer, the coagulation pathway, and inflammation. *J. Transl. Med.*, 3, 25.
51. Li, J., Orlandi, R., White, C.N., Rosenzweig, J., Zhao, J., Seregini, E., Morelli, D., Yu, Y., Meng, X.Y., Zhang, Z., Davidson, N.E., Fung, E.T., and Chan, D.W. (2005) Independent validation of candidate breast cancer serum biomarkers identified by mass spectrometry. *Clin. Chem.*, 51, 2229-2235.
52. Diamandis, E.P. and van der Merwe, D.E. (2005) Plasma protein profiling by mass spectrometry for cancer diagnosis: opportunities and limitations. *Clin. Cancer Res.*, 11, 963-965.
53. Koomen, J.M., Li, D., Xiao, L.C., Liu, T.C., Coombes, K.R., Abbruzzese, J., and Kobayashi, R. (2005) Direct tandem mass spectrometry reveals limitations in protein profiling experiments for plasma biomarker discovery. *J. Proteome. Res.*, 4, 972-981.
54. Villanueva, J., Shaffer, D.R., Philip, J., Chaparro, C.A., Erdjument-Bromage, H., Olshen, A.B., Fleisher, M., Lilja, H., Brogi, E., Boyd, J., Sanchez-Carbayo, M., Holland, E.C., Cordon-Cardo, C., Scher, H.I., and Tempst, P. (2006) Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J. Clin. Invest.*, 116, 271-284.
55. Liotta, L.A. and Petricoin, E.F. (2006) Serum peptidome for cancer detection: spinning biologic trash into diagnostic gold. *J. Clin. Invest.*, 116, 26-30.
56. Matrisian, L.M., Sledge, G.W., Jr., and Mohla, S. (2003) Extracellular proteolysis and cancer: meeting summary and future directions. *Cancer Res.*, 63, 6105-6109.

57. Bank,U., Kruger,S., Langner,J., and Roessner,A. (2000) Review: peptidases and peptidase inhibitors in the pathogenesis of diseases. Disturbances in the ubiquitin-mediated proteolytic system. Protease-antiprotease imbalance in inflammatory reactions. Role of cathepsins in tumour progression. *Adv.Exp.Med.Biol.*, 477, 349-378.
58. Song,J., Patel,M., Rosenzweig,C.N., Chan-Li,Y., Sokoll,L.J., Fung,E.T., Choi-Miura,N.H., Goggins,M., Chan,D.W., and Zhang,Z. (2006) Quantification of Fragments of Human Serum Inter- α -Trypsin Inhibitor Heavy Chain 4 by a Surface-Enhanced Laser Desorption/Ionization-Based Immunoassay. *Clin Chem.*, 52, 1045-1053.
59. Marshall,J., Kupchak,P., Zhu,W., Yantha,J., Vrees,T., Furesz,S., Jacks,K., Smith,C., Kireeva,I., Zhang,R., Takahashi,M., Stanton,E., and Jackowski,G. (2003) Processing of serum proteins underlies the mass spectral fingerprinting of myocardial infarction. *J Proteome.Res.*, 2, 361-372.
60. Steinert,R., Buschmann,T., van der,L.M., Fels,L.M., Lippert,H., and Reymond,M.A. (2002) The role of proteomics in the diagnosis and outcome prediction in colorectal cancer. *Technol. Cancer Res.Treat.*, 1, 297-304.
61. Yu,J.K., Chen,Y.D., and Zheng,S. (2004) An integrated approach to the detection of colorectal cancer utilizing proteomics and bioinformatics. *World J.Gastroenterol.*, 10, 3127-3131.
62. Engwegen,J.Y., Helgason,H.H., Cats,A., Harris,N., Bonfrer,J.M., Schellens,J.H., and Beijnen,J.H. (2006) Identification of serum proteins discriminating colorectal cancer patients and healthy controls using surface-enhanced laser desorption ionisation-time of flight mass spectrometry. *World J Gastroenterol.*, 12, 1536-1544.
63. Steinert,R., Von Hoegen,P., Fels,L.M., Gunther,K., Lippert,H., and Reymond,M.A. (2003) Proteomic prediction of disease outcome in cancer : clinical framework and current status. *Am.J.Pharmacogenomics.*, 3, 107-115.
64. Gulmann,C., Sheehan,K., Kay,E., Liotta,L., and Petricoin,E., III (2006) Array-based proteomics: mapping of protein circuitries for diagnostics, prognostics, and therapy guidance in cancer. *J Pathol.*, 208, 595-606.
65. Miller,J.C., Zhou,H., Kwekel,J., Cavallo,R., Burke,J., Butler,E.B., Teh,B.S., and Haab,B.B. (2003) Antibody microarray profiling of human prostate cancer sera: antibody screening and identification of potential biomarkers. *Proteomics.*, 3, 56-63.
66. Etzioni,R., Urban,N., Ramsey,S., McIntosh,M., Schwartz,S., Reid,B., Radich,J., Anderson,G., and Hartwell,L. (2003) The case for early detection. *Nat.Rev.Cancer*, 3, 243-252.
67. Kolch,W., Mischak,H., and Pitt,A.R. (2005) The molecular make-up of a tumour: proteomics in cancer research. *Clin Sci.(Lond)*, 108, 369-383.
68. Srinivas,P.R., Verma,M., Zhao,Y., and Srivastava,S. (2002) Proteomics for cancer biomarker discovery. *Clin.Chem.*, 48, 1160-1169.
69. de Noo,M.E., Liefers,G.J., and Tollenaar,R.A. (2005) Translational research in prognostic profiling in colorectal cancer. *Dig.Surg.*, 22, 276-281.
70. Ludwig,J.A. and Weinstein,J.N. (2005) Biomarkers in cancer staging, prognosis and treatment selection. *Nat.Rev.Cancer*, 5, 845-856.

