Serum Proteomic Profiling by Matrix-Assisted Laser Desorption-Ionization Time-of-Flight Mass Spectrometry for Cancer Diagnosis: Next Steps

Eleftherios P. Diamandis

Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, and Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada

Approximately 4 years ago, Petricoin et al. published a new approach for diagnosing ovarian cancer by using surface-enhanced laser desorption-ionization time-of-flight mass spectrometry (SELDI-TOF-MS; ref. 1). The principle of this method is relatively straightforward. It has been hypothesized that proteins or protein fragments released by tumor cells or their environment may enter the general circulation. By using a protein chip, which performs a crude extraction of proteins from whole serum, groups of proteins may be immobilized and then detected by using SELDI (a derivative of matrix-assisted laser desorption-ionization, MALDI), in association with a mathematical algorithm. Since that time, the method has been used by numerous investigators to diagnose other malignancies such as breast, prostate, bladder, pancreatic, head and neck, lung, melanoma, liver, nasopharyngeal cancers, gliomas, etc. (2). Invariably, all these articles reported impressive diagnostic sensitivities and specificities, in many cases approaching 100%. None of the currently available serum cancer biomarkers is characterized by such sensitivity/specificity. It is thus natural that this method has created tremendous excitement among scientists, clinicians, the public, and the media (3).

Soon after publication of the first report, this author, and others, identified methodologic and bioinformatic shortcomings (4–12). The merits and shortcomings of this technology have been widely debated in the literature and repetition is not warranted (13–16). Time will be the ultimate judge. But some issues could be addressed now: how can this technology move from the proof-of-principle stage to validation and, eventually, the patients?

The good news first. A multicenter evaluation of the reproducibility of SELDI-TOF for the detection of prostate cancer indicated that patterns obtained by the same platform in different laboratories can be satisfactorily reproduced (17). However, the bad news is rather overwhelming.

The original data published by Petricoin et al. in *Lancet* have never been reproduced. Discriminatory peaks have not been identified. Moreover, re-analysis of the raw data identified bioinformatic artifacts which could invalidate the original conclusions (12, 18). More recently, careful evaluation of methods of sample collection and processing for proteomic analysis by SELDI-TOF revealed that preanalytic variables such as sample handling could markedly influence results (10). Another report highlighted the effect of bias in influencing serum proteomic

pattern analysis for breast cancer diagnosis (9). A recent study further attempted to validate three previously identified breast cancer biomarkers, by SELDI-TOF, using patients from another institution (19). Among the three previously identified candidate biomarkers, one was not reproduced. The other two were shown to be fragmented complement C3a, a highly abundant protein in the serum, produced by the liver. Interestingly, a very large number of candidate cancer biomarkers previously identified using this technology are also liver-derived products and, like complement C3a, are acute phase reactants (6). As I have indicated earlier, it is highly unlikely that high-abundance proteins produced by the liver in response to acute inflammation are going to succeed as cancer biomarkers (5–7). Apparently, we seem to be rediscovering nonspecific cancer biomarkers that were also identified approximately 40 years ago as part of the acute phase reaction (4). All these markers have already been abandoned for their lack of specificity. Despite the fact that many authors speculate that fragmented proteins may come from the tumor microenvironment, this has not been shown conclusively, and the more likely explanation seems to be processing of these proteins by amino- and carboxy-peptidases present in the serum.

In addition to the rather unsuccessful validation studies published to date (9, 19), others continue to publish proof-of-principle studies with this methodology for various types of cancer. A recent report published in this journal concerns pancreatic cancer (20). Although these authors attempted to improve on some of the previously identified shortcomings by using higher resolution mass spectrometers, multiple protein chips for sample pretreatment, and a clinical design that included samples from three different institutions, major questions still exist. For example, no effort was made to positively identify any of the discriminatory peaks, leaving this author to speculate, based on previous experiences, that these molecules represent high-abundance proteins, which are not likely to originate from tumor cells. For more explanations, see ref. 5.

Honda et al. attempted to validate their algorithm by using an independent set of samples from a different institution. However, no conclusions could be drawn from this validation set because the number of cancers and nonmalignant diseases was very small and the data could not be generalized. One major concern with this and related studies is that for a large number of discriminatory peaks, the intensity seems to be reduced in the cancer population, in comparison to normal subjects. There are currently no useful cancer markers whose concentration is actually decreased in the serum of patients with cancer. The most useful cancer biomarkers originate from tumor cells and their concentration is increased in the circulation and correlates with tumor burden. There are no good hypotheses that could explain a serum concentration decrease of an apparent tumor marker. A possible

gene down-regulation effect in tumor cells would have been highly speculative unless the marker under discussion has absolute tissue specificity and the tumor overtakes the normal tissue, thus reducing the marker concentration in serum. A more likely explanation is that these molecules represent highly abundant proteins produced by the liver or other organs, whose concentration is decreased due to cancer cachexia or malnutrition, nonspecific effects for many cancer types.

For good cancer biomarkers, it is known that there is a proportional relationship between biomarker concentration and tumor stage. In most of the published studies, this important relationship has not been shown.

Early concerns with this technology have now been enhanced by recent reports which further underline that this methodology is very sensitive to sample collection and processing procedures, which could lead to artifactual (and usually overoptimistic) sensitivities and specificities (9). Coupled with the lack of serious validation of any of the previously published diagnostic procedures based on this methodology, I can still assume that it is highly premature to conclude from the newly published reports that this approach represents an important new diagnostic method for cancer. I believe that it is time to establish minimal requirements for future publications in the field. Some proposals are listed below:

1. Examine the effect of sample collection and storage conditions on the data (9, 10).
2. Examine other variables that could potentially bias the data, either biologically or bioinformatically, as previously described (11, 12).
3. Every derived diagnostic algorithm should be tested without re-training sets, on an equally large set of samples from different institutions and/or countries to verify its robustness.
4. It is essential to positively identify the discriminatory peaks and link them to disease pathobiology (do the identified discriminatory peaks make biological sense?). If a peak cannot be positively identified, it should not be considered as a useful marker.
5. Correlate peak intensities with tumor burden (e.g., stage).
6. Provide clues for peak intensity decreases in patients with cancer versus normal subjects.
7. Validation studies of previously reported data, even if negative, should be published, as done by Karsan et al. (9), so that this apparently highly promising technology is put into perspective.
8. Good experimental design and laboratory practices should be exercised in executing the study. This should include (*a*) randomized experimental and control samples prior to analysis, (*b*) it is essential that test sets are analyzed in a blinded fashion, (*c*) provide sufficient experimental details so that others could reproduce the study, (*d*) define how frequently the mass spectrometer has been calibrated, and (*e*) define how experimental consistency and reproducibility throughout the experiment has been monitored and controlled.

It seems to me that it is time for reviewers and editors to exercise increased scrutiny of reports dealing with this technology so that previously identified shortcomings could be avoided and further progress is facilitated.

References

1. Petricoin EF III, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to identify ovarian cancer. Lancet 2002;359:572–5.
2. Simpkins F, Czechowicz JA, Liotta L, Kohn EC. SELDI-TOF mass spectrometry for cancer biomarker discovery and serum proteomic diagnostics. Pharmacogenomics 2005;6:647–53.
3. Check E. Proteomics and cancer: running before we can walk? Nature 2004;429:496–7.
4. Diamandis EP. Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems. J Natl Cancer Inst 2004;96:353–6.
5. Diamandis EP. Proteomic patterns in biological fluids: do they represent the future of cancer diagnostics? Clin Chem 2003;49:1272–8.
6. Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. Mol Cell Proteomics 2004;3: 367–78.
7. Diamandis EP, van der Merwe D-E. Plasma protein profiling by mass spectrometry for cancer diagnosis: opportunities and limitations. Clin Cancer Res 2005;11: 963–5.
8. Master SR. Diagnostic proteomics. Back to basics? Clin Chem 2005;51:1333–4.
9. Karsan A, Eigl BJ, Flibotte S, et al. Analytical and preanalytical biases in serum proteomic pattern analysis for breast cancer diagnosis. Clin Chem 2005;51:1525–8.
10. Banks RE, Stanley AJ, Cairns DA, et al. Influences of blood sample processing on low-molecular-weight proteome identified by surface-enhanced laser desorption/ionization mass spectrometry. Clin Chem 2005;51: 1637–49.
11. Ransohoff DF. Bias as a threat to the validity of cancer molecular-marker research. Nat Rev Cancer 2005;5:152–49.
12. Baggerly KA, Morris JS, Edmonson SR, Coombes KR. Signal in noise: evaluating reported reproducibility of serum proteomic tests for ovarian cancer. J Natl Cancer Inst 2005;97:307–9.
13. Petricoin E III, Liotta LA. Counterpoint: the vision for a new diagnostic paradigm. Clin Chem 2003;49:1276–8.
14. Liotta LA, Lowenthal M, Mehta A, et al. Importance of communication between producers and consumers of publicly available experimental data. J Natl Cancer Inst 2005;97:310–4.
15. Petricoin EF, Liotta LA. Mass spectrometry-based diagnostics: the upcoming revolution in disease detection. Clin Chem 2003;49:533–4.
16. Ransohoff DF. Lessons from controversy: ovarian cancer screening and serum proteomics. J Natl Cancer Inst 2005;97:315–9.
17. Semmes OJ, Feng Z, Adam BL, et al. Evaluation of serum protein profiling by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry for the detection of prostate cancer: I. Assessment of platform reproducibility. Clin Chem 2005;51:102–12.
18. Sorace JM, Zhan M. A data review and re-assessment of ovarian cancer serum proteomic profiling. BMC Bioinformatics 2003;4:24.
19. Li J, Orlandi R, White CN, et al. Independent validation of candidate breast cancer serum biomarkers identified by mass spectrometry. Clin Chem 2005;51: 2229–35.
20. Honda K, Hayashida Y, Umaki T, et al. Possible detection of pancreatic cancer by plasma protein profiling. Cancer Res 2005;65:10613–22.