Abstract  Malignant mesothelioma is an asbestos-induced, aggressive tumour with limited treatment options and very poor outcome. Currently, there are no tumour biomarkers in widespread clinical use for this disease. Soluble mesothelin is the most intensively investigated mesothelioma biomarker and has been approved by the US FDA primarily as a tool for monitoring patient response and progression. Mesothelin is elevated in the blood and effusions of patients with mesothelioma, and is rarely elevated in people with benign disease with normal renal function. However, the sensitivity of mesothelin limits its use as a stand-alone tool for the screening of the asymptomatic asbestos-exposed population—one of the primary aims of mesothelioma biomarker studies. Thus, there is an intense research effort focused on the identification of new and/or novel biomarkers for mesothelioma. Some of the challenges associated with biomarker discovery in mesothelioma are discussed.

Keywords  Mesothelioma · Biomarker · Mesothelin

Introduction

Intense interest has focussed on identifying and validating new biomarkers for disease, including for cancer. The first of the so-called traditional circulating blood cancer biomarkers, carcinoembryonic antigen (CEA) and alphafetoprotein were identified in the 1960s. Since then, only 13 blood-based and four urine-based cancer biomarkers have received approval for clinical use from the US Food and Drug Administration (FDA), a level of approval which is considered the “gold standard” for clinically useful biomarker discovery [1].

A large number of candidate biomarkers have been reported in the literature. For these new biomarkers to be brought to the clinic, they must offer advantages over existing biomarkers or clinical procedures in terms of improved sensitivity and specificity or prognostic or predictive ability and must fulfil specific clinical needs, as well as being cost-effective [2]. The majority of biomarkers approved by the FDA are approved for use in a monitoring setting. However, biomarkers have been proposed for many different settings of disease management, including diagnosis, prognosis, risk stratification, guiding therapy selection, overall management and population screening. Interests in biomarker discovery not only stem from the potential health benefits but also the huge economic potential following the introduction of a new test into the health care system [3].

Malignant mesothelioma

For mesothelioma, an asbestos-induced tumour with very poor patient outcome, 5-year survival rates of less than 5% are routinely reported [4]. Finding a biomarker that could enable early detection of mesothelioma, before symptoms develop, has been a driving focus for researchers. There is
evidence that screening can impact on cancer-specific mortality particularly for breast, bowel and cervical cancers with reports of significant decreases in disease-specific mortality following screening programmes [5]. However, for mesothelioma, no effective treatment is available. Proponents of early diagnosis for mesothelioma suggest that a timely diagnosis provides the opportunity to offer treatment to patients at an earlier time when tumours are smaller, localised and more accessible to treatments. There is some limited evidence to support this; intra-cavity immunotherapeutic trials from France showed that if malignant mesothelioma was detected early, the disease could potentially be treated in a less aggressive manner and more successfully [6]. Also, in a non-trial setting, a selected group of patients, those presenting with early stage disease and treated with multimodality therapy (surgery, postoperative chemotherapy with or without radiation therapy), have been reported to have a 46 % 5-year survival [7].

Clearly, differences exist in the requirements for a biomarker intended as a screening tool for a population compared to a biomarker for triaging symptomatic individuals. In individuals already presenting with symptoms, the risks associated with surgical confirmation of a malignant diagnosis would generally be acceptable whereas the rate of false positives in a screening programme would need to be minimised if surgery was to be a confirmatory tool. Thus, a biomarker with very high specificity would be required in such a setting. False positives could be minimised by only screening high risk individuals, and for mesothelioma, this would be those exposed to a moderately high level of asbestos. However, data on asbestos exposure is not well standardised. Many biomarker studies in the literature present patient-reported categorical data (i.e. exposed/unexposed/unknown), as opposed to data derived from lung asbestos fibre burden, job exposure matrices or correlates with atmospheric asbestos loads.

**Mesothelin**

We have been studying mesothelin as a biomarker for malignant mesothelioma for over 12 years. Mesothelin was originally identified as a glycoprophatidylinositol-linked glycoprotein of approximately 40 kDa present on the cell surface of mesothelial cells and some cancer types including ovarian, pancreatic and lung carcinomas; squamous cell carcinomas of the oesophagus and cervix and mesothelioma [8]. In collaboration with Ingegerd and Karl Eric Hellstrom, we identified that a soluble form of mesothelin was present in the blood at high levels in mesothelioma patients [9]. At the time, this protein was termed soluble mesothelin-related peptide (SMRP). Subsequently, it has been found that the detected protein is in the majority a soluble form of the cell surface mesothelin identified by Pastan and colleagues in 1994 that is shed into the circulation [10, 11].

In our original publication, we described elevated levels of mesothelin in the blood of 84 % of a relatively modest total number of mesothelioma patient samples (n=44) using an in-house assay [9]. For comparison, we examined levels of the protein in samples collected in our clinics from patients with other malignant and benign pulmonary diseases, some of whom had been exposed to asbestos, and also from a cohort of younger, female, healthy laboratory volunteers. In these early studies, a positive cut-off value was defined as three standard deviations above the average mesothelin value for the non-asbestos-exposed controls. Using this cut-off, 1 of 22 patients with asbestosis and 1 of 22 patients with idiopathic pulmonary fibrosis, plus 1 of 22 patients with non-small cell lung cancer were mesothelin positive; 157 patients with inflammatory or malignant pulmonary or pleural diseases other than MM were negative. Of the positive patients, one had an elevated creatinine at the time of the sample, a factor since shown to elevate blood mesothelin levels.

In hind-sight, there were a number of issues with the original design and implementation of the 2003 study. Firstly, the cut-off was defined in the same sample set as the sensitivity and specificity was calculated from and also that the control “healthy group” were markedly different from the cases. However, a strength of the study was the relatively large number of clinically relevant controls examined, controls that were collected in the same centre, under the same conditions and in the same time frame as the cases.

The mesothelin assay was advanced for commercial development, and the availability of this standardised assay which was simple and easy to use enabled many different laboratories around the world to independently verify that mesothelin was elevated in the blood of mesothelioma patients. Indeed, a meta-analysis examining the data from 16 studies at eleven different laboratories from 1026 individual mesothelioma patients and 3465 controls of various aetiologies, including healthy individuals, those with benign disease and those with malignancy, concluded that elevated serum mesothelin is a strong predictor for mesothelioma, but that the marker lacks the sensitivity to be used in early diagnosis [12••].

**Biomarker discovery**

Various genomic, proteomic, immunomic, imaging and other tools have been used to identify, quantify and characterise novel biomarkers for many different diseases. There are advantages and disadvantages to the various biomarker discovery platforms. Common problems for all platforms include the high false discovery rate and the identification of tens to hundreds of potential biomarkers in discovery phases that need to be validated. There is also a poor rate of conversion of
commercialisation and retrospective clinical validation and prospective clinical evaluation consists of six steps: discovery, validation, verification, assay development, standardisation in discovery platforms is being addressed by the National Cancer Institute’s Clinical Proteomic Technology Assessment for Cancer Network [20–22]. Filtering of false positive candidates identified in the discovery stage can also be achieved through the use of orthologous techniques in the secondary validation and verification stages using assays with purportedly greater assay precision (such as antibody-based techniques), though some of the problems associated with sampling can be carried over to this step. Another factor that is becoming clearer now is the accuracy of commercially available “research-only” ELISA kits. In our own studies, we have found significant discrepancy between two different ELISAs for the mesothelioma biomarker, megakaryocyte potentiating factor (MPF), whereby in matching samples one assay resulted in a sensitivity of 29 % and the other 52 % at screening, a non-symptomatic but otherwise matched population is required. Another useful control for screening studies is longitudinal samples from the same individual.

Further influencing the choice of “normal” control group used for comparative purposes in biomarker discovery studies is the biospecimen type used (i.e. blood, tumour, proximal fluid or cell lines). For example, normal pleural fluid is difficult to ethically obtain; non-malignant (normal) cells in culture generally have significantly different proliferation rates than cancerous cell lines, and the use of immortalised, for example viral transformed cell lines, is particularly artificial. In a simple case-control comparison, the demonstration of a statistically significant difference in the mean level of the biomarker between groups is not necessarily sufficient to indicate that a biomarker would be clinically useful. Indeed, in some cases, a $t$ test $p<0.05$ is simply reflective of low variability in the control group. Recently, it was elegantly described how the variability of biomarker levels as well as quantitative difference in means between groups gives a better predictor of potential biomarker clinical value. After accounting for such differences and the proportion of cases expressing a given biomarker, Skates and colleagues estimated, for a given discovery set sample size, the probability of validating a given biomarker in the verification stage [17••]. This approach highlights that more factors than the results of a Student’s $t$ test need to be accounted for in biomarker discovery studies and that many “true” biomarkers do not necessarily meet the demands of the clinic.

A commonly recognised problem in biomarker studies is that the original study shows great promise but subsequent studies do not have such strong results. Possible reasons for the lack of success in identifying new biomarkers can occur in each of the steps in the discovery pipeline including problems and biases related to experimental design, data analysis, sample collection, processing and/or storage [18, 19]. Furthermore, problems arise from the high variability and the lack of replication performed in some discovery platforms which result in a high rate of false positive candidates. Lack of standardisation in discovery platforms is being addressed by the National Cancer Institute’s Clinical Proteomic Technology Assessment for Cancer Network [20–22]. Filtering of false positive candidates identified in the discovery stage can also be achieved through the use of orthologous techniques in the secondary validation and verification stages using assays with purportedly greater assay precision (such as antibody-based techniques), though some of the problems associated with sampling can be carried over to this step. Another factor that is becoming clearer now is the accuracy of commercially available “research-only” ELISA kits. In our own studies, we have found significant discrepancy between two different ELISAs for the mesothelioma biomarker, megakaryocyte potentiating factor (MPF), whereby in matching samples one assay resulted in a sensitivity of 29 % and the other 52 % at
Another potential candidate biomarker identified following mRNA expression studies was fibulin-3, a secreted glycoprotein. Fibulin-3 mRNA was found to be on average significantly overexpressed by 7.36-fold from 37 surgical mesothelioma tumours compared with matched normal peritoneum [31]. In a follow-up study to examine fibulin-3 as a fluid-based biomarker, an initial impressive diagnostic accuracy for mesothelioma was reported; plasma fibulin-3 had a sensitivity of over 96 % at a specificity of 95 %, and pleural effusion fibulin-3 a 84 % sensitivity at a 93 % specificity using a commercial ELISA for the protein (USCN Life Science) [29]. In subsequent sample sets, however, sensitivity was noticeably less, estimated to be 40 % in Canadian samples [29] and 22 % in Australian samples [37] at a 95 % specificity. This is another example of a biomarker that showed great promise in early studies but, as is common with biomarkers, has not subsequently been proven to be useful.

For mesothelioma, as in other disease settings, the vast majority of published biomarker discovery papers report the discovery of a marker with expression levels significantly higher in cases relative to controls. That is statistical significance is the favoured study end-point rather than clinical or practical relevance. Whilst there have been suggestions that biomarker studies not be published until independent validation in external samples is performed, preferably in a clinical setting, we feel this is an unrealistic expectation given the push not only of researchers, but research institutes, universities, the media and even funding bodies for ground-breaking research with a direct clinical impact to be rapidly published. How to reconcile these two aims is not clear. Also, in many instances, if preliminary biomarker evidence remains unpublished due to imperfect validation, then other researchers with the appropriate resources to investigate these biomarkers with their own clinical samples may remain unaware of the potential of the biomarker, and a validation opportunity consequently lost. It is imperative that all suitable studies for a biomarker are reported and subsequently evaluated in a meta-analysis before the final utility of a biomarker can be decided.

Would we have found mesothelin today using a discovery platform?

Mesothelin was identified as a biomarker for mesothelioma primarily based upon existing biological understanding of the tumour. An interesting question however is, if we were to examine some of the data available from high-throughput screening platforms being used in the untargeted search for novel mesothelioma biomarkers, would we have found mesothelin?

The probability of mesothelin being identified in an unbiased biomarker study can be estimated using the assumptions described by Skates and colleagues [17••]. The available data
for this purpose for mesothelin as a blood biomarker is that approximately 50% of mesothelioma cases express the marker at diagnosis, that there is relatively limited variability of mesothelin levels in control samples and there are four or more standard deviations of the mean difference between the mean mesothelin levels in controls relative to cases (Fig. 1a). Using these assumptions, a biomarker discovery experiment limited to ten case and control samples (because of either cost constraints or availability of “normal” tissue) and with 100 samples available for the verification stage, would have approximately a 50% probability of identifying mesothelin.

Indeed examining the publically available mRNA expression data [38] (GEO accession GSE2549), mesothelin was clearly expressed in mesothelioma surgical specimens whilst not being measurable in four out of five normal pleural tissue samples analysed (Fig. 1b). Although the focus of the Gordon manuscript was not to identify biomarkers but rather improve understanding of mesothelioma tumourgenesis and pathobiology, in supplementary data they show that mesothelin was overexpressed approximately 14-fold in tumours relative to normal lung and pleural tissue ($p<0.005$). However, examining our own unpublished dataset of mRNA expression in mesothelioma cell lines and normal mesothelial cell cultures, we found no difference in the levels of mesothelin mRNA between the groups. This finding possibly reflects the observation that mesothelin is one of several molecules whose expression is downregulated in tissue culture [39].

Examining publically available quantitative mass spectrometry-based proteomic data, mesothelin was found to be present in the proteome of mesothelioma effusions (Fig. 1c) but was not identified as a candidate biomarker following analysis comparing proteins differentially expressed between mesothelioma and lung cancer [40] (ProteomeXchange identifier PXD000531). In our own unpublished proteomic dataset, mesothelin was only shown to be differentially expressed in one of three experiments.

Thus, whether or not mesothelin would have appeared on any list of candidate biomarkers following an unbiased analysis on a high throughput discovery platform would have been dependent, as expected, upon the biospecimen used (either tumour tissue, cell lines or effusions) as well as the comparator groups evaluated.

**Conclusion**

At present, the only example of a FDA-approved biomarker that was identified through a high-throughput screening biomarker discovery strategy is the OVA1® (Vermillion, USA) assay [41•]. Of note, OVA1 is not approved for use in a diagnostic setting but in a specific clinical setting of triaging women with pelvic masses to specific centres for surgical intervention. The OVA1 test consists of four proteins identified by SELDI-TOF mass spectrometry plus CA125, a well-established ovarian cancer biomarker. The panel of markers has a greater specificity than CA125 used alone. We hope that a similar approach, using a panel of biomarkers, might be successful for mesothelioma. Despite a large number of biomarkers being reported for mesothelioma in the literature, mesothelin remains the single-best blood-based biomarker for the cancer and is considered to be the “gold standard” against which new biomarkers need to be judged and remains the only one with FDA approval.
Compliance with Ethics Guidelines

Conflict of Interest Dr. Jenette Creaney, Dr. Ian M. Dick and Bruce Robinson do not have any conflicts of interests to disclose.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance
• Of major importance

1. Pavlou MP, Diamandis EP. Validation of candidate protein biomarkers. In: Ginsburg GS, Willard HF, editors. Genomic and personalized medicine. 1. 2nd ed. London: Elsevier; 2013. p. 263–70.
2. Ioannidis JP. Biomarker failures. Clin Chem. 2013;59:202–4.
3. MarketsandMarkets. Global Biomarker Market 2009. Available from: http://www.pr.com/press-release/176143.
4. Musk AW, Olsen N, Alfonso H, et al. Predicting survival in malignant mesothelioma. Eur Respir J. 2011;38:1420–4.
5. Moyer VA. Force USPST. Screening for ovarian cancer: U.S. Preventive Services Task Force reaffirmation recommendation statement. Ann Intern Med. 2012;157:900–4.
6. Boutin C, Nussbaum E, Monnet I, et al. Intrapleural treatment with recombinant gamma-interferon in early stage malignant pleural mesothelioma. Cancer. 1994;74:2460–7.
7. Sugarbaker DJ, Flores RM, Jaklitsch MT, et al. Resection margins, extrapleural nodal status, and cell type determine postoperative long-term survival in trimodality therapy of malignant pleural mesothelioma: results in 183 patients. J Thorac Cardiovasc Surg. 1999;117:54–63. discussion -5.
8. Chang K, Pastan I. Molecular cloning and expression of a cDNA encoding a protein detected by the K1 antibody from an ovarian carcinoma (OVCAR-3) cell line. Int J Cancer. 1994;57:90–7.
9. Robinson BW, Creaney J, Lake R, et al. Mesothelin-family proteins and diagnosis of mesothelioma. Lancet. 2003;362:1612–6.
10. Hellstrom I, Raycraft J, Kanan S, et al. Mesothelin variant 1 is released from tumor cells as a diagnostic marker. Cancer Epidemiol Biomarkers Prev. 2006;15:1014–20.
11. Ho M, Onda M, Wang QC, et al. Mesothelin is shed from tumor cells. Cancer Epidemiol Biomarkers Prev. 2006;15:1751.
12. Hollevoet K, Reitsma JB, Creaney J, et al. Serum mesothelin for diagnosing malignant pleural mesothelioma: an individual patient data meta-analysis. J Clin Oncol. 2012;30:1541–9. Meta-analysis confirming the consistency of mesothelin as a mesothelioma biomarker.
13. Pepe MS, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer. J Natl Cancer Inst. 2001;93:1054–61.
14. 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;100:1432–8.
15. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics. 2002;1:845–67.
16. Simpson RJ, Bernhard OK, Greuning DW, et al. Proteomics-driven cancer biomarker discovery: looking to the future. Curr Opin Chem Biol. 2008;12:72–7.
17. 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. A mathematical approach to determining the requirements for biomarker validation.
18. Diamandis EP. Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems. J Natl Cancer Inst. 2004;96:353–6.
19. Ransohoff DF. Bias as a threat to the validity of cancer molecular-marker research. Nat Rev Cancer. 2005;5:142–9.
20. Rodriguez H, Tezak Z, Mesri M, et al. Analytical validation of protein-based multiplex assays: a workshop report by the NCI-FDA interagency oncology task force on molecular diagnostics. Clin Chem. 2010;56:237–43.
21. Regnier FE, Skates SJ, Mesri M, et al. Protein-based multiplex assays: mock presubmissions to the US Food and Drug Administration. Clin Chem. 2010;56:165–71.
22. Addona TA, Abbatiello SE, Schilling B, et al. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. Nat Biotechnol. 2009;27:633–41.
23. Creaney J, Sneddon S, Dick IM, et al. Comparison of the diagnostic accuracy of the MSLN gene products, mesothelin and megakaryocyte potentiating factor, as biomarkers for mesothelioma in pleural effusions and serum. Dis Markers. 2013;35:119–27.
24. Creaney J, Yoonman D, Demelker Y, et al. Comparison of osteopontin, megakaryocyte potentiating factor, and mesothelin proteins as markers in the serum of patients with malignant mesothelioma. J Thorac Oncol. 2008;3:851–7.
25. Yamaguchi N, Hattori K, Oh-edo M, et al. A novel cytokine exhibiting megakaryocyte potentiating activity from a human pancreatic tumor cell line HPC-Y-5. J Biol Chem. 1994;269:805–8.
26. Koijima T, Oh-edo M, Hattori K, et al. Molecular cloning and expression of megakaryocyte potentiating factor cDNA. J Biol Chem. 1995;270:21984–90.
27. Hollevoet K, Nackaerts K, Thimpont J, et al. Diagnostic performance of soluble mesothelin and megakaryocyte potentiating factor in mesothelioma. Am J Respir Crit Care Med. 2010;181:620–5. Meta-analysis confirming the consistency of mesothelin as a mesothelioma biomarker.
28. Ostroff RM, Mehan MR, Stewart A, et al. Early detection of malignant pleural mesothelioma in asbestos-exposed individuals with a noninvasive proteomics-based surveillance tool. PLoS One. 2012;7:e46091.
29. Pass HI, Levin SM, Harbut MR, et al. Fibulin-3 as a blood and effusion biomarker for pleural mesothelioma. N Engl J Med. 2012;367:1417–27.
30. Zhang X, Shen W, Dong X, et al. Identification of novel autoantibodies for detection of malignant mesothelioma. PLoS One. 2013;8:e72458.
31. Pass HI, Liu Z, Wali A, et al. Gene expression profiles predict survival and progression of pleural mesothelioma. Clin Cancer Res. 2004;10:849–59.
32. Greilier L, Baas P, Welch JJ, et al. Biomarkers for malignant pleural mesothelioma: current status. Mol Diagn Ther. 2008;12:375–90.
33. Holloway AJ, Diyagama DS, Opeskin K, et al. A molecular diagnostic test for distinguishing lung adenocarcinoma from malignant mesothelioma using cells collected from pleural effusions. Clin Cancer Res. 2006;12:5129–35.
34. Gordon GJ. Transcriptional profiling of mesothelioma using microarrays. Lung Cancer. 2005;49 Suppl 1:S99–103.
35. Pass HI, Lott D, Lonardo F, et al. Asbestos exposure, pleural mesothelioma, and serum osteopontin levels. N Engl J Med. 2005;353:1564–73.

36. Grigoriu BD, Scherpereel A, Devos P, et al. Utility of osteopontin and serum mesothelin in malignant pleural mesothelioma diagnosis and prognosis assessment. Clin Cancer Res. 2007;13:2928–35.

37. Creaney J, Dick IM, Meniawy TM, et al. Comparison of fibulin-3 and mesothelin as markers in malignant mesothelioma. Thorax. 2014;69(10):895–902.

38. Gordon GJ, Rockwell GN, Jensen RV, et al. Identification of novel candidate oncogenes and tumor suppressors in malignant pleural mesothelioma using large-scale transcriptional profiling. Am J Pathol. 2005;166:1827–40.

39. Zhang J, Qiu S, Zhang Y, et al. Loss of mesothelin expression by mesothelioma cells grown in vitro determines sensitivity to anti-mesothelin immunotoxin SS1P. Anticancer Res. 2012;32:5151–8.

40. Mundt F, Johansson HJ, Forshed J, et al. Proteome screening of pleural effusions identifies galectin 1 as a diagnostic biomarker and highlights several prognostic biomarkers for malignant mesothelioma. Mol Cell Proteomics. 2014;13:701–15.

41. Hogdall C, Fung ET, Christensen IJ, et al. A novel proteomic biomarker panel as a diagnostic tool for patients with ovarian cancer. Gynecol Oncol. 2011;123:308–13. The only approved biomarker identified using a proteomics platform.