Proteomic analysis of seminal and vaginal fluid for developing improved biomarkers in contraceptive research

(Andrzej Kulczycki, William Grizzle and Denise Oelschlager)

Biomarkers of semen exposure were first developed in forensic medicine. As part of legal investigations, sperm and prostate-specific antigen (PSA) have been used, among other biomakers, to establish if ejaculation occurred during an assault on a woman (Kulczycki 2008). More recently, there is growing recognition that biomarkers of semen exposure can help validate reports of sexual behavior and assist development of new vaginal methods of contraception and STI/HIV prevention. The twin imperatives to expand contraceptive options and expedite contraceptive research and development have provided motivation. So has the HIV pandemic, which has emphasized the need for cervical and vaginal barrier methods (both physical and chemical). Semen biomarkers can expedite such development. Several biomarkers have been proposed, but each has its disadvantages (Lawson et al.1998; Mauck and Doncel 2007).

This paper presents the rationale for a new approach to biomarker discovery for detecting vaginal exposure to semen and for assessing the presence of vaginal fluid in male genital secretions after intercourse. Further, we report initial results of an exploratory assessment of such a methodology, one which shows very exciting potential for moving the entire field forward and which may have several valuable applications to the reproductive health field. For example, such exploratory work may hold direct relevance to improving measurement of condom failure, as well as for gauging the effectiveness of other barrier methods, including candidate microbicides.

Use of biomarkers in clinical, demographic and reproductive health research

Biomarkers are typically defined as physical, functional or biochemical indicators of a physiological or disease process that has diagnostic and/or prognostic utility. Biomarkers have been rapidly adopted in clinical research with multiple applications (e.g. diagnostic, prognostic, predictive, monitoring) and as surrogate endpoints in clinical studies. In biomedicine, novel and innovative technologies for biomarker discovery and validation are rapidly replacing more traditional methods. Proteomics technology is being explored for potential uses in cancer and other disease fields, including research for proteins that may serve as biomarkers. Proteomics additionally offers potentially major new opportunities for developing biomarkers of semen and vaginal fluid. To the best of our knowledge, our study represents the first attempt to assess this potential.

There has also been rapid growth of interest in the application of biomarkers in epidemiologic and demographic research. An increasing number of multipurpose household surveys collect biological data alongside more familiar interviewer respondent information (Weinstein et al. 2008). This is particularly apparent in the growth of biomarker applications in survey work such as the DHS enterprise (Boersma 2001), which use field-friendly technologies and low-cost, quality laboratory assays. Biomarker data provide much needed information about the prevalence of a variety of health conditions, such as anemia or sexually transmitted infections (STIs), including HIV, and

may therefore improve public health knowledge and program capacity. Other applications include improved monitoring of reproductive hormones in population research (Valeggia 2007). In addition, the new field of biodemography is contributing new insights from biologic data to understanding human health and well-being especially at older ages, where it has also challenged evolutionary views of longevity (Crimmins et al. 2008; Wachter 2008). Technological advances make application of biomarkers more feasible than before, although there remain many challenges to realizing the potential of biomarkers.

The use of semen biomarkers

Biomarkers may overcome various reporting biases and inaccuracies that are increasingly questioned, and which cannot be fully overcome by behavioral approaches alone, such as in the reporting of sexual behaviors (Gallo et al. 2006; Hewett et al. 2008). Semen biomarkers further hold potential for more objective assessment of barrier contraceptive efficacy and effectiveness. This includes measurement of condom effectiveness, which remains a major public health question, as well as a topic dogged by controversy in the United States and other countries. Precise estimates of condom effectiveness against specific STIs are unavailable, not least because of reliance on self-reported sexual activity with its questionable validity and because of ethical constraints on choice of study designs.

Semen biomarkers could be useful in vaginal product development and related research settings in four more ways. These include, firstly, in early stage clinical trials so as to evaluate the safety of a new physical or chemical barrier. Second, to indicate early on whether diaphragms or condoms, for example, are effective physical barriers. For example, detection of the semen biomarker in the cervix in the absence of barrier method use and the inability to detect it following product use, would confirm that the barrier (e.g. a diaphragm or condom) worked. Third, to assess product compliance, which requires a biomarker that has low false positive and negative rates. A fourth application could be in microbicide effectiveness trials, where detection of a semen biomarker indicates the failure to use condoms consistently and/or correctly, and there is an elevated risk of STI/HIV transmission and of unintended pregnancy than if they had been. The field of HIV prevention has recently come to realize that biomarkers of semen exposure could hold tremendous potential in the near future for improved evaluation of microbicide efficacy and use (Mauck and van der Straten 2008).

There are many likely gains from such work, but there are also profound challenges. In short, there is a continued need for unique, reliable, quantifiable, easily measured, relatively inexpensive, noninvasive biomarkers. The approach to biomarker discovery adopted in this research involves open-ended discovery-based research, as adopted in technologies such as genomics, proteomics and other high-throughput approaches.

Current semen biomarkers fall into two broad types: biomarkers of either seminal plasma or of spermatozoa and other cells present in semen. The former include prostate-specific antigen (PSA), semenogelins, and acid phosphastase (AP); the latter include spermatozoa

and Y-chromosome DNA. One possible problem with using Yc DNA is that oral or manual stimulation by the male may result in detectable values in sample from the female (Zenilman et al. 2005; Ghanem et al., 2007). Spermatozoa and other cells present in semen make up less than 5% of sperm. Over 90% of human ejaculate comprises seminal plasma, a heterogenous mixture of secretions from several glands. These secretions contain a number of proteins, including PSA and potentially a number of unknown candidate biomarkers. AP, once in common use, was earlier ruled out as an accurate biomarker of semen exposure because of its low specificity and sensitivity (Lawson et al., 1998).

PSA is presently the standard biomarker of semen exposure. It has been used to assess reliability of self-reported sexual behavior and compare different interview techniques. It has also been used in studies of condom efficacy. PSA studies have successfully detected PSA in 3 settings: after unprotected intercourse (Walsh et al. 1999); in conditions of simulated intercourse, after vaginal inoculation with semen (Macaluso et al. 1999; 2003; Galvao et al., 2005); and after coitus during which a physical barrier (condom or cervical barrier) was used (Walsh et al. 1999; 2003). PSA has the advantage of being stable in dry or frozen specimens and highly standardized, effective and inexpensive tests exist for its detection. However, quantitative PSA tests are expensive and require specialized equipment usually restricted to central laboratories. Also, methods for PSA detection vary in their lower limit of detection and there is some uncertainty as to the biological significance of different levels of PSA presence. More information is needed on several dimensions of the basic PSA approach – especially on the dose-decay curve – to gain its wider acceptance in such work. In addition, second-generation biomarkers of semen exposure need to be developed. This study aimed to advance the research agenda on both fronts, with this paper focused on the second aim.

The promise of proteomics for developing new markers

Proteomic analysis is being used to identify new biomarkers and has already found multiple applications in the discovery of new diagnostic, prognostic and therapeutic targets. Common to all the many definitions of proteomics is a central concern with the study of proteins in a cell, tissue or organism. Proteomics permits both rapid identification of protein patterns in living organisms and protein characterization. Proteomic-based discovery of disease markers has included quantitative measurement of disease-specific proteins in body fluids. The proteome is a rich source of biological information because proteins are involved in almost all biological activities. However, only a small percentage of the thousands of proteins in human cells have been sequenced or identified, and the field of proteomics has not captivated public and scientific attention as much as geneomics, which has benefited from the successful sequencing of the human genome. But the promise of proteomics is every bit as great.

New technology exists for high resolution, high sensitivity detection and analysis of such proteins. The basic approach involves removal of most interfering proteins from body fluids, separating and displaying the remaining low abundance proteins as a map, and then analyzing such maps to construct a database. It is important to assess the potential

of such an approach for our purpose, although it may still be somewhat premature to expect substantial results from proteomic analysis applied to the field of sexual activity.

SELDI-TOF-MS (**S**urface **E**nhanced **L**aser **D**esorption/**I**onisation **T**ime **o**f **F**light **M**ass **S**pectrometry) is a useful, proven proteomic approach that has facilitated the discovery of disease-specific protein profiles (Grizzle WE et al. 2003; 2005.). SELDI-TOF-MS is well-suited for high-throughput protein profiling because it is able to rapidly analyse samples containing vast amounts of proteins by generating patterns that these proteins produce. It shows differences between these patterns for proteins expressed in different tissues, or in tissues during different disease states. Thus, this is a mass spectometry (MS) technique that produces a mass spectral fingerprint that can distinguish differences in protein expression levels, such as between diseased and normal samples. This has enabled use of SELDI-TOF-MS to identify, at an early stage, individuals with specific cancers (e.g. ovarian, endometrium, cervical, prostate).

Exploratory study

The present study sought to establish whether the SELDI-TOF-MS method can determine the distinct patterns of protein peaks (or protein "fingerprints") of semen and vaginal fluids that indicate intercourse. The approach involves on-chip separation of complex mixtures together with mass spectrometry. Using a laser, proteins on the chip are desorbed, causing them to be launched as ions. The time-of-flight (TOF) of the ion before detection by an electrode is a measure of the m/z (mass-to-charge ratio) value of the ion. Peptides with a larger m/z move more slowly down the flight tube and therefore have a longer TOF.

The study was nested into a larger analysis that recruited 48 couples to further elucidate the PSA dose-response decay curve. Samples were assayed for PSA and, in the case of 17 couples, additional samples were collected and assayed for analysis using the SELDI-TOF-MS system to identify seminal proteins. These samples were therefore also used to determine the feasibility of using SELDI-TOF-MS to study both semen exposure in women and -- in an additional novel aspect of the study – to assess the feasibility of studying exposure to vaginal fluid in men.

We expect to detect differences between genders corresponding to distinct profiles of vaginal and seminal fluid. Data collection and testing have been completed, and analysis is ongoing, with initial results already available. Full results will be available early in the new year and certainly well within time for the PAA meetings. It is already clear that our experiment was a success and that the SELDI-TOF-MS runs showed good results.

Material and methods

We recruited 17 couples under a protocol approved by the IRBs of the University of Alabama at Birmingham, Brookwood Medical Hospital in Birmingham, and CDC. In all, samples were collected from 10 couples at the campus clinic and 7 more at an off-campus infertility clinic. All participants completed the protocol and were included in the study.

All samples were collected following strict standard operating procedures, which were developed in advance through piloting of our sample collection, specimen storage and lab assay.

Enrolled couples collected pre- and post-coital vaginal secretions, and swab samples were also collected for analysis from the men. Men were asked to collect a post-coital semen sample from the head of their penis immediately after intercourse. These samples were tested for possible traces of exposure to women's vaginal secretions (simulating the case of condom leakage). All women and men were also asked to collect one more sample at 30 minutes post-intercourse (thereby providing another data point to compare with a PSA reading in future data analysis). The pre-coital semen sample was supplied earlier by the unused semen left over from the first phase of the study. Thus, each of the 17 couples provided six samples taken at three different intervals.

Samples were brought to the lab within 24 hours of being returned to the clinic at which the couple was enrolled. They were stored for a maximum of 24 hours in the refrigerator at -20C, prior to transfer to the laboratory deep freezer. All 17 sets of samples were aliquoted and assayed in the UAB Department of Pathology, which has a SELDI-TOF-MS system. The samples were tested using, separately, SELDI IMAC30 and CM10 chips. Samples from all couples were randomized for allocation on the SELDI-TOF-MS array, with the same randomized order used for both the IMAC30 and CM10 chips. The analysis was run using first swabs in all cases.

This paper also reports on the approaches taken to identifying and classifying spectral peaks. We first identified signal peak locations through graphical analysis before following up with our statistical analysis. The graphical output permits identification and approximate measurement of peak intensities through a comparison of mass spectroscopy profiles between samples collected at two of the three different time points. For each comparison of time points, five representative samples were chosen at random for graphical representation. The statistical analysis aimed to develop a classifier to help identify spectra as belonging to the correct group in terms of pre-intercourse, immediately post-intercourse, or 30 minutes after intercourse.

Findings

Only 1 of the 17 swabs did not look distinctly different, perhaps because we did not get a good swab. But in general, all study participants seemed reliable in obtaining samples.

The graphs indicated that for both sexes, a number of peaks were evident in the spectral profiles. For women, vaginal swab samples collected before intercourse have considerably different spectral profiles from those collected immediately after, or 30 minutes after, coitus. In particular, the pre-coital signals tend to be much flatter. Examination of the spectral profiles revealed several areas of more intense peaks post-coitus, which for the most part also appear to be strong at 30-minute after intercourse. Some peaks appeared to be inconsistent, that is, they were only apparent in several

samples at a particular time point, which may suggest that these may not be significant markers.

For male swabs, analysis of the mass spectroscopy profiles for pre-coital and post-coital profiles appeared to show discernible differences in at least four peaks. Comparisons for other time-points revealed peaks in the same vicinities that generally appeared to be of approximately similar intensity. For example, the comparison of post v. 30 minute post-coital samples shows peaks in the same spectral regions. The intensity of these signals appears to vary little, however, between the immediately post and 30 minutes-post coital samples.

Although the graphical analysis is very suggestive, it is often not possible to determine from the graphical output which peaks are less or more intense at which time points. The statistical approach developed includes an algorithm for distinguishing significant peaks ($p<0.0001$) to assist in identifying the more informative peaks that separate male proteins found in semen from proteins found in vaginal fluid.

These analyses are still ongoing, but we have already established that that there were six statistically significant peaks that discriminated well between the female pre- and post-readings. These picked up fully 100% of the pre-coital readings and 94% of the post-coital peaks. However, our initial analyses do not appear to discriminate (using the same level of statistical significance) between peaks for other time-point comparisons.

For the samples collected from men, it is clear that the comparison of pre- and immediately post-coitus samples reveals the largest number of peaks and that the SELDI-TOF-MS system can likewise reveal statistically significant differences between the signals. Again, at least six statistically significant peaks were highlighted, though we prefer a more parsimonious solution involving only four such peaks. Comparisons between the other time points are currently being studied.

Discussion

There is an ongoing need to identify new potential biomarkers, to test the validity of candidate biomarkers and to develop more reliable assays. Until recently, the search for new semen biomarkers was very slow. This is changing as researchers are turning to biomarkers for use in contraception and microbicide research trials. Researchers have made progress in overcoming the problems of objectively measuring the consistency or correctness of condom use through pioneering use of objective markers of semen exposure. Semen markers, notably (to date) PSA, can help determine if intercourse took place and if it was protected. They can reduce reliance on self-reported sexual behavior

Proteomics offers potentially major new opportunities for developing biomarkers of semen and vaginal fluid. We provide the first results of the application of such an approach which may help develop an alternative biomarker for measuring semen in vaginal fluids post-coitus, and a biomarker for measuring vaginal fluids in semen post-coitus. Multiple protein peaks were observed in the mixed signals seen from women and

men after intercourse, indicating that the application of SELDI for semen analysis works. We can see, for example, clear differences in the protein signals in the vaginal vault between the female before- and after-sex swabs. We have also collected the first known data for biomarker analysis of exposure to vaginal fluids. This method has great potential for further analysis.

There are many challenges to realizing the potential of biomarkers. There are also many technical challenges to proteomic research. The science is still not at full potential, and there is the cost and complexity of such research.

Through this study, however, we have already developed successful means of specimen collection, selection of appropriate technology platforms, and laboratory procedures, and modes of analyses. Future work will need to standardize these further, establish more reliable assays, as well as determine the validity of proposed biomarkers. Although this study lacked sufficient samples and resources for a more comprehensive assessment, the research team will be well positioned to undertake in a larger subsequent project.

References

Boerma JT, Holt E, Black R. 2001. Measurement of Biomarkers in Surveys in Developing Countries: Opportunities and Problems. *Population and Development Review*, 27(2): 303-314.

Crimmins EM et al. 2008. A comparison of biological risk factors in two populations: the United States and Japan. *Population and Development Review*, 34(3): 457-482.

Galvao LW, Oliveira LC, Diaz J, Kim DJ, Marchi N, van Dam J, Castilho RF, Chen M, Macaluso M. 2005. Effectiveness of female and male condoms in preventing exposure to semen during vaginal intercourse: a randomized trial. *Contraception*, 71(2):130-6.

Grizzle WE et al. (2003) "Serum protein expression profiling for cancer detection: Validation of a SELDI-based approach for prostate cancer," *Disease Markers*, 19: 185-195.

Grizzle WE et al. (2005) "The Early Detection Research Network surface-enhanced laser desorption and ionization prostate cancer detection study: A study in biomarker validation in genitourinary oncology," *Urologic Oncology: Seminars and Original Investigations*, 22: 337-343.

Hewett PC et al. 2008. Using sexually transmitted infection biomarkers to validate reporting of sexual behavior within a randomized, experimental evaluation of interviewing methods. *American Journal of Epidemiology*, 168(2):202-211.

Kulczycki A. Forensic Medicine. In Zhang, Yawei (Ed.). *Encyclopedia of Global Health*. Thousand Oaks, CA: Sage, 2008, pp. 682-684.

Lawson ML, Macaluso M, Bloom A, Hortin G, Hammond KR, Blackwell R. 1998. Objective biomarkers of condom failure. Sex Transm Dis. 25(8):427-32.

Macaluso M, Lawson L, Akers R, Valappil T, Hammond K, Blackwell R, Hortin G. 1999. Prostate-specific antigen in vaginal fluid as a biologic biomarker of condom failure. *Contraception*, 59(3):195-201.

Macaluso M, Lawson ML, Hortin G, Duerr A, Hammond KR, Blackwell R, Bloom A. 2003. Efficacy of the female condom as a barrier to semen during intercourse. Am J Epidemiol., 157(4):289-97.

Macaluso M et al. 2007. Efficacy of the male latex condom and of the female polyurethane condom as barriers to semen during intercourse: A randomized clinical trial. *American Journal of Epidemiology*, 166(1): 88-96.

Mauck CK and van der Straten A. 2008. Using objective markers to assess participant behavior in HIV prevention trials of vaginal microbicides. *J Acquir Immune Defic Syndr*., 49(1):64-69.

Mauck CK, Doncel GF, Biomarkers of Semen Exposure Clinical Working Group. 2007. Biomarkers of semen in the vagina: applications in clinical trials of contraception and prevention of sexually transmitted pathogens including HIV. *Contraception*, 75(6):407-419.

Valeggia CR. 2007. Taking the lab to the field: Monitoring reproductive hormones in population research. *Population and Development Review*, 33(3): 525-542.

Wachter, KW. 2008. Biodemography comes of age. *Demographic Research*, 19(40): 1501-1512.

Walsh TL, Frezieres RG, Nelson AL, Wraxall BG, Clark VA. 1999. Evaluation of prostate-specific antigen as a quantifiable indicator of condom failure in clinical trials. *Contraception*, 60(5):289-98.

Walsh TL, Frezieres RG, Peacock K, Nelson AL, Clark VA, Bernstein L, Wraxall BG. 2003. Use of prostate-specific antigen (PSA) to measure semen exposure resulting from male condom failures: implications for contraceptive efficacy and the prevention of sexually transmitted disease. *Contraception*, 67(2):139-50.

Weinstein M, Vaupel, JW and Wachter, KW (eds.). 2008. *Biosocial Surveys*. National Academy Press, Washington, D.C.

Zenilman J M, Yuenger J, Galai N, Turner, CF, Rogers SM. 2005. Polymerase chain reaction detection of Y chromosome sequences in vaginal fluid: preliminary studies of a potential biomarker for sexual behavior. *Sex Transm Dis*, 32(2): 90-94.