Published in final edited form as:

Occup Environ Med. 2011 October ; 68(10): 759–764. doi:10.1136/oem.2010.058966.

SELDI-TOF derived serum biomarkers failed to differentiate between patients with beryllium sensitisation and patients with chronic beryllium disease

B C Tooker¹, R P Bowler², J M Orcutt², L A Maier^{3,4}, H M Christensen¹, and L S Newman¹

¹University of Colorado Denver, Colorado School of Public Health and Division of Allergy and Clinical Immunology, Department of Medicine, School of Medicine, Aurora, Colorado, USA

²Department of Medicine, National Jewish Health, Denver, Colorado, USA

³Division of Environmental and Occupational Health Sciences, National Jewish Health, Denver, Colorado, USA

⁴University of Colorado Denver, Colorado School of Public Health and Division of Pulmonary and Critical Care, Department of Medicine, School of Medicine, Aurora, Colorado, USA

Abstract

Background—People exposed to beryllium may develop beryllium sensitisation (BeS) and, in some cases, progress to chronic beryllium disease (CBD).

Objectives—The objective of this study was to test the ability of proteomic technology to identify patterns of serum protein biomarkers that allow differentiation between BeS and CBD and thus remove the need for invasive bronchoscopic procedures.

Methods—Initially, SELDI-TOF methodology and analysis was performed on serum samples from 30 CBD and 31 BeS patients.

Results—This ‘starter set’ yielded two distinct biomarker pattern sets with eight candidate proteins. The first set differentiated between BeS and CBD with 83.3% sensitivity and 82.3% specificity, with 10-fold cross-validation of 75% and 79%, respectively. The second set of biomarkers yielded higher sensitivity (90.0%) and higher specificity (90.3%), with 10-fold cross-validation of 71.7% and 82.3%, respectively. Due to its greater sensitivity and specificity, the second set of biomarkers was used as the framework for differentiating between CBD and BeS in a second set of serum samples from 450 patients with BeS and CBD. When this larger set of samples was subjected to the biomarker framework in a blinded fashion, it yielded a sensitivity of 43.53% and a specificity of 38.93%.

Copyright Article author (or their employer) 2011. Produced by BMJ Publishing Group Ltd under licence.

Correspondence to Brian C Tooker, University of Colorado Denver, Allergy and Clinical Immunology, Building: RC2, Mail Stop: B164, 12700 E. 19th Ave, Aurora, CO 80045, USA; brian.tooker@ucdenver.edu.

Competing interests None.

Ethics approval This study was conducted with the approval of National Jewish Health.

Provenance and peer review Not commissioned; externally peer reviewed.

Conclusions—Due to these low sensitivity and specificity values, we have concluded that, currently, the unique set of SELDI-TOF derived biomarkers does not possess the qualities that would allow it to differentiate between a CBD patient and a BeS patient using serum protein biomarkers. Future refinements in sample collection or proteomic technology may be needed to improve biomarker discovery.

INTRODUCTION

Proteomic technologies have been suggested as an important set of tools to differentiate disease from non-disease.^{1–11} However, relatively few studies have attempted to validate the use of proteomics in identifying work-related disorders.²³ As a first step towards the field application of such technology, we conducted a study to determine if proteomics can differentiate workers who have become sensitised to beryllium (BeS) from workers who have chronic beryllium disease (CBD).

Inhalation of particulate forms of beryllium metal, beryllium ceramics and beryllium alloys can lead to BeS,¹ a cell mediated immune response to beryllium,^{12–15} in a proportion of exposed individuals. BeS has been shown to progress to CBD at a rate of 6–8% per year.¹⁶ BeS diagnosis requires a confirmed positive blood beryllium lymphocyte proliferation test (BeLPT) or a positive bronchoalveolar lavage (BAL) BeLPT.^{17,18} In CBD, individuals have evidence of BeS and demonstrate an inflammatory process in the lung characterised by non-caseating granulomas and/or mononuclear cell infiltrates in lung tissue.¹⁸ Currently, disease diagnosis usually requires bronchoscopy with BAL and lung biopsy.¹⁶ There is a body of literature documenting differences in proinflammatory cytokine levels such as IFN- γ , IL-2, IL-6 and TNF- α ^{19–23} between patients with BeS and those with CBD. These differences led us to hypothesise that patterns of protein expression in blood could be used to determine disease status without subjecting the patient to invasive bronchoscopy.

SELDI-TOF is a technology that combines matrix assisted laser desorption/ionisation time-of-flight mass spectrometry with unique protein binding chemistries that allows complex samples, such as blood, saliva, urine or BAL fluid,^{7,24–28} to be reduced in complexity for biomarker analysis. We hypothesised that by using SELDI-TOF protocols and classification and regression trees (CART), we would be able to discriminate between patients with CBD and patients with BeS.

METHODS AND MATERIALS

Study population

As previously published, we defined an individual with two or more positive blood BeLPT and/or a positive BAL BeLPT and negative transbronchial lavage lung biopsy as a case of BeS. A positive CBD diagnosis was defined as a positive blood and/or BAL BeLPT with pathology consistent with CBD.^{17,18} CBD was also diagnosed in individuals with a positive BAL BeLPT and evidence of lymphocytosis on lung biopsy in the absence of granulomas on biopsy.¹⁸ This protocol was approved by the National Jewish Health Human Subjects Institutional Review Board. All subjects provided informed consent prior to participation at the time of their clinical evaluation for BeS and CBD. Participants provided demographic

information and underwent venipuncture for performance of blood BeLPT and storage of serum, spirometry, body plethysmographic lung volume determination, diffusing capacity for carbon monoxide (DL_{co}) measurement and clinical bronchoscopy, including BAL cell count and differential, transbronchial lung biopsies and lavage BeLPT. Methods for the performance of these procedures were as previously published.¹⁶ The study cohort consisted of patients with BeS and CBD, diagnosed at National Jewish Health from September 1989 to October 2007. From among the samples, 31 BeS and 30 CBD serum samples were randomly chosen and used as a 'training set' for developing the SELDI-TOF protocols and CARTanalysis framework that would be used to test the blinded samples of the validation set. This validation set consisted of an additional 450 serum samples (267 BeS and 183 CBD participants) obtained from the same frozen serum repository and subjected to SELDI-TOF protocols and CART analysis as described below.

SELDI-TOF biomarker discovery training set

To avoid freeze/thaw cycles, all samples were thawed once on ice and 20 µl aliquots added to 96-well plates. Serum was fractionated using the Serum Fractionation Kit (Bio-Rad, Hercules, California, USA) as described by the manufacturer. This kit separates proteins by their isoelectric point using elution buffers with increasingly basic pH. Separating serum proteins into six fractions reduces the overall complexity of the serum sample. SELDI-TOF chip type, serum fraction and wash stringencies were empirically determined from seven randomly chosen serum samples from the training set of 30 CBD and 31 BeS samples (see figure 1) according to Bio-Rad protocols. This procedure sought to maximise the number of identifiable protein peaks while reducing peak-to-peak interference. IMAC30 copper coated SELDI-TOF chips were pre-incubated in several different solutions prior to the addition of the serum fraction as follows. Initial incubation for 10 min with 50 µl 0.1 M copper sulphate with mixing on a MicroMix 5 (settings 20, 7) (Siemens Healthcare Diagnostics, Deerfield IL, USA) was followed with excess copper sulphate solution being removed and the chip washed for 60 s with 250 µl Milli-Q water (Millipore, Billerica, Massachusetts, USA). The Milli-Q water was removed and 250 µl 0.1 M sodium acetate pH 4 buffer added to each chip for a 5 min incubation on the MicroMix 5. Following the sodium acetate solution removal, the chip was again washed twice in Milli-Q water as above. Then, 250 µl of binding buffer (low stringency: 0.1 M sodium phosphate, 0.5 M NaCl, pH 7) was added to the chips twice for 5 min of mixing on the MicroMix 5. The chip was then ready for the addition of the serum fraction.

Serum fractions 1 and 5 were added in duplicate (45 µl) along with 100 µl of binding buffer to randomly selected spots on the chips for 60 min incubations at room temperature with shaking on the MicroMix 5. Following incubation, all chips were washed three times with 250 µl of binding buffer and then allowed to dry for 15–20 min. Next, 1 µl of the energy absorbing molecule (EAM) α-cyano-4-hydroxycinnamic acid dissolved in 50:50 1% trifluoroacetic acid:acetonitrile (v/v) was added to each chip spot and air dried for 5 min. This addition of EAM was then repeated and again allowed to dry. All-In-One Protein Standard II (Bio-Rad) was suspended in the EAM and added to spot 'H' on each chip. Chips were held in the dark at room temperature until read on the SELDI-TOF.

All chips were scanned on a Protein Biosystems IIc instrument (Ciphergen Biosystems, Fremont, California, USA) which was calibrated to focus on proteins from 3000 to 35 000 Da (empirically determined as recommended by the manufacturer). Peak sizes were standardised using the All-In-One Protein Standard II and peaks were grouped into clusters using the auto-detection function of ProteinChip software (Bio-Rad). These clusters were saved and used during the biomarker validation experiments (see below).

Statistics and data analysis

Demographic and clinical parameters were analysed using SAS software for PCs. Results are expressed as median or mean values. We tested the difference between means for continuous variables using the Student t test and used Wilcoxon's rank-sum test for non-normally distributed data. We tested the difference between proportions for categorical data using χ^2 or Fisher's exact tests when appropriate. Statistical significance was defined by a two-tailed p value <0.05.

CART analysis was used to identify a possible panel of biomarkers²⁹ derived from protein peaks observed in the training set SELDI-TOF data. The target (dependent) variable was group status (CBD or BeS). All peak intensities were considered as predictor (independent) variables. Since the target variable is categorical, classification trees were used. For building our classification trees, duplicates were considered individually as equal contributors to the CART analysis and we assumed equal prior on the target variable and no penalties on predictor variables. For selecting the optimal tree, we considered the tree with the minimum cost (lowest misclassification rate and the best prediction rates (sensitivity and specificity) via 10-fold cross-validation). Cost in this instance is defined as how much an individual sample misclassification affected the overall model. Different splitting functions were tested, but the Gini classification method with no preference on even splits yielded the best cost benefit. We built trees with all the peak intensities as predictor variables with and without using automatic best predictor discovery where best predictor discovery selected different predictors based on predictor importance. Receiver operating characteristic (ROC) curves were also used to evaluate the results of a given prediction. Mean peak intensities of each member of the biomarker panel are presented in table 1 and a repeated measures analysis of variance was used to compare intensities between the two groups.

Biomarker validation using the second set

For the biomarker validation, 45 μ l of serum fraction 5 from each of the 450 samples were added to IMAC30 chips pre-incubated with copper sulphate and washed in the low stringency binding buffer (0.1 M sodium phosphate, 0.5 M NaCl, pH 7). Following the binding and washing steps (see above), the All-In-One Protein Standard II was added to spot 'H' and α -cyano-4-hydroxycinnamic acid was added to each chip spot. All chips were scanned on a Protein Biosystems IIc instrument which was calibrated to focus on proteins from 3000 to 35 000 Da as recommended by the manufacturer. Protein peak sizes were standardised as above. Following size calibration, peaks were clustered using the saved cluster information from the training set.

RESULTS

Study population training set

Table 2 summarises the demographic and key clinical data for the BeS and CBD groups for the training sample set. Patients in this training set demonstrated no statistically significant differences in the four demographic criteria of age, gender, race and smoking status between BeS and CBD (see table 2).

As expected, compared to BeS subjects, CBD patients showed increases in immunological markers in their BAL: higher mean lymphocytes (27.6×10^6 vs 8.3×10^6 , $p < 0.0001$), a higher percentage of abnormal lavage BeLPTs (63.3% vs 41.9%, $p = 0.02$) and a higher percentage of abnormal blood BeLPTs (50.0% vs 19.3%, $p = 0.015$), based on test results obtained at the time of diagnosis. There were no differences in absolute white blood cell count (WBC) between the two groups ($p = 0.67$). In terms of lung function, CBD patients had a higher mean FEV₁ compared to BeS patients (95.1 vs 86.7, $p = 0.04$), although both were still normal on average. There were no differences in total lung capacity, forced vital capacity or DL_{co} between groups.

Study population validation set

Table 2 also shows the demographic and key clinical data for the BeS and CBD groups in the validation sample set. Patients in the validation set showed no statistically significant differences in the demographic criteria of age, gender and smoking status between the two groups. However, there was a statistically significant difference in racial makeup between the two groups ($p < 0.0025$). The CBD group had a higher percentage of white subjects (97.1%) compared to the BeS group (91.3%).

As expected, CBD patients showed increases in immunological markers compared to BeS patients: higher mean lymphocytes (25.5×10^6 vs 10.0×10^6 , $p < 0.0001$), a higher percentage of abnormal lavage BeLPTs (70.1% vs 11.3%, $p = 0.022$) and a higher percentage of abnormal blood BeLPTs (56.1% vs 42.2%, $p = 0.026$). BeS patients had higher mean macrophage counts compared to CBD patients (88.5×10^6 vs 72.4×10^6 , $p < 0.0001$). There were no differences in absolute WBC between the two groups. There were no statistically significant differences in lung function between the two groups.

Biomarker discovery by SELDI-TOF: training set

In this study, the training set SELDI-TOF generated protein peak data analysed by CART identified two separate clusters of serum protein peaks with high sensitivity and specificity to distinguish CBD from BeS patients. The first classification tree was generated using data from IMAC30 fraction 1. Three protein peaks were used to generate classification trees with minimal costs and the best prediction rates (sensitivity) and specificity via 10-fold cross-validation. Figure 2 shows one possible CART tree for IMAC30 fraction 1 using each protein peak multiple times. This tree for IMAC30 fraction 1 has a cost of 0.461 but an area under the ROC curve of 0.880, a sensitivity of 93.5% and a specificity of only 80.0%. The second cluster of serum protein biomarkers, obtained using IMAC30 fraction 5, yielded even more power to differentiate between patients with CBD and those with BeS.

Figure 3 shows a CART tree for IMAC30 fraction 5. By using these five peaks (see table 1) as multiple nodes, this biomarker panel has a sensitivity of 90.3% and a specificity of 90.0%. The cost of node construction was optimised and the area under the ROC curve was significantly higher (0.930) than the ROC curve generated from the IMAC30 fraction 1.

Biomarker validation using the validation set

To confirm the results obtained with the training set, we performed a SELDI-TOF analysis of serum fraction 5 for the entire 450 samples by binding them to IMAC30 chips coated with copper (see above), since this was the serum fraction–chip combination with the best predictive ability for differentiating BeS from CBD in our training set. When the second set of samples was subjected to the clustering and CART analysis framework generated by the training set, the framework correctly identified 74 of the 183 CBD patients and 171 of the 267 BeS patients, yielding a sensitivity of 43.53%. However, the framework also identified 109 BeS patients as having CBD and 96 CBD patients as having BeS. These incorrect identifications yielded a specificity of 38.93%, with a negative predictive value of 35.96% and a positive predictive value of 40.44%.

DISCUSSION

Exposure to beryllium can lead to BeS and progression to CBD. Testing for BeS is straightforward and relatively non-invasive using a blood test, the BeLPT. However, a definitive diagnosis of CBD requires confirmation with bronchoscopy with BAL and transbronchial lung biopsy. The objective of this study was to examine whether proteomic technology could be employed as a non-invasive diagnostic method for differentiating patients with CBD from those with BeS. We considered this study design to be a necessary first step before embarking on costly field trials. The initial sample set of 61 patient samples yielded promising results. We identified two distinct sets of protein biomarkers with high sensitivity and specificity for differentiation of patients with BeS and CBD. We used the set of protein biomarkers from serum fraction 5 that bound the IMAC30 chip as the benchmark for the second, much larger set of serum samples. This serum fraction–chip combination yielded a sensitivity of 90.0% and a specificity of 90.3% in the training sample set but, unfortunately, when the larger validation set of 450 serum samples was subjected to protein peak clustering and CART analysis in a blinded fashion, the framework correctly identified only 74 of the 176 CBD patients and only 171 of the 266 BeS patients, yielding a sensitivity of 43.53%. The framework also identified 109 BeS patients as having CBD and 96 CBD patients as having BeS, resulting in low specificity and low predictive value.

This biomarker framework most likely failed to adequately distinguish BeS patients from CBD patients in the validation set because the number of samples used in the training set of serum samples was quite low. This small sample (61 total) yielded an inflated specificity that was resolved in the follow-up validation set.³⁰ Serum contains thousands of proteins that are not expressed in equal amounts. In fact, the 10 most abundant serum proteins account for ~98% of the total serum protein content.³¹ The abundance of these highly expressed proteins can make it difficult, if not impossible, to find meaningful protein biomarkers due to the highly abundant proteins consuming the entire binding area on the

SELDI-TOF chips.^{32–34} To address this issue, we reduced the complexity of the total serum sample by performing pH based fractionation of the serum proteins and by using various binding buffers and the specific binding chemistries of the different SELDI-TOF chips to selectively bind proteins with high affinity for the specific binding chip chemistries. Furthermore, we bound each serum fraction to each chip type to determine empirically which serum fraction–chip combinations maximised protein binding within our target region (3–35 kDa) without crowding the protein peaks until they were indistinguishable from each other. Serum fraction 5, bound to the IMAC30 copper coated chips, yielded the densest number of protein peaks without overlapping to the point where the protein peaks became unusable. It is probable that in using only these protein peaks, we may have excluded protein peaks that would have increased both the sensitivity and specificity for differentiation of CBD from BeS patients.

Serum collection and processing can introduce variability and potential bias in any experiment in which samples are acquired over a significant period of time.³⁵³⁶ Our serum samples were collected at one location (National Jewish Health) over a 12-year period using a standardised protocol which, we believe, has minimised collection bias. Furthermore, serum sample processing and SELDI-TOF methodologies were performed by a single technician using standard protocols. However, it is not possible to eliminate all variation when collecting samples over such a period of time. Furthermore, serum samples were not stored in a protease inhibitor. This can allow random protein destruction even when samples are stored at –70°C. Samples were used for SELDI-TOF analysis within days of preparation to reduce variation due to time.³⁷ Even with standardised protocols to minimise variation, there are protein binding issues that are inherent to SELDI-TOF.

Test sensitivity of the blood BeLPT ranges from 80.3% to 91.9% with a specificity of 97%³⁸³⁹ in published studies, making it unlikely that these SELDI-TOF derived biomarkers, with a sensitivity of 43.5%, can replace existing BeLPT in this regard. If we had found higher sensitivity and specificity values in the validation set, our next step would have been to isolate the protein biomarkers using a regimen of fractionation by pH, hydrophobicity and size. However, in light of these results we have decided that pursuing the protein identities is unnecessary at this time.

Acknowledgements

The authors would like to thank the patients who participated in our research, Dr Misoo Ellison for her statistics help, Dr Tristan Barnes for his help in establishing this collaboration and Margaret Mroz and May Gillespie for their help in retrieving both patient samples and patient information.

Funding P01 ES011810. Beryllium: Exposure, Immune and Genetic Mechanisms.

REFERENCES

1. Baggerly KA, Morris JS, Coombes KR. Reproducibility of SELDI-TOF protein patterns in serum: comparing datasets from different experiments. *Bioinformatics*. 2004; 20:777–785. [PubMed: 14751995]
2. Harezlak J, Wu MC, Wang M, et al. Biomarker discovery for arsenic exposure using functional data. Analysis and feature learning of mass spectrometry proteomic data. *J Proteome Res*. 2008; 7:217–224. [PubMed: 18173220]

3. Zhai R, Sh S, Lu X, et al. Proteomic profiling in the sera of workers occupationally exposed to arsenic and lead: identification of potential biomarkers. *Biometals*. 2005; 18:603–613. [PubMed: 16388400]
4. Toraason M, Albertini R, Bayard S, et al. Applying new biotechnologies to the study of occupational cancer—a workshop summary. *Environ Health Perspect*. 2004; 112:413–416. [PubMed: 15033588]
5. Zhang X, Jin M, Wu H, et al. Biomarkers of lupus nephritis determined by serial urine proteomics. *Kidney Int*. 2008; 74:799–807. [PubMed: 18596723]
6. Zinkin NT, Grall F, Bhaskar K, et al. Serum proteomics and biomarkers in hepatocellular carcinoma and chronic liver disease. *Clin Cancer Res*. 2008; 14:470–477. [PubMed: 18223221]
7. Bons JA, Drent M, Bouwman FG, et al. Potential biomarkers for diagnosis of sarcoidosis using proteomics in serum. *Respir Med*. 2007; 101:1687–1695. [PubMed: 17446058]
8. Akashi T, Nishimura Y, Wakatabe R, et al. Proteomics-based identification of biomarkers for predicting sensitivity to a PI3-kinase inhibitor in cancer. *Biochem Biophys Res Commun*. 2007; 352:514–521. [PubMed: 17137555]
9. Lakhan SE. Schizophrenia proteomics: biomarkers on the path to laboratory medicine? *Diagn Pathol*. 2006; 1:11. [PubMed: 16846510]
10. Tomosugi N. Discovery of disease biomarkers by ProteinChip system; clinical proteomics as noninvasive diagnostic tool. *Rinsho Byori*. 2004; 52:973–979. [PubMed: 15675343]
11. Li J, Zhang Z, Rosenzweig J, et al. Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clin Chem*. 2002; 48:1296–1304. [PubMed: 12142387]
12. Infante PF, Newman LS. Beryllium exposure and chronic beryllium disease. *Lancet*. 2004; 363:415–416. [PubMed: 14962519]
13. Kelleher PC, Martyny JW, Mroz MM, et al. Beryllium particulate exposure and disease relations in a beryllium machining plant. *J Occup Environ Med*. 2001; 43:238–249. [PubMed: 11285872]
14. Maier LA, Martyny JW, Liang J, et al. Recent chronic beryllium disease in residents surrounding a beryllium facility. *Am J Respir Crit Care Med*. 2008; 177:1012–1017. [PubMed: 18244954]
15. Newman LS, Kreiss K. Nonoccupational beryllium disease masquerading as sarcoidosis: identification by blood lymphocyte proliferative response to beryllium. *Am Rev Respir Dis*. 1992; 145:1212–1214. [PubMed: 1586067]
16. Newman LS, Mroz MM, Balkissoon R, et al. Beryllium sensitization progresses to chronic beryllium disease: a longitudinal study of disease risk. *Am J Respir Crit Care Med*. 2005; 171:54–60. [PubMed: 15374840]
17. Kreiss K, Newman LS, Mroz MM, et al. Screening blood test identifies subclinical beryllium disease. *J Occup Med*. 1989; 31:603–608. [PubMed: 2788726]
18. Newman LS, Kreiss K, King TE, et al. Pathologic and immunologic alterations in early stages of beryllium disease. Re-examination of disease definition and natural history. *Am Rev Respir Dis*. 1989; 139:1479–1486. [PubMed: 2729754]
19. Bost TW, Riches DW, Schumacher B, et al. Alveolar macrophages from patients with beryllium disease and sarcoidosis express increased levels of mRNA for tumor necrosis factor-alpha and interleukin-6 but not interleukin-1 beta. *Am J Respir Cell Mol Biol*. 1994; 10:506–513. [PubMed: 8179912]
20. Maier LA, Sawyer RT, Tinkle SS, et al. IL-4 fails to regulate in vitro beryllium-induced cytokines in berylliosis. *Eur Respir J*. 2001; 17:403–415. [PubMed: 11405518]
21. Tinkle SS, Kittle LA, Newman LS. Partial IL-10 inhibition of the cell-mediated immune response in chronic beryllium disease. *J Immunol*. 1999; 163:2747–2753. [PubMed: 10453017]
22. Tinkle SS, Kittle LA, Schumacher BA, et al. Beryllium induces IL-2 and IFN-gamma in berylliosis. *J Immunol*. 1997; 158:518–526. [PubMed: 8977230]
23. Tinkle SS, Newman LS. Beryllium-stimulated release of tumor necrosis factor-alpha, interleukin-6, and their soluble receptors in chronic beryllium disease. *Am J Respir Crit Care Med*. 1997; 156:1884–1891. [PubMed: 9412570]
24. Cadieux PA, Beiko DT, Watterson JD, et al. Surface-enhanced laser desorption/ionization-time of flight-mass spectrometry (SELDI-TOF-MS): a new proteomic urinary test for patients with urolithiasis. *J Clin Lab Anal*. 2004; 18:170–175. [PubMed: 15103681]

25. Lin Z, Jenson SD, Lim MS, et al. Application of SELDI-TOF mass spectrometry for the identification of differentially expressed proteins in transformed follicular lymphoma. *Mod Pathol*. 2004; 17:670–678. [PubMed: 15098007]
26. Ryu OH, Atkinson JC, Hoehn GT, et al. Identification of parotid salivary biomarkers in Sjogren's syndrome by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry and two-dimensional difference gel electrophoresis. *Rheumatology (Oxford)*. 2006; 45:1077–1086. [PubMed: 16522680]
27. Schaub S, Wilkins J, Weiler T, et al. Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. *Kidney Int*. 2004; 65:323–332. [PubMed: 14675066]
28. Schweigert FJ, Wirth K, Raila J. Characterization of the microheterogeneity of transthyretin in plasma and urine using SELDI-TOF-MS immunoassay. *Proteome Sci*. 2004; 2:5. [PubMed: 15341658]
29. Breiman, L.; Friedman, JH.; Olshen, RA., et al. *Classification and Regression Trees*. New York: Chapman and Hall; 1984.
30. Liu P, Hwang JT. Quick calculation for sample size while controlling false discovery rate with application to microarray analysis. *Bioinformatics*. 2007; 23:739–746. [PubMed: 17237060]
31. Zolg JW, Langen H. How industry is approaching the search for new diagnostic markers and biomarkers. *Mol Cell Proteomics*. 2004; 3:345–354. [PubMed: 14749446]
32. Diamandis EP. Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems. *J Natl Cancer Inst*. 2004; 96:353–356. [PubMed: 14996856]
33. Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003; 422:198–207. [PubMed: 12634793]
34. Gillette MA, Mani DR, Carr SA. Place of pattern in proteomic biomarker discovery. *J Proteome Res*. 2005; 4:1143–1154. [PubMed: 16083265]
35. Marshall J, Kupchak P, Zhu W, et al. Processing of serum proteins underlies the mass spectral fingerprinting of myocardial infarction. *J Proteome Res*. 2003; 2:361–372. [PubMed: 12938926]
36. 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; 51:1637–1649. [PubMed: 16002455]
37. Rogers MA, Clarke P, Noble J, et al. Proteomic profiling of urinary proteins in renal cancer by surface enhanced laser desorption ionization and neural-network analysis: identification of key issues affecting potential clinical utility. *Cancer Res*. 2003; 63:6971–6983. [PubMed: 14583499]
38. Frome EL, Newman LS, Cragle DL, et al. Identification of an abnormal beryllium lymphocyte proliferation test. *Toxicology*. 2003; 183:39–56. [PubMed: 12504341]
39. Stange AW, Furman FJ, Hilmas DE. The beryllium lymphocyte proliferation test: Relevant issues in beryllium health surveillance. *Am J Ind Med*. 2004; 46:453–462. [PubMed: 15490468]

What this paper adds

- ▶ Currently, differentiating between chronic beryllium disease (CBD) and beryllium sensitivity (BeS) requires bronchoscopy with bronchoalveolar lavage and lung biopsy with a beryllium stimulated lymphocyte proliferation test.
- ▶ There is a body of literature documenting differences in proinflammatory cytokine levels such as IFN- γ , IL-2, IL-6 and TNF- α between patients with BeS and CBD.
- ▶ Performing SELDI-TOF on a pilot set of serum samples generated a biomarker framework that differentiated between CBD and BeS with 90.0% sensitivity and 90.3% specificity.
- ▶ Validating the findings with a larger set of serum samples yielded a sensitivity of 43.53% and a specificity of 38.93%.
- ▶ Although this particular biomarker framework was not validated, the use of SELDI-TOF derived serum biomarkers for differentiating between CBD and BeS may still be possible.

Figure 1.

Representative spectra from serum fraction–chip determination. Some fraction–chip combinations yielded very low spectrum peaks (CM10 fraction 4) and some yielded minimal numbers of peaks (H50 fraction 6). Overall, the combination of IMAC30 chips incubated with serum fraction 5 yielded the most robust peak set.

Figure 2.

Classification tree two using IMAC30 fraction 1 data with multiple nodes for each protein peak. There are three peaks, which were used to construct the tree (5851.18 Da, 8486.60 Da and 3499.95 Da). Each node is sequentially labelled and shows splitting criteria.

Figure 3.

Best classification tree using IMAC30 fraction 5 data. There are five peaks, which were used to construct the tree (4255.84 Da, 5835.93 Da, 8042.86 Da, 13 163.30 Da and 15 662.00 Da). There are eight nodes due to the fact that protein peaks 5835.93 Da, 13 163.30 Da and 15 662.00 Da are used twice as nodes.

IMAC30 fraction 5 derived protein biomarkers

Training sample set					Validation sample set					
Mass (Da)	Beryllium sensitised		Chronic beryllium disease		Repeated ANOVA p Value	Beryllium sensitised		Chronic beryllium disease		Repeated ANOVA p Value
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
4255.84	2.13	2.69	0.95	1.55	0.0268	1.94	4.68	3.29	10.17	0.0727
5835.93	2.59	3.21	0.84	0.90	0.0015	1.85	1.66	1.65	1.58	0.1699
8042.86	2.76	1.69	2.40	1.24	0.2563	1.99	1.09	1.88	1.10	0.3302
13 163.30	12.72	4.10	13.13	3.84	0.6482	7.58	3.28	7.94	3.49	0.2364
15 662.00	0.82	0.64	1.85	3.81	0.1370	2.89	3.39	3.24	4.91	0.3587

The table shows the masses and relative abundance of the protein biomarkers in both the training sample set and the validation sample set. It also contains the repeated ANOVA p values for the five proteins that were used to generate the classification and regression trees (figures 2 and 3) and the repeated ANOVA p values for the same five proteins using all 450 serum samples.

Table 2
Summary of demographic and clinical characteristics at diagnosis of BeS and CBD patients*

	Training sample set		Validation sample set			
	Beryllium sensitised (n = 31)	Chronic beryllium disease (n = 30)	p Value	Beryllium sensitised (n = 266)	Chronic beryllium disease (n = 176)	p Value
Demographics						
Age, mean years (SD) (D)	53.15 (10.98)	57.33 (12.57)	0.1700	53.45 (11.41)	52.90 (9.77)	0.5923
Gender, male, n (%)	22 (70.96)	21 (70.00)	0.9300	207 (77.82)	146 (82.95)	0.1797
Race, white, n (%)	28 (90.32)	28 (93.33)	0.3300	241 (91.29)	170 (97.14)	0.0025
Ever smoked, n (%)	17 (54.84)	15 (50.00)	0.7000	155 (58.27)	89 (50.57)	0.1126
Immunological markers						
Absolute WBC, mean, $\times 10^6$ (SD) (F)	31.33 (29.05)	34.05 (20.53)	0.6695	37.13 (36.51)	39.88 (29.97)	0.3932
Lymphocytes, mean, $\times 10^6$ (SD) (F)	8.25 (10.60)	27.56 (22.38)	<0.0001	10.03 (8.30)	25.53 (19.17)	<0.0001
Macrophages, mean, $\times 10^6$ (SD) (F)	91.75 (7.15)	74.40 (18.88)	<0.0001	88.50 (8.85)	72.35 (19.87)	<0.0001
Abnormal lavage BeLPT, n (%) (G)	13 (41.9)	19 (63.3)	0.0223	29 (11.34)	117 (70.06)	<0.0001
Abnormal blood BeLPT, n (%) (H)	6 (19.3)	15 (50)	0.0149	108 (42.19)	97 (56.07)	0.02645
Lung function tests						
TLC (% pred), mean (SD) (B), (H)	108.03 (12.08)	108.30 (14.03)	0.9369	107.19 (13.96)	105.81 (14.34)	0.3268
FEV (% pred), mean (SD) (B), (I)	86.73 (15.17)	95.10 (17.19)	0.0445	92.79 (16.45)	92.91 (16.75)	0.9384
FCP (% pred), mean (SD) (G)	99.83 (12.52)	104.36 (11.80)	0.1452	104.51 (11.31)	104.25 (10.94)	0.98081
DLco (% pred), mean (SD) (A), (C), (L)	81.92 (15.14)	77.41 (15.37)	0.2586	90.77 (18.54)	93.30 (17.76)	0.1796

(A) CBD=29; (B) BeS=30; (C) BeS=29; (D) BeS=256, CBD=175; (E) BeS=264, CBD=175; (F) BeS=259, CBD=173; (G) BeS=255, CBD=165; (H) BeS=257, CBD=173; (I) BeS=259, CBD=174; (K) BeS=259, CBD=173; (L) BeS=225, CBD=158.

* Demographics, immunological markers and lung function tests for BeS and CBD patients from both the training set and the validation set.

BeLPT, beryllium lymphocyte proliferation test; DLco, diffusing lung capacity for carbon monoxide; FEV₁, forced expiratory volume at 1 s; FCP, forced vital capacity; TLC, total lung capacity; WBC, white blood cell count.