Prediction of clinical outcomes in mantle cell lymphoma using the ACB2000 pyrolysis system

Arati A. Inamdar¹, Parag Borgaonkar², Yvonne K. Remache¹, Shalini Nair², Waleed Maswadeh³, Al Limaye², A. Pete Snyder³, Andrew Pecora⁴, Andre Goy⁴ and K. Stephen Suh¹,*

1. The Genomics and Biomarkers Program, The John Theurer Cancer Center, Hackensack University Medical Center, Hackensack, NJ 07601
2. AC Birox, LLC, Newark, NJ 07102
3. U.S. Army Edgewood Chemical Biological Center (ECBC), Aberdeen Proving Ground, MD 21010
4. Clinical Divisions, John Theurer Cancer Center, Hackensack University Medical Center, Hackensack, NJ 07601

*Corresponding author: K. Stephen Suh, PhD, The Genomics and Biomarkers Program, John Theurer Cancer Center at Hackensack University Medical Center
40 Prospect Avenue
David Jurist Research Building, Hackensack, New Jersey 07601
Email: ksuh@HackensackUMC.org
Tel: 201-336-8214

Running Title: Using ionization signatures as biomarkers to predict clinical outcomes
## Email addresses

<table>
<thead>
<tr>
<th>Name</th>
<th>Email Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arati A. Inamdar</td>
<td><a href="mailto:Alnamdar@HackensackUMC.org">Alnamdar@HackensackUMC.org</a></td>
</tr>
<tr>
<td>Parag Borgaonkar</td>
<td><a href="mailto:borgaonkar.parag@gmail.com">borgaonkar.parag@gmail.com</a></td>
</tr>
<tr>
<td>Yvonne K. Remache</td>
<td><a href="mailto:YRemache@HackensackUMC.org">YRemache@HackensackUMC.org</a></td>
</tr>
<tr>
<td>Shalini Nair</td>
<td><a href="mailto:shalini.nair@acbirox.com">shalini.nair@acbirox.com</a></td>
</tr>
<tr>
<td>Waleed Maswadeh</td>
<td><a href="mailto:wmmaswad2001@yahoo.com">wmmaswad2001@yahoo.com</a></td>
</tr>
<tr>
<td>Amit Limaye</td>
<td><a href="mailto:al.limaye@logistic-solutions.com">al.limaye@logistic-solutions.com</a></td>
</tr>
<tr>
<td>Pete Snyder</td>
<td><a href="mailto:Arnold.P.Snyder.civ@mail.mil">Arnold.P.Snyder.civ@mail.mil</a></td>
</tr>
<tr>
<td>Andrew Pecora</td>
<td><a href="mailto:APecora@HackensackUMC.org">APecora@HackensackUMC.org</a></td>
</tr>
<tr>
<td>Andre Goy</td>
<td><a href="mailto:AGoy@HackensackUMC.org">AGoy@HackensackUMC.org</a></td>
</tr>
<tr>
<td>K. Stephen Suh</td>
<td><a href="mailto:KSuh@HackensackUMC.org">KSuh@HackensackUMC.org</a></td>
</tr>
</tbody>
</table>
ABSTRACT

Background

Biological and molecular heterogeneity in cancer contributes to variations in treatment response, clinical outcome, and survival. The addition of new disease- and condition-specific biomarkers to existing clinical markers to track cancer heterogeneity provides possibilities for further assisting clinicians in predicting clinical outcomes and making choices of treatment options. Ionization patterns derived from biological specimens can be adapted for use with existing clinical markers for early detection, patient risk stratification, treatment decision making, and monitoring disease progression. The instrument ACB2000 (ACBirox universal detector 2000, ACBirox LLC, NJ, USA) utilizes a system capable of generating ionized spectral signals obtained through pyrolysis, gas chromatography, and differential mobility spectrometry (Py-GS-DMS). For conceptual studies, we have used mantle cell lymphoma (MCL) as a disease model, and have extrapolated the ionization patterns to develop a multivariate algorithm comprising a qualitative ionized biochemical signature. We have tested and validated the capacity of this approach to predict the probability of a good or poor clinical outcome as a means of estimating the likely success of a particular treatment option.

Methods

A qualitative identification and classification study was conducted on serum samples from MCL patients (stage IV) with known clinical outcomes. Multiple signatures of each sample were obtained to ensure repeatability. The acquired gas chromatography and differential mobility spectrometry (GS-DMS) data were processed using a new multivariate data analysis approach in which variable selection and reduction steps were followed by receiver operating characteristic curve (ROC) analysis. Sample replicates were analyzed by randomly assigning
them to either a training or blind set to assess the precision of the resulting classification for clinical outcome.

Results

The instrument generated ionization signatures that qualitatively differentiated between MCL patient groups and estimated the probability that chemotherapy would be successful. Prediction of good outcomes (responders) (GO) or poor outcomes (non-responders) (PO) could be made with 90% accuracy.

Conclusion

Ionization signatures of patient serum samples obtained with the ACB2000 provides an efficient detection system capable of differentiating between patients who will have a good or poor outcome. The system facilitates the cross examination of current clinical biomarkers and clinical parameters to predict clinical outcomes.

Key words: Biomarker; Clinical outcome; Differential Mobility Spectrometry; Ionization signature; Gas Chromatography; Mantle cell lymphoma; Pyrolysis
Despite advances in the treatment of hematological malignancies, a subset of patients continues to develop progressive disease, to be refractory to treatment, or to relapse shortly after treatment. Hematological malignancies often pose a challenge in terms of early detection, diagnosis, and prediction of clinical outcomes [1]. With robust biomarker-based detection methods, extremely sick patients can be stratified prior to treatment and recognized as a high-risk group with poor prognosis. However, conventional methods to detect clinical biomarkers usually involve invasive procedures for procuring blood or tissue biopsy, and require lengthy laboratory diagnostic confirmation [2]. Recently, corroborative research efforts in the field of oncology have led to use of molecular biomarkers for early diagnosis and accurate prediction of prognosis for various cancers [3]. However, no simple, rapid, and non-invasive test exists that is capable of performing initial patient selection, accurately predicting treatment response, and assigning the probability of good or poor outcome. We used a technology that combines pyrolysis, gas chromatography, and differential mobility spectrometry (Py-GC-DMS) for detecting ionized particles from human biospecimens derived from Mantle cell lymphoma (MCL) patients as an example to assess the diagnostic ability of this technology to directly predict or supplement other biological biomarkers to predict clinical outcomes.

MCL is cytogenetically characterized by a t(11;14) translocation and a bcl-1 rearrangement resulting in the overexpression of cyclin D1 [4]-[6]. Presence of complex karyotypes, deregulation of key cell cycle regulators associated with p53 mutations, CDK4 activation, p16/CDKN2A inactivation, and inactivation of DNA damage response pathways are associated with aggressive clinical behavior and poor prognosis [7]-[9]. A high proliferation index as assessed with Ki67/MIB-1 immunostaining, SOX11 overexpression, p53 alterations, and blastoid morphology have been reported to predict a poor outcome in
MCL patients [10]-[12]. Other serum biomarkers such as higher levels of βeta-2 microglobulin, free immunoglobulin light chain in serum (either monoclonal or polyclonal), IL-2Rα, IL-8, and MIP-1β correlate with poor prognosis and inferior outcomes [13]-[15]. Although morphological features along with genetic and serum biomarkers appear to be useful for predicting outcomes, their prognostic significance has not been confirmed in all studies [16],[17]. Furthermore, assessment of overall prognosis using the MCL International Prognostic Index (MIPI) lacks accuracy and is dependent on the treatment regimen [18]. In view of these deficits, it is evident that there is a lack of diagnostic methodology that can predict the clinical outcome before initiation of treatment for MCL.

Gas chromatography (GC)-related technologies have shown cost-effective utility in clinical diagnostic settings to resolve issues posed by the limitations of current clinical diagnostic procedures and laboratory associated costs. Several GC and other related “sniff and tell” technologies were previously developed for detection and biosensor related purposes [19]-[21]; these have included detection of infections and diagnosis of metabolic disorders [22],[23]. A combination of gas chromatography and mass spectrometry (GC-MS) has been used for cancer detection via a “breath analyzer” that detects volatile organic compounds (VOC), which are elaborated from different types of cancer, including lung and breast cancers [24]-[26]. More recently, GC-related technologies have been more widely applied in clinical settings to identify serum metabolites and biomarkers for different cancer types, including gastric, colorectal, esophageal, and pancreatic cancers [27]-[30]. Similarly, GC and MS technologies have also been developed for diagnosis of hematological malignancies [31]. Ionization pattern recognition has been used in detection technologies, and has recently been under investigation for potential use in diagnosis of cancer and infectious diseases and for rapid analysis of potential bioterror agents [32]-[36]. For detection of pathogens or markers of interest, approaches to ionization pattern recognition have used a multivariate algorithm that can immediately differentiate the unknown from the known [37].
In this study, we analyzed serum samples of MCL patients using “ionized”
biochemical signatures produced by the ACB2000 instrument. Signatures were first
quantified, and the signal intensities were then differentiated using new multivariate data
analysis algorithm approach and were finally associated with the clinical phenotype of
interest to generate an ionization signature map. In combination with analysis of DNA,
protein, or metabolite expression, the ionization signature map provides a biomarker that
permits patient-stratification and prediction of clinical outcome with high sensitivity and
specificity. The ionization signature map generated by this technology provides a “second
opinion” that is independent of existing clinical biomarkers and established scoring systems,
thus, enhancing clinical decision-making.
Methods

Study Population

This study was approved by the Hackensack University Medical Center’s (HUMC) Institutional Review Board (IRB) (reference # CR00002851) and was performed under MCL study protocol number, PRO00001689 titled “Biomarker studies on Mantle Cell Lymphoma”. Informed consent was obtained from all patients prior to enrollment. Aliquots of serum samples obtained under standard of care procedures were registered in the HUMC Tissue Repository and stored at -80°C. Serum samples from untreated, newly diagnosed, chemo-naive (CN) (n = 15) and relapsed, chemo-exposed (CE) (n = 6) MCL patients (total n=21) were studied. Clinical characteristics of the patients used in this study include leukemic phase status (increased WBC/lymphocyte count); blastoid (presence of blastoid cells) status; MIPI score (low, intermediate, and high); pre-treatment history (CN versus CE), type of treatment regimen administered, clinical outcome [good outcome (GO) versus poor outcome (PO)], progression-free survival (PFS, from date of first line treatment), and current status (alive or deceased) (Additional file 1, Table 1).

To evaluate the system, MCL patients were divided into GO (n = 12/21) and PO (n = 9/21) based on a median overall survival cut-off of 35 months irrespective of their prior treatment status. Among 15 Chemo-naive MCL patients, 11 patients who subsequently received chemotherapeutic regimen, HCVAD (Hyper-fractionated – Cyclophosphamide, Vincristine, Adriamycin and Dexamethasone) with or without rituximab as first line treatment (n=11) were further divided into good outcome (GO, n = 6/11) or poor outcome (PO=5/11) according to the following clinical outcome criteria: GO=relapse-free/progression free survival > 35 months after initial treatment, or PO=more than 2 relapses or death < 35 months after initial treatment (Additional file 1, Table 1, gray portion). Since presence of
leukemic phase, blastoid condition, and high MIPI score in MCL are associated with poor prognosis, we performed contingency analysis to determine the association of these clinical features with the prognosis for the MCL cohort used in our study. Our analysis suggested that there was no significant association between these clinical features and clinical outcome (as defined above) in this study (Figure 1).

ACB2000 Py-GC-DMS Detection System

A serum sample (2.5 µl) was applied to the sample probe, which was inserted into the pyro-tube (1) and heated to 300°C to generate pyrolytic fragments in the vapor phase. Vaporized pyrolytic analytes were transported by a flow module (5) which injected into the GC module (3) where the analytes were separated. The separated analytes entered an ionization chamber inside the Differential Mobility Spectrum (DMS) module (4), and the ionized analytes were detected (Figure 2). Typical DMS settings used for separation of analytes were based on ion differential mobility in alternating electric fields, in the compensation voltage (V) linearly increased from -40 V to +15 V in approximately one second. Positive and negative ion signal intensities were simultaneously recorded by DMS at each compensation voltage. The output spectrum generated by the DMS consisted of 60 signal intensity values each from the negative and positive ions (120 points total) acquired at an average rate of 1.5 spectra per second. The system generated a single relative intensity (RI) for each sample by adding the intensity of signals corresponding to the detected ions. Each sample was run in triplicate, and average relative intensity from three runs was calculated for each sample.

A new multivariate data analysis approach: variable selection and reduction using receiver operating characteristics (VSR-ROC) curve method algorithm
A new multivariate algorithm was developed using ROC statistics described by Maswadeh and Snyder (2012) [37] for the classification and identification of any two groups; in the current study, the groups were good outcome (GO) and poor outcome (PO). The data analysis algorithm comprised two components. The first component of the algorithm was variable ranking and selection; the second component was multivariate analysis, performed by combining the highest ranked variables into a single variable, which was used to generate a ROC curve. The ACD (area between the curve and the diagonal) was calculated to separate GO and PO groups at the maximum degree. The major variables, True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) of ROC curve statistics [37] as well as new variables unique to this analysis such as Percentage Separation, Positive Population Center likelihood (PPCL), Negative Population Center likelihood (NPCL), Confidence level (CL) were further calculated for the two groups (Additional file 2, part A and B).

Algorithm component 1: variable selection and ranking:

Using the ACB2000 system, experimental data files were generated for serum samples of 20 MCL patients with known clinical outcomes (11 with good outcomes, 9 with poor outcomes). As described in Additional file 1 (part A) each experimental data file generated a relative intensity for each sample. Upon subtracting the variables (pixels) corresponding to ionic patterns found to occur in common in both GO and PO groups, unique ionization signatures consisting of 572 variables (pixels) were recognized. These variables were further used to calculate the ACD [37] between the GO and PO groups at each of 572 variables. For example, Table A1 shows the average intensity of signal for pixel (variable) # 195 for MCL patients belonging to GO and PO groups. Figure A1 shows the frequency plot generated using pixel signals listed in Table A1 (See Additional file 2, part A). Here, the frequency plot
indicates a moderate separation of good (green diamond, ◢) from poor (red box, □) clinical outcome (Additional file 2, Figure A1). To improve the distinction between GO and PO patients, the ACD were calculated for these 572 variables, and the variables that gave the best possible degree of separation between data-points of the GO and PO groups were used for further analysis. The low ACD variables were considered to be noise or to represent variables that were not sufficiently sensitive to discriminate between GO and PO groups [37]. In total, 59 variables, representing unique biomarkers, were identified that were capable of distinguishing between the two groups with a high degree of separation. These 59 variables, arranged in order of highest to lowest ACD, are listed in Table A2 (Additional file 2, part A).

Algorithm component 2, top-variables combination and ROC statistics:

The second component of our new multivariate data analysis approach utilizes mathematical steps (Steps 1 through 4, which are explained in Additional file 2, part B) described by Maswadeh and Snyder (2012) [37]. In brief, a new ACD was calculated from the ACDs of first two variables (for example, #195 and #197) as described in Step 1 and Step 2 (Additional file 2, part A, Figure A2) and degree of separation was calculated (Additional file 2, part A, Figure A3). The new ACD was then used to calculate a second new ACD with the next variable (Additional file 2, Steps 3 and 4). This iterative process was carried on in a sequential manner, ultimately reducing all 59 variables to a single variable. The final single variable was used to calculate an ACD representing the maximum degree of separation between GO and PO patients (Additional file 2, part A, Figure A4). The corresponding ROC was used to calculate major variables (TP, FP, TN, and FN) of ROC curve statistics as well as new variables unique to this study (% Separation, PPCL, NPCL, CL) between the GO and PO groups (See the description of steps in Additional File 2, part B along with Figure B1 and Table B1).
Data Analysis

MCL patient serum samples with known, good or poor clinical outcomes (Additional file 1, Table 1) were analyzed using the ACB2000 system. Each experiment was performed in triplicate to generate the raw data set. Raw data files of all samples were further processed to identify ionization signature patterns restricted to either good or poor outcome by eliminating the variables with low ACD or ionization signatures that were shared between the two groups. This processing by the first component of the VSR-ROC algorithm identified 59 variables. The second component of the VSR-ROC algorithm reduced these 59 variables to a single variable representing the best possible separation between GO and PO groups. The unique ionization signatures generated after application of the VSR-ROC algorithm were assigned to either the good or poor outcome group based on known clinical outcome. Then a second known set of MCL serum samples (n=21) was tested to verify the ability of the VSR-ROC algorithm to identify patients with GO and PO.

Statistical analysis

Data are summarized as mean ± standard error of the mean (SEM) for relative intensity (RI) along with the probability (% P) of patients experiencing favorable prognostic factors so as to be labeled as good outcome (GO) or patients experiencing poor prognostic factors and therefore having a poor outcome (PO). The correlation between the clinical features and clinical outcome for the cohort was assessed using contingency table. Statistical analysis was performed using GraphPad Prism software. The details of the statistical tests are indicated in the figure legends.
Results

Py-GC-DMS System Generates a Three-Dimensional Ionization Signature Based on DMS Signal Intensity

Three-dimensional DMS signals were plotted as a function of compensation voltage (\(x\) axis), retention time (\(y\) axis), and an ion abundance (DMS signal intensity; \(z\) axis). A 2.5 µl sample of patient serum was sufficient to generate 120 points that subsequently, separated into 60 negative and 60 positive ionized spectral points. Typical waterfall display consistent with 3D-\(x, y, z\) format (Figure 3A) shows that ionized spectra represent ionized materials from patient serum sample in continuous spectrum format that is divided into positive and negative areas. These same raw data points can be converted into a 2-D ((\(x, y\), \(z\)) pixel array format as shown in Figure 3B where each peak represents a sum of data points quantifying an average intensity of pixels. All data output from the triplicate analyses were quantitatively analyzed by Py-GC-DMS system and presented as relative intensity for each patient serum sample used in this study (Additional file 1, Table 2).

Relative Signal Intensity of Ionization from Each Outcome Group is Similar but Ionized Pattern is Different

An individual relative intensity value for each MCL case was generated by averaging signal intensities of variables with good signal (>9% of optimized signal) from triplicate runs. The average RI values for GO and PO groups were calculated by averaging the relative intensity of all MCL patients sorted according to their treatment outcomes (GO vs PO) (Additional file 1, Table 2). These average RI values were not significantly different between GO (33.33 ± 5.62, N=12) and PO (40.50 ± 8.629, N=9) where \(p = 0.4139\), suggesting that RI values alone cannot differentiate clinical outcomes. The average RI value was determined for newly
diagnosed chemo naïve (CN)-MCL patients prior to HCVAD (Table 2, boxed data from GO and PO groups). The CN-MCL patients with GO had an average RI of $32.85 \pm 6.675$, $N=6$, and CN-MCL patients with PO had an average RI of $29.19 \pm 9.257$, $N=5$, $p=0.3743$ indicating that differences between GO and PO groups were not significant (Figure 4B).

**In Silico Modification of Ionized Signatures and Optimization of the Multivariate Algorithm Increases Accuracy for Predicting Clinical Outcome.**

Analysis of ionization signatures permitted identification of a total of 572 pixels as independent variables specific to each ionization signal on the output grid that were unique to good or poor outcome groups but were not shared between the two groups. Therefore, the signature output from each GO and PO sample was only represented by group-specific ionization points. Receiver Operating Characteristic curve analysis was performed on these 572 pixels, and the 59 pixels with the highest ACD were selected (Additional File, Table A2). These 59 variables and corresponding ionization signatures were further optimized by application of the VSR-ROC algorithm (Additional file 2: Part A and B). The steps (Additional file 2: Part B) were repeated with all 59 selected pixels (variables) which reduced the set of variables to a single, aggregate variable for each outcome group. The final frequency plot demonstrated the best possible separation of patient populations with good versus poor clinical outcome (Additional file 2, Figure A4). The degree of separation between good and poor outcomes from the aggregate variable was 90%, and showed significant differences in the pattern and location of ionized signatures between the good and poor outcome groups that could be visually identified (Figure 5 A,B).

**A new algorithm and VSR-ROC statistics on ionization signatures identify good and poor clinical outcome MCL patients with high predictability and accuracy**
The VSR-ROC algorithm used to construct a database of ionization pixels with best ACD was tested against a set of MCL samples with known clinical outcomes. Based on the analysis of the ionized-signature pattern of this set of MCL samples, the probability (\(\%P\)) of obtaining a unique ionization signature pattern specific for good outcome or poor outcome groups was calculated using percent GO and percent PO from Additional file 1, Table 3. The probability that a sample from a good outcome patient would show the ionization signatures specific for good outcome (defined as true GO) was 88% (\(p<0.0009\)) for all GO patients irrespective of prior treatment (Figure 6 A). The probability that a sample from a poor outcome patient would show the ionized signature specific for poor outcome (defined as true PO) was 89% (\(p<0.0081\)) for all PO patients irrespective of treatment (Figure 6 A). Similarly, the predictability for samples from chemo-naïve patients treated with HCVAD regimen being true GO and PO was 92% (\(p=0.004\)) and 80% (\(p=0.1043\)), respectively (Figure 6 B). Thus, in the analysis of this set of patients, the VSR-ROC algorithm differentiated ionization spectra that were specific to good (11/12) or poor (8/9) clinical outcomes with accuracy levels of 92% and 89%, respectively, at 70% confidence. Thus, we anticipate that the spectral analysis will provide 100% accuracy for predicting clinical outcome if the confidence level is lowered.

**Discussion**

Cancer is the second most common cause of death in the U.S., and accounts for nearly one of every four deaths [38]. Although there have been advancements in cancer diagnosis and management, clinical trials and research studies indicate the need for innovative personalized treatment regimens, especially because not all patients respond to treatment or else succumb to multiple relapses. Unfortunately, the conventional scientific approach alone is minimally effective, and the research communities need to consider alternate and innovative approaches that will supplement and support the clinical findings [39]. Discovery of biomarkers for
predicting clinical outcomes is gaining momentum, but a robust analytical tool that can support biomarker-based conclusions as an independent secondary method would enhance the accuracy of diagnosis, and would be especially helpful for differentiating patients who are at high risk for poor clinical outcome. The availability of such a tool would facilitate patient selection, would accelerate delivery of precision medicine, and would likely result in better cancer care.

Pyrolysis is a form of thermolysis that couples extremely high temperatures and an oxygen-depleted environment for optimal vaporization of samples. Pyrolyzed analytes are further separated by Py-GC-MS, and the resulting ionized signature represents a unique biochemical fingerprint of that specific sample. This technology has previously been used to test biospecimens from humans for various purposes. Three decades of Py-GC-MS research have improved the minimal volume required for accurate analyses [40]-[42], leading to a variety of uses. For example, Py-GC-MS has been used as a diagnostic tool to discriminate between leukemic and normal white blood cells [43],[44]; screening bacterial species [45]; analysis of DL-lactin/glycolic acid composition for orthopedic use [46], and structural analysis of neuromelanin in Parkinson’s disease [47]. More recent uses of Py-GC-MS include detection of volatile organic compounds in exhaled breath [48] and isolation of malignant B cells from patients with chronic lymphocytic leukemia (CLL) for prognostic studies [31]. Similar to these applications, our data show that this technology is likely to be useful for predicting clinical outcomes from analysis of MCL patient sera, which would have implications for selecting the most appropriate therapeutic approach.

We have selected a liquid tumor model to test and validate the use of the Py-GC-DMS system. Because of its pathological heterogeneity and complexity, MCL typically carries a poor prognosis with a propensity to develop drug resistance, primary treatment failure, and early relapse, which results in shortened survival ranging from a few months to four years [49],[50]. To provide better alternative treatment options, patients who are at high risk to
experience a poor outcome must be identified early and with high accuracy; however, this is a
difficult task with current clinical methods. Current clinical markers for MCL include cyclin
D1/D2/D3, translocation status t(11;14)(q13;q32), Ki67 and SOX11. These parameters, in
combination with flow cytometric analysis for immunophenotypic patterns, cell morphology,
and MIPI scores [51], [52], are used for diagnostic and prognostic purposes. For example,
SOX11 is the most recent clinical diagnostic marker for mantle cell lymphoma, and is
capable of detecting both cyclin D1-positive and -negative MCL cases [53]. Similarly,
clinical parameters such as MIPI score and cell morphology, although routinely used to
determine the prognosis for MCL patients, are of marginal prognostic significance, and lack a
high degree of sensitivity or specificity [18]. Even in our cohort, we observed that patients
with a low MIPI score were as likely to die as were patients with a high MIPI score.
Similarly, high beta-2 microglobulin, leukemic, and blastoid status failed to correlate
significantly with clinical outcome (Additional file 1, Table 1 and Figure 1).

Our data showed that the multivariate algorithm based on Py-GC-MS approach can
potentially be applied with current diagnostic and prognostic methods for improving accuracy
for stratification of patients and predicting clinical outcome with high accuracy and
reliability. Optimization through multiple repetitions using modified multivariate statistical
analysis algorithms, including a two-component VSR-ROC algorithm, was carried out to
increase the accuracy of the ionization signatures to produce a most robust and consistent
high-intensity ionization pattern, obviating the need to maintain a large set of signatures.
After optimization, typical and visually distinguishable patterns of ionization signatures of
MCL patients belonging to good and poor clinical outcomes were obtained (Figure 5 A,B).
The results of our study indicate that the optimized VSR-ROC algorithm and the visual
display of ionization signatures can be used in combination with clinical data sets to
differentiate patient samples according to their likely outcome.
The optimized ionization signatures specific for GO and PO groups were then used to determine the ability to predict the clinical outcome on a second set of patients (Additional file 1, Table 3). Our results showed that the ionization signatures can be further stratified to support prediction of clinical outcome for all treatments combined or for patients receiving HCVAD treatment. In this dataset, the probability between groups of identifying true non-responders/PO and false responders/GO was found to be non-significant, mainly due to small sample size (n=5) (Figure 6 A,B); however, a larger cohort with greater than 100 MCL biospecimen would provide false negatives and positives. We expect that the high accuracy of prediction (>90%, p<0.005) from the current data will not be significantly reduced as the size of cohort is increased if the clinical outcome criteria remain the same as used in this study. Specific ionized fragments that differentiated good versus poor outcome groups could be further purified via high resolution chromatographic methods and this remains as our future goal.

We find several advantages of Py-GC-DMS-based ionization signatures for supporting current clinical MCL markers as an independent analytical tool for guiding choices of individualized treatment for patients. The analytic methods employed in our study (1) required only 2.5 µl of serum per run, (2) took less than 90 seconds for detection time, (3) eliminated any special handling and chemical reagents, (4) provided automation of all data processing, and (5) have the inherent capacity for addition of new signatures to the existing database to improve accuracy. Our study suggests that Py-GC-DMS ionization signatures in combination with existing clinical markers could support clinical decisions for treatment strategies, risk stratification, and reducing the cost of cancer care. As a predictive tool, this system promises to provide information that will help avoid unnecessary treatment expenses for patients with a high likelihood of being poor responders, providing for early consideration...
of adjuvant therapies for such high-risk patients, and offering a convenient and rapid method
for predicting the response of an individual patient to a therapeutic agent.

Conclusion

The Py-GC-DMS system can be readily adapted to existing clinical markers to support
stratification of patients for diagnosis, prognosis, and prediction of clinical outcomes in MCL
and possibly in other cancers. The algorithm can be extrapolated to include more ionization
profiles. The systematic analysis of ionization signatures can provide prognostic information
of high specificity and sensitivity relevant to the patient selection process. The ionization
profiling can be incorporated into current clinical decision-making as an independent
technological approach, and this technology can potentially improve point-of-screening and
point-of-care clinical protocols.

List of Abbreviations

1. bcl-1 b-cell lymphoma-isoform 1
2. CDK4 Cyclin dependent kinase 4
3. CDKN2A Cyclin dependent kinase inhibitor 2A
4. MIB-1 Mindbomb 1
5. SOX-11 Sry-related HMG box
6. IL-2Rα Interleukin-2 receptor alpha
7. IL-8 Interleukin-8
8. MIP-1β Macrophage inflammatory protein-1 beta
Competing Interests

The authors declare that they have no competing interests.

Authors’ Contributions

AI analyzed data and drafted the manuscript. PB and YR carried out AC2000 experiments and generated preliminary version of the manuscript. SN, WM, AL and APS participated in analyses of data. AP and AG participated in discussion. KSS designed and directed the project, supervised all experiments, analyzed research data, edit and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We thank the John Theurer Cancer Center and Hackensack University Medical Center for supporting funding and preparation of this manuscript. We also thank the Elyssa and Jack Schecter Family Fund, Leon Lowenstein Foundation, Walter & Louise Sutcliffe Foundation and Newman’s Own Foundation for supporting research at the Genomics and Biomarkers Program of HackensackUMC.
References


analysis of cell proliferation: validation of a simplified method suitable for multi-


Figure Legends

Figure 1. Lack of association between clinical outcome and clinical features. Contingency analysis was performed to determine the association between clinical outcome (good or poor) in MCL patients and clinical markers including leukemic phase (A) (p= 0.575), blastoid condition (B) (p= 0.4135), MIPI score (C) (p=0.7141), and β₂-microglobulin (D) (p= 0.8454). Where no-significant association between clinical outcome and defined clinical features was found contingency analysis was performed and analyzed using Fisher's exact test or chi-square test.

Figure 2. ACB2000 Py-GC-DMS detection system. The system and close-up view of the components involved with sample processing: (1) Pyro-tube, (2) GC injector, (3) GC, (4) DMS, (5) flow control module, (6) electronic control boards, (7) internal PC adapter, (8) DC input.

Figure 3. The display pattern generated by the ACB 200 Py-GC-DMS detection system. A. Typical three-dimensional (3D) waterfall display of raw data with an average rate of 1.5 spectra/s demonstrates significant signatures in both the negative ion-space (right yellow axis) and the positive ion-space (left green axis). B. A two-dimensional (2D) pixel format obtained from 3D raw data representing 572 pixels as variables being incorporated into the multivariate algorithm for calculations.

Figure 4. Analysis of correlation between relative intensity (RI) of signals and clinical outcomes. (A) Comparison of RI of GO and PO groups irrespective of previous treatment history. (B) Comparison of RI of GO and PO groups in patients who were chemo-naïve prior
to starting treatment. Each value is the average of RI from 12 GO and 9 PO samples; each sample was run in triplicate. No significant differences in the GO and PO outcome groups were found. Error bars indicate SEM.

Figure 5. Ionization signature-pattern of MCL samples. Ionization signatures for an MCL patient with good outcome (A) and a patient with poor outcome (B) have distinct patterns.

Figure 6. The ionization signature pattern specific for GO and PO groups is potentially capable of accurately discriminating between good and poor outcome groups among MCL patients. (A) The probability that a sample from any patient with a GO signature pattern irrespective of prior treatment history would truly belong to the GO group (true GO); the probability that a sample with a PO signature pattern would truly belong to the PO group (true PO); the probability that a sample with GO signature would be falsely considered under the PO group (false PO); the probability that a sample would be falsely considered under the GO group (false GO). The significance of differences between true GO and false PO groups, and true PO and false GO was analyzed using two-tailed paired t test; **p<0.01, ***p<0.001. (B) The probability that a sample from a chemo-naive patient would be a true GO, false PO, true PO or false GO. Significance was analyzed using two-tailed paired t test. NS, non-significant; **p<0.01. Error bars indicate SEM.
Table description:

Table 1: Summary of the clinical information of the MCL sample cohorts used in this study. MCL patients were classified into two groups, GO and PO, based on clinical criteria: \[(GO=\text{relapse-free/progression free survival} > 35 \text{ months}) \text{ and } (PO=\text{more than 2 relapses or death} < 35 \text{ months})\].

Table 2: Summary of the clinical information of all patients and the relative intensity of the corresponding sample detected using the ACB2000 Py-GC-DMS detection system. Patients who were chemo-naïve are boxed.

Table 3: Classification output (analysis) of the MCL training test sets. A data file generated from raw data collected in triplicate on the ACB2000 Py-GC-DMS detection system and analyzed by multivariate algorithm. Samples \((n=21)\) with a known clinical outcome were tested. Probability of being GO or PO indicated by gray shading.
Figure 1.
Figure 3.
Figure 4
Figure 5.
Figure 6

A

B
Additional files provided with this submission:

Additional file 1: Inamdar et al Additional File 1_Tables.docx, 27K
http://www.biomedcentral.com/imedia/1689595571746309/supp1.docx
Additional file 2: Inamdar et al Additional File 2.docx, 139K
http://www.biomedcentral.com/imedia/2204322751746311/supp2.docx