PAA 2010

In Search of the Holy Grail: Improving Assessments of Sexual Activity in Population Surveys through Collecting Biomarkers of Semen Exposure

Andrzej Kulczycki

It is widely agreed that better measures are needed to improve upon measures of self-reported sexual activity, which are subject to social desirability and recall bias. However, social desirability bias in particular has been a neglected problem in the family planning field (Stuart 2009), even though it has long been suspected that study subjects may frequently fail to report unprotected sex. Slowly but surely, it is increasingly recognized that to improve on the accuracy of self-reported data, measures should include biomarkers, as well as techniques for eliciting honest reporting. To date, demographers and family planning researchers have mostly attempted to refine self-administered questionnaires through such techniques as audio computer-assisted self-interviewing (ACASI) (Hewett et al. 2008), but they have shown little interest in developing biomarkers of sexual activity and adding them to population surveys.

This seems odd given the importance of demographic and health surveys and the rapid growth of interest in the application of biomarkers in demographic research. An increasing number of multipurpose household surveys collect biological data alongside more familiar interviewer respondent information (Weinstein et al. 2008). The DHS enterprise has rapidly grown biomarker applications in the past decade, using field-friendly technologies and low-cost, quality laboratory assays (Boersma 2001). Biomarker data provide much needed information about the prevalence of a variety of health conditions, such as anemia or certain sexually transmitted infections (STIs), including HIV, and may therefore improve public health knowledge and program capacity. To date, the challenge of obtaining more accurate measures of self-reported sexual behavior data has been mostly taken up by researchers in the sexually transmitted infection and HIV prevention fields (Zenilman 1995; Lawson et al. 1998; Gallo 2007; Mauck and Doncel 2007). Clearly, there is need for demographers and family planning researchers to be more engaged in this new and pressing line of research.

Biomarkers of semen exposure were first developed in forensic medicine, notably to establish if ejaculation occurred during an assault on a woman (Kulczycki 2008). More recently, such biomarkers have been proposed to help validate reports of sexual behavior and assist development of new vaginal methods of contraception and STI/HIV prevention. They may also come to be used to evaluate interventions in areas or population groups where the incidence of particular outcomes is low or where behavior change is needed, such as adolescent sexual health intervention or for the prevention of intimate-partner violence and HIV (Ross et al; 2007; Pronyk et al. 2006). In particular, detection of seminal biomarkers in vaginal fluid provides objective evidence of a woman's recent exposure to semen. Several biomarkers have been proposed, though each has its disadvantages (Lawson et al.1998; Mauck and Doncel 2007).

This paper has several aims. First, it briefly reviews the current state of the field, including problems with existing approaches. It discusses an initial attempt to come up

with a rapid test for detecting semen exposure in women. It proceeds to report on an alternative approach spearheaded by the author, which is the first known application of proteomic analysis to assess the mixture of semen and vaginal fluid, opning up the possibility of new biomarker discovery. The rationale for this new approach as a way forward toward improving measures of sexual activity is described. Initial results are very encouraging, but much more research is needed. In addition to reporting findings of this exploratory assessment, this paper will describes the procedures used to identify the protein components transmitted between opposite-sex sexual partners. The potential of a proteomic approach is emphasized. 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. However, it may still be somewhat premature to expect substantial results from proteomic analysis applied to the field of sexual activity.

Present use and new applications of semen biomarkers

Semen biomarkers offer the potential for more objective assessment of sexual activity over traditional self-reported measures. For example, they could assist in vaginal product development, such as in early stage clinical trials so as to evaluate the safety of a new physical or chemical barriers; in the assessment of product compliance, which requires a biomarker that has low false positive and negative rates; and 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 (Mauck and van der Straten 2008).

Prostate-specific antigen (PSA) is presently the standard biomarker of semen exposure. It can be measured in vaginal samples collected after intercourse and has been used widely in rape investigations. High PSA concentrations (100–10000 ng PSA/mL vaginal swab eluate) are detectable immediately after exposure, and levels return to baseline (<1.0 ng/mL) within 24 to 48 hours PSA has been validated as a reliable marker of semen exposure in studies of vaginal specimens obtained after unprotected intercourse (Walsh et al. 1999); after vaginal insemination with different volumes of 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). It has been used as a proxy measure for condom efficacy (Macaluso et al. 2007) as well as to assess reliability of self-reported sexual behavior and compare different interview techniques ((Minnis et al. 2009). For example, it been used to corroborate self-reports of coitus and barrier contraceptive use, demonstrating that self-reports of sexual behavior cannot be assumed to be valid measures (Gallo et al. 2003; 2007a and b; Zenilman et al. 1995). PSA can be detected in urine but levels are low, so that the likelihood of contamination by fluids from other parts of the body is small (Walsh et al., 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.

Although the quantitative IMx PSA assay is the 'gold standard' for PSA detection, rapid immunochromatographic strip tests for detection of PSA are available commercially. They are completely portable, easy to use, relatively inexpensive, and require no instrumentation. Relatively good performance of the rapid ABAcard p30 test (Abacus Diagnostics, West Hills, CA, USA) compared to a quantitative assay (IMx PSA, Abbott Laboratories, Abbott Park, IL, USA) for detection of PSA in vaginal swabs has recently been reported in two specimen sets (Hobbs et al. 2009). The ABAcard test performed well, but not perfectly, and requires careful consideration of the implications for false positive and false negative test results (in particular, high concentrations of PSA can give false negative results).

Nonetheless, this is the first, simple and relatively inexpensive test for identifying a marker of recent semen exposure in vaginal swabs. Further testing is planned in a forthcoming randomized trial of female STD clinic attendees. The rapid test results can be interpreted semi-quantitatively to approximate different levels of likely semen exposure. Rapid PSA detection requires no instrumentation and can be performed easily and economically. Also, having rapid PSA results available immediately following the main part of the interview may permit ongoing interview probing of any discrepancies between the objective marker of recent semen exposure and self-reported behaviors. However, regarding use of PSA as a marker of semen exposure in general, there is some uncertainty as to the biological significance of different levels of PSA presence. Another proposed semen biomarker, Y-chromosome DNA, is detectable for up to 2 weeks post-coitus, but oral or manual stimulation by the male may result in detectable values in samples from the female (Zenilman et al. 2005; Ghanem et al., 2007).

Proteomic analysis and the promise of new markers

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. Open-ended discovery-based research, as adopted in technologies such as genomics, proteomics and other high-throughput approaches, seems a particularly fruitful approach to biomarker discovery and is adopted in the research here. Proteomics permits both rapid identification of protein patterns in living organisms and protein characterization. The proteome is a rich source of biological information because proteins are involved in almost all biological activities. Proteomic-based discovery of disease markers has included quantitative measurement of disease-specific proteins in body fluids. 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.

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. SELDI-TOF-MS (Surface Enhanced Laser Desorption/Ionisation Time of Flight Mass Spectrometry) is a useful, proven proteomic approach that has facilitated the discovery of disease-specific protein profiles.

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 mass spectometry (MS) technique 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), as well as applications in infectious and cardiovascular diseases (Grizzle et al. 2003 and 2005; Seibert et al. 2005; Hodgetts et al. 2007; Cieniewski-Bernard et al. 2008).

An exploratory proteomic analysis of sexual activity

The present CDC-funded 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. The study was nested in a larger study 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 hypothesized that this technology would permit detection of differences between genders corresponding to distinct profiles of vaginal and seminal fluid. Ongoing analysis has already yielded significant findings and it is already clear that the SELDI-TOF-MS runs showed good results.

Of the 17 couples recruited, all participants completed the protocol and were included in the study. All samples were collected following strict standard operating procedures, 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 to test for possible traces of exposure to women's vaginal secretions. In sum, 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. Vaginal swab specimen preparation methods were developed and all 17 sets of samples were aliquoted and assayed. The samples were tested using, separately, SELDI IMAC30 and CM10 chips (with the former proving more useful). Samples from all couples were randomized for allocation on the SELDI-TOF-MS array, with the same randomized order used for both sets of chips. The analysis was run using first swabs in all cases.

SELDI-TOF-MS data consist of matched-pairs of m/z (mass-to-charge ratio) and signal intensity values (Dijsktra et al. 2007). The signal intensity is a relative value without unit and is associated with the amount of protein components in the sample. Intensity values can be plotted against m/z values to obtain a raw mass spectrum to show the peaks that

represent the masses and the amounts of the sample components. Consequently, an m/z value can be used to locate a peak and the corresponding intensity, i.e., the height of the peak, can be used to measure the amount of the sample components. A sample component may be identified as a protein component based on a certain criterion.

This paper will report briefly on the procedures taken to identifying and classifying spectral peaks, as well as findings. Signal peak locations were first identified through graphical analysis and then more precisely through 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. 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. Overall, 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.

The graphical output presents a good overview, but does not permit precise determination of 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. For the samples collected from both the women and their partners, 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. On average, 45% of the post-coital protein components found in the vaginal specimens were transmitted from their male partners, 43% were undefined, 6% were completely carried over by the female

herself, and the rest were found to have been carried over by both partners. The 'undefined' values may represent material produced by the woman's sexual response or catabolic products from either the woman or the man.

Further points

Researchers have often despaired over whether sufficiently valid self-reported behavioral data can be obtained in pregnancy, STI and HIV prevention research that depend on reports of sexual activity. There is a high level of discrepant reporting found in multiple studies, regardless of interview mode. Consistent and correct condom use is often over-reported, confounding information that could be important to STI/HIV prevention programming. Recently, biomarkers of semen exposure have enabled progress to be made in objectively measuring the consistency or correctness of condom use. Semen biomarkers can reduce reliance on self-reported sexual behavior. In particular, use of PSA can help determine if intercourse took place and if it was protected. A simple, rapid, and inexpensive test for PSA has recently been tested that would facilitate recent semen biomarker evaluation. Initial reports suggest that it may be reliable for the detection of semen in vaginal secretions in various research settings, including those in which laboratory facilities are not available.

However, this does not imply that we are close to resolving the seemingly intractable problems with behavioral assessments of sexual activity. We are still not close to having a biomarker that could measure contraceptive prevalence rates, and there are problems inherent even in using a biomarker such as PSA to detect a behavior such as condom use at last intercourse. PSA detects only recent semen exposure, but it too may be imperfect for such purpose (e.g. in detecting user errors and mechanical failures associated with condom use), and other measures may assess condom-use behaviors more thoroughly. Pregnancies and STIs may also result from unreported unprotected intercourse without PSA data. Because PSA has a rapid decay curve, it is generally not detectable after 24 hours and can therefore underestimate the overall prevalence of discrepant results (and possible misreporting) because some PSA results may be false negatives. The threshold level of PSA at which there is a risk of pregnancy or STI acquisition is still unknown, although for certain applications, this may not matter. There is clear need for more objective biologic measures of sexual activity and product use that can be employed in research that requires valid assessments of sexual activity.

Proteomic analysis offers potentially major new opportunities for developing biomarkers of semen and vaginal fluid. This study provides 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, as well as offer the first biomarker for detecting and 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. Clear differences are evident in the protein signals in the vaginal vault between the female before- and after-sex swabs. Additionally, modes of specimen collection, selection of appropriate technology platforms, and laboratory procedures, and modes of analyses, have been identified. Future work will need to

standardize these further, establish more reliable assays, as well as determine the validity of proposed biomarkers. Like PSA and other currently available biomarkers, SELDI detects only recent semen exposure, but the proteomic approach developed has great potential for further analysis.

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 self-reporting of sexual activity and behaviors. However, there remain numerous challenges to realizing the potential of biomarkers. These include many technical challenges to proteomic research. Although this study lacked sufficient samples and resources for a more comprehensive assessment, it has highlighted the feasibility of a proteomic approach, which may one day offer more precise measurement of recent sexual activity. But the quest for the holy grail regarding measurement of sexual activity remains elusive. Biomarkers of sexual activity are still not ready to be integrated into a survey, and we may never be able to dispense with measuring self-reported sexual activity. On the other hand, it is possible to see that used with caution, they may in the not too distant future, be a handy addition to population surveys.

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.

Cieniewski-Bernard C, Acosta A, Dubois E, Lamblin N, Beseme O, Chwastyniak M, Amouyel P, Bauters C, Pinet F. 2008. Proteomic analysis in cardiovascular diseases. *Clin Exp Pharmacol Physiol*. 35(4):362-6.

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.

Gallo MF, F.M. Behets and M.J. Steiner *et al.*, Prostate-specific antigen to ascertain reliability of self-reported coital exposure to semen, *Sex Transm Dis* **33** (2006), pp. 476–479.

Gallo MF, F.M. Behets and M.J. Steiner *et al.*, Validity of self-reported 'safe sex' among female sex workers in Mombasa, Kenya-PSA analysis, *Int J STD AIDS* **18** (2007), pp. 33–38.

Gallo MF, M.J. Steiner and L. Warner *et al.*, Self-reported condom use is associated with reduced risk of chlamydia, gonorrhea, and trichomoniasis, *Sex Transm Dis* **34** (2007), pp. 829–833.

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.

Hodgetts A, Levin M, Kroll JS, Langford PR (2007). Biomarker discovery in infectious diseases using SELDI. *Future Microbiol* 2:35-49.

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.

Pronyk P, Hargreaves JR, Kim JC, et al. Effect of a structural intervention for the prevention of intimate-partner violence and HIV in rural South Africa: a cluster randomised trial. *Lancet*. 2006;368(9551):1973–1983.

Ross DA, Changalucha J, Obasi AI, et al. Biological and behavioural impact of an adolescent sexual health intervention in Tanzania: a community-randomized trial. *AIDS.* 2007; 21(14): 1943–1955.

Seibert V, Ebert MP, Buschmann T (2005). Advances in clinical cancer proteomics: SELDI-ToF-mass spectrometry and biomarker discovery. *Brief Funct Genomic Proteomic* 4(1):16-26.

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 JM, Weisman CS, Rompalo AM et al.. 1995. Condom use to prevent incident STDs: the validity of self-reported condom use, *Sex Transm Dis* 22: 15–21.

Zenilman JM, 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.