Expert Review of Proteomics

ISSN: 1478-9450 (Print) 1744-8387 (Online) Journal homepage: informahealthcare.com/journals/ieru20

Translating clinical proteomics: the importance of study design

Evelyne Maes, William C Cho & Geert Baggerman

To cite this article: Evelyne Maes, William C Cho & Geert Baggerman (2015) Translating clinical proteomics: the importance of study design, Expert Review of Proteomics, 12:3, 217-219, DOI: 10.1586/14789450.2015.1041512

To link to this article: https://doi.org/10.1586/14789450.2015.1041512

Published online: 30 Apr 2015.

Translating clinical proteomics: the importance of study design

Expert Rev. Proteomics 12(3), 217–219 (2015)

Evelyne Maes

Center for Proteomics, University of Antwerp/Flemish Institute for Technological Research (VITO), Groenenborgerlaan 171, 2020 Antwerp, Belgium

William C Cho

Author for correspondence: Department of Clinical Oncology, Queen Elizabeth Hospital, Kowloon, Hong Kong williamcscho@gmail.com

Geert Baggerman

Center for Proteomics, University of Antwerp/Flemish Institute for Technological Research (VITO), Groenenborgerlaan 171, 2020 Antwerp, Belgium

Mass spectrometry-based clinical proteomics approaches were introduced into the biomedical field more than two decades ago. Despite recent developments both in the field of mass spectrometry and bioinformatics, the gap between proteomics results and their translation into clinical practice still needs to be closed, as implementation of proteomics results in the clinic appears to be scarce. An extra focus on the importance of the experimental design is therefore of crucial importance.

In the post-genome era, 'omics' technologies are a central part of biomedical research. As these tools provide an integrative approach for the research of an organism, they are often preferred over traditional biochemical approaches in many biological areas of study. Evidently, much can be learned by using proteomics in a clinical context. Clinical proteomics aims to understand the pathobiology of diseases at a protein level, to characterize new protein targets for drug development and therapeutic intervention, and/or to identify protein biomarker candidates for the (early) diagnosis of diseases and the prognosis or prediction of the therapeutic response [1]. Hence, numerous clinical proteomics studies have been published over the last two decades which have improved our understanding of many diseases [2]. Despite fast technological developments in the field of both mass spectrometry (MS) and bioinformatics, which steadily improved the sensitivity, specificity and throughput, limited translation to clinical practice has been achieved. This hurdle may be attributed to several causes.

First, this is due to insufficient attention paid to the study design in the discovery step. Arbitrary decisions regarding protein effect sizes (expression threshold for fold change of differentially expressed proteins) or sample sizes lead to poor (often underpowered) designs which direct the experiments to inconclusive results. When compared to genomics studies, where the importance of study design has been well documented, reports on the significance of the experimental design in proteomics studies are minimal, and only recently the subject has gained more attention [3]. The studies of Oberg and Vitek in 2009 [4], Cairns [5] and Levin [6] in 2011 were among the first key publications arguing the necessity of performing power calculations in proteomics discovery experiments given a certain technical and biological variance. Particularly in clinical proteomics discovery experiments, where small cohorts of complex human samples are used to elucidate protein expression, inter- and intra-individual variations and systematic effects would obscure a differential analysis leading to high false discovery rates and irrelevant results when an improper study design is applied. Besides the correct design, the avoidance of bias and confounding factors is also essential [7]. Indeed, pre-analytical variables (such as differences in sample storage or processing) can affect the sample quality and, thus, influence the overall quality of data. Luckily, there is an increasing awareness of the importance of controlling these variables while banking clinical samples [3]. In addition to these controllable pre-analytical variables, the effect of uncontrollable factors

Keywords: biomarker • clinical proteomics • quality control • study design • validation

EXPERT REVIEWS

 10.1586/14789450.2015.1041512 ISSN 1478-9450

(e.g., demographic characteristics) must be accounted for in the study design by implementing randomization, replication and blocking schemes [7]. Since both bias and confounding factors can be laboratory-specific, multi-center verification of the results with independent sample cohorts can increase the success in concrete validation.

Second, the lack of a well-defined research question is another underestimated reason why validation of proteomics results is lagging. A clearly defined research question and a falsifiable hypothesis will impact the choice of the study population of interest. Especially in clinical proteomics, where disease heterogeneity induces an enormous variety, a deep understanding of the disease pathology may be required in order to select the most appropriate individuals for the study [8]. For example, it is crucial to understand whether the study population is suited to test the disease-positive cases versus disease-negative controls for differential markers. As phenotypic heterogeneity across studies makes it difficult to generalize the obtained results or to replicate them in independent cohorts (which can partially explain the lack of validation), a clearly defined and focused research hypothesis as well as an adequate sample size with suitable control groups can increase the homogeneity and reduce the observed variability [8].

A third factor that jeopardizes the validation potential of proteomics studies is the lack of statistical rationale in the analysis of results in the initial stages of the discovery phase. Although adequate data analysis is crucial to provide conclusive results, the underlying assumptions of statistical tests are sometimes ignored, which leads to long lists of dubious markers that cannot be validated. Due to an increasing awareness of this problem, bioinformaticians/biostatisticians usually join proteomics team to ensure statistical results of high quality.

In addition, even when statistically sound data are obtained, model over-fitting due to detection of hundreds of proteins in small sample sizes (test set) in the discovery experiment may easily lead to false correlations (high false discovery rate) and over-interpretation of proteomic data [9]. Leave-one-out cross validation is needed and confirmation of the detected differences in a follow-up independent patient cohort (validation set) reflecting the targeted population heterogeneity is mandatory. In these validation phases, the statistical design must be implemented as indicated by the regulatory authorities to evaluate the classification accuracy of the marker combination [1]. Yet, several study design guidelines such as the prospective-specimen-collection, retrospective-blinded-evaluation rules are available [10].

A well-defined proteomic experimental design, however, should not only be statistically sound and sufficiently powered, but also requires robust tools to systematically assess the instrument performance [11]. The lack of standardization and inter-laboratory transferability in discovery and verification can, therefore, be seen as the fourth hurdle in translational proteomics. Since the complex proteome can be studied with a highly diverse toolbox of mostly complex proteomics approaches and equipment, appropriate guidelines and protocols for evaluating the quality of the measurements and lab-to-lab differences are needed to ensure that data can be reproduced by others [12]. Although efforts such as the Minimal Information about a Proteomics Experiment reporting guidelines are a first step toward more standardization [13], implementing a quality control of the appropriate performance criteria is primordial since poor system performance results in poor reproducibility of the measurements. Recent publications of Tabb [14] and Bereman [15] on quality control in proteomics are, therefore, of high importance. Other key publications of Paulovich *et al.* [16] and Abbatiello *et al.* [17] reported inter-laboratory studies where standards for benchmarking the performance of discovery (liquid chromatography-MS) and verification (multiple reaction monitoring-MS) platforms are described. Implementing these quality control guidelines will be beneficial for the transferability of the results.

Last but not the least, validation of the results in the discovery phase is often a problem, as the costs of the verification/validation procedure are usually high [3]. This largely explains the enormous amount of published biomarker candidates that did not reach clinical practice. However, with the transition in medicine toward prevention, prediction and personalized treatment, biomarkers are of growing interest for clinical practice and pharma industry. In order to meet this demand, it is of utmost importance that sufficient evidence is generated in a well-designed discovery and verification study to support the investment for a large-scale validation. Only when substantiated results are obtained, investors can be convinced in walking the uncertain path of validation with a potentially low return on investment in early stage as a large number of candidates need to be validated before a protein biomarker can be found. Fortunately, current developments in reproducible targeted MS-based procedures are facilitating the validation of these long lists as highly multiplexed MS-based assays are now realized [18,19].

The growing interest for the issues of quality in data acquisition, analysis and experimental design indicates that we are on an important turning point. Hopefully, many success stories can follow the examples of two recent US FDA-approved proteomics-based biomarker panels: the OVA-1 and ROMA multi-marker blood tests for predicting malignancy in women with an adnexal mass [20]. Along with the technological developments in the field and an increasing focus on well-designed studies, we envision that the above hurdles can finally be overcome and bring clinical proteomics to a new horizon.

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

References

1. Fuzery AK, Levin J, Chan MM, et al. Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. Clin Proteomics 2013;1:13
2. Oliveira BM, Schmitt A, Fallkai P, et al. Is clinical proteomics heading towards to "bench to bedside"? Translational Proteomics 2013;1:53-6
3. Skates SJ, Gillette MA, LaBaer J, et al. Statistical design for biospecimen cohort size in proteomics-based biomarker discovery and verification studies. J Proteome Res 2013;12:5383-94
4. Oberg AL, Vitek O. Statistical design of quantitative mass spectrometry-based proteomic experiments. J Proteome Res 2009;5:2144-56
5. Cairns DA. Statistical issues in quality control of proteomic analyses: good experimental design and planning. Proteomics 2011;6:1037-48
6. Levin Y. The role of statistical power analysis in quantitative proteomics. Proteomics 2011;12:2565-7
7. Oberg AL, Mahoney DW. Statistical methods for quantitative mass spectrometry proteomic experiments with labeling. BMC Bioinformatics 2012;13(Suppl 16):S7
8. Wallstrom G, Anderson KS, LaBaer J. Biomarker discovery for heterogeneous diseases. Cancer Epidemiol Biomarkers Prev 2013;5:747-55
9. Borrebaeck CA. Viewpoints in clinical proteomics: when will proteomics deliver clinically useful information? Proteomics Clin Appl 2012;7-8:343-5
10. Pepe MS, Feng Z, Janes H, et al. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. J Natl Cancer Inst 2008;20:1432-8
11. Frantzi M, Bhat A, Latosinska A. Clinical proteomic biomarkers: relevant issues on study design & technical considerations in biomarker development. Clin Transl Med 2014;1:7
12. Ivanov AR, Colangelo CM, Dufresne CP, et al. Interlaboratory studies and initiatives developing standards for proteomics. Proteomics 2013;6:904-9
13. Martinez-Bartolome S, Binz PA, Albar JP. The Minimal Information about a Proteomics Experiment (MIAPE) from the Proteomics Standards Initiative. Methods Mol Biol 2014;765-80
14. Tabb DL. Quality assessment for clinical proteomics. Clin Biochem 2013;6:411-20
15. Bereman MS. Tools for monitoring system suitability in LC MS/MS centric proteomic experiments. Proteomics 2015;15(5-6): 891-902
16. Paulovich AG, Billheimer D, Ham AJ, et al. Interlaboratory study characterizing a yeast performance standard for benchmarking LC-MS platform performance. Mol Cell Proteomics 2010;2:242-54
17. Abbatiello SE, Schilling B, Mani DR, et al. Large-scale inter-laboratory study to develop, analytically validate and apply highly multiplexed, quantitative peptide assays to measure cancer-relevant proteins in plasma. Mol Cell Proteomics 2015, pii:mcp. M114.047050
18. Chambers AG, Percy AJ, Simon R, et al. MRM for the verification of cancer biomarker proteins: recent applications to human plasma and serum. Expert Rev Proteomics 2014;2:137-48
19. Domon B, Gallien S. Recent advances in targeted proteomics for clinical applications. Proteomics Clin Appl 2015;9:423-31
20. Grenache DG, Heichman KA, Werner TL, et al. Clinical performance of two multi-marker blood tests for predicting malignancy in women with an adnexal mass. Clin Chim Acta 2015;358-63