Clinical Chemistry 42:6
853–857 (1996)

Ultrasensitive prostate-specific antigen assays and their clinical application

ELEFTHERIOS P. DIAMANDIS,* HE YU, and DIMITRIOS N. MELEGOS

The contributions of Stamey in the area of prostate cancer diagnosis, monitoring, and treatment and in the prostate-specific antigen (PSA) literature are numerous and noteworthy.[1] His accompanying Opinion article *[1]* deals with aspects of analytical performance as well as clinical interpretations of PSA assays and raises concern that clinicians may be misled by recently published data. We agree with many of his concerns and disagree with others. In areas of disagreement, we would like to offer alternative views that are based on extensive published and unpublished data from our own research group.

Analytical and Functional Sensitivity of PSA Assays

During the last few years both clinical investigators and commercial concerns have made efforts to produce PSA assays with improved detection limits (e.g., *[2–11]*). The well-accepted rationale behind more-sensitive PSA assays is that cancer relapse will be detected earlier if patients are monitored with more-sensitive methods. An ultrasensitive PSA assay is characterized by its lowest limit of detection (LLD), which, by definition, is the least amount of analyte that can be detected with a predetermined confidence, usually 95% or 99%. The LLD commonly is calculated by analyzing the zero calibrator and the lowest-concentration calibrator of the assay many times (e.g., 12–20 replicates). From these data, the SD of the zero signal is calculated. The LLD is the analyte concentration that corresponds to the zero signal + 2 SD (for 95% confidence) or + 3 SD (for 99% confidence), calculated from the slope of the calibration curve between the zero and the first calibrators. From these calculations, one can determine the factors that affect the LLD. In a "sandwich-type" assay with variable zero signal, variable precision of the zero signal, and variable signal of the lowest-concentration calibrator (Table 1), the best LLD (case D, Table 1) is associated with (*a*) lower zero (background) signal, (*b*) greater precision of the zero signal, and (*c*) the calibration curve with the greatest slope (i.e., sensitivity).

Several properties contribute to these three factors. The zero signal depends on the detection technique. For example, with fluorescence, the background signal includes the intrinsic fluorescence of the measuring cuvette and solvents, scatter from the excitation light *[12]*, and fluorescence from nonspecific binding of the labeled reagent. This nonspecific fluorescence in turn depends on the nature and the amount of labeled reagent added, its diluent, the incubation time and temperature, the nature and method of blocking of the solid-phase used, the efficiency of washing, and so forth. With an enzyme label, nonenzymatic hydrolysis and intrinsic fluorescence of the substrate also contribute. Lot-to-lot variability of the reagents contributes to the fluctuation of the background signal, as do differences in washing efficiencies of microtiter wells. Added to this are the pipetting precision, intertechnologist variation, variabilities of plastic capacity, and homogeneity of coating. The slope of the calibration curve depends in part on the affinities of the antibodies used; lower-affinity antibodies allow loss of signal during washing. The method of detection (e.g., absorbance, fluorescence, chemiluminescence) determines how much signal is generated.

When all the potential variables are considered, it is not surprising that different investigators report different LLD values, even when evaluating the same commercially available method. Eliminations of outliers may further affect the final value of an LLD and allow the calculation of better LLDs.

Should LLD be calculated at all? The answer is yes, but LLD should not form the basis for clinical decisions. LLDs serve as rough estimates of the capabilities of a single method, and as indicators of the quality of the antibodies used and the power of the detection technique. In other words, the LLD helps to define the *potential* of a method for certain clinical use. By also defining a lower range of concentrations that cannot reliably be distinguished from zero, it thus defines the lower limit of the reporting range of the method.

Which approach should be used to guide clinical decisions? Vessella et al. *[5]* introduced the term "biological detection limit" (BDL), which they determined by adding the LLD and 2 SD of the patients' sample interassay imprecision data generated

Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, 600 University Ave., Toronto, Ontario, Canada M5G 1X5 (address for correspondence), and Department of Clinical Biochemistry, University of Toronto, 100 College St., Toronto, Ontario, Canada M5G 1L5.

*Author for correspondence. Fax 416-586-8628; e-mail epd@eric.on.ca.

[1] Nonstandard abbreviations: PSA, prostate-specific antigen; BSA, bovine serum albumin; LLD, lowest limit of detection; BDL, biological detection limit; FS, functional sensitivity; AFU, arbitrary fluorescence units; and RCDL, residual cancer detection limit.

Received November 27, 1995; accepted March 7, 1996.

Table 1. Hypothetical PSA assay and calculated lowest limits of detection (LLD).

Case	Zero calibrator, mean (SD)[a]	CV, %	10 μg/L calibrator[a]	LLD, μg/L (95% confidence)[b]
A	1000 (100)	10	2000	2
B	1000 (200)	20	2000	4
C	2000 (200)	10	3000	4
D	1000 (100)	10	3000	1

[a] Signals for mean (SD) zero calibrator and lowest-concentration calibrator (10 μg/L) are in arbitrary units.

[b] The LLD is calculated by multiplying 2 SD (for 95% confidence) for the zero calibrator by the concentration of the lowest-concentration calibrator (10 μg/L in this example) and dividing by the difference of the signal between the lowest-concentration and zero calibrators. For case A, LLD = (2 × 100) × 10 μg/L/(2000 − 1000) = 2 μg/L.

at concentrations close to the detection limit. For example, they calculated the LLD of the Tandem-R® assay (Hybritech, San Diego, CA) to be 0.09 μg/L. The interassay SD for samples in the PSA concentration range of 0.24–0.36 μg/L was found to be 0.05 μg/L. They then calculated the BDL to be 0.09 + 2 × 0.05 = 0.19 μg/L (rounded to 0.20 μg/L). For the IMx® assay (Abbott Labs., Abbott Park, IL), the LLD was 0.03 μg/L and the interassay SD at PSA concentrations of 0.04–0.06 μg/L was 0.013 μg/L. Thus, the BDL for the IMx assay was 0.03 + 2 × 0.013, or ~0.06 μg/L.

Klee et al. *[6]* preferred "functional sensitivity" (FS), which is defined as the lowest concentration of PSA that can be assessed with an interassay imprecision of 20%. Because this proposed term, like BDL, is based on interassay imprecision data, we have compared the BDL and FS. Vessella et al. reported a BDL of 0.06 μg/L for the IMx; their interassay precision profile at PSA concentrations of 0.04, 0.05, 0.06, and 0.10 μg/L were 24%, 20%, 18%, and 8%, respectively. As defined by Klee et al., the FS of Vessella et al. would have been 0.05 ± 0.01 μg/L, identical to their BDL. The data from Klee et al. support a FS of 0.06–0.13 μg/L *[6]*.

These data allow us to propose the following recommendations. (*a*) The LLD should not be used to make clinical decisions. (*b*) Clinically useful cutoff points are usually 2–4 times higher than the LLD. (*c*) The clinical decision limits for PSA (whether called BDL or FS or otherwise) should be based on interassay imprecision profiles established with patients' sera at concentrations 2–4 times the LLDs; in general, interassay imprecision at the clinically useful cutoff point should be ≤20%. (*d*) Further, neither BDL nor FS should be considered "gold standards" for making clinical decisions. Specifically, for PSA applications, we recommend that, for diagnosis of patient relapse during monitoring, a set of simple criteria should be formulated incorporating further conservative measures (see below).

We emphasize three caveats in the clinical applications of BDL and FS. Precision profiles should be established with the same protocols as the ones used for patient monitoring. For example, if patients' sera are run in singletons, the precision profile should also be established in singletons, with instrument or method recalibrations made at the frequency used for patients' samples. Secondly, the precision data should be monitored over long periods because BDL or FS may change due to lot-to-lot reagent variation, instrumental malfunction, or other variables. Finally, the clinical value of any BDL or FS should be tested with clinical investigations. Stamey and Prestigiacomo have pointed out one assay with low BDL or FS that could not detect relapse of prostate cancer as efficiently as other ultrasensitive assays *[8]*.

The criteria outlined above may be modified in research settings in which other definitions are more appropriate. For example, we recently measured PSA in sequential sera from postprostatectomy patients to establish criteria for biochemical evidence of relapse of prostate cancer. All sera from the same patient were analyzed in the same assay run, in triplicate. In this setting, we adopted the following principles for establishing a relapse: (*a*) Only values above the FS or BLD were considered meaningful; (*b*) intraindividual variation for PSA ≤30% was allowable; and (*c*) PSA variation was determined as intraassay (not interassay) imprecision. Because clinical experience with highly sensitive PSA assays is limited, we have also included further conservative measures; e.g., a biochemical relapse will be considered to be established only after two consecutive sequential samples show a PSA increase that at least doubles the initial PSA concentration. These criteria must be evaluated clinically; however, a similar set of interpretative criteria were recently shown to work well in a small group of patients whose clinical outcomes were known *[13]*.

PSA Calibrator Matrix and Quality-Control Sera

Stamey suggests *[1]* that LLD values based only on assays with bovine serum albumin (BSA)-based reagents bear no relationship to relevant clinical values in humans. We believe this is an overstatement. BSA-based calibrators are now used not only for PSA but also for many other analytes and are well established. The advantages of BSA matrix are well-recognized: e.g., no infectivity, low cost, excellent reproducibility and consistency, absence of analyte, and extended analyte stability. For PSA assays, other alternatives could be animal sera, female sera, and male sera devoid of PSA. Animal sera are used already and the same argument as is made against BSA can be used against them. Female sera are not always devoid of PSA, as we have shown already *[14]*, and sera from pregnant women have an even higher median PSA (unpublished). Moreover, postprostatectomy serum has, on average, a higher PSA concentration than serum from nonpregnant women *[14]*.

The role of a zero calibrator is to set an average assay signal that corresponds to no analyte presence. In our assay system, a 60 g/L BSA matrix devoid of PSA generates a signal of ~1000 arbitrary fluorescence units (AFU), and a PSA calibrator of 0.001 μg/L in the same matrix generates a signal of ~1500 AFU. If the AFU of the zero calibrator were well below the real average AFU for a sample containing zero PSA—female serum devoid of PSA, or male serum from postprostatectomy patients also devoid of PSA—then we would expect that all or nearly all female sera and postprostatectomy sera should read positive for PSA because of this zero shift. By contrast, of 1161 samples from

nonpregnant women analyzed with our older PSA assay *[2, 14]*, >80% had apparent PSA less than the BLD (0.01 μg/L). In a newer series of 212 female sera analyzed with our latest and more-sensitive assay (BLD ~0.001 μg/L), also calibrated with 60 g/L BSA-based samples, the median concentration found was 0.002 μg/L and 32% of the sera had PSA ≤0.001 μg/L. Thus, the notion that the 60 g/L BSA solution sets the zero signal too low is not supported by the experimental data.

Contrary to the concern that the zero signal is too low, the 1000 AFU (by definition) for the zero calibrator in 60 g/L BSA may have been set at a higher point than the actual clinical samples in our assay, thus causing an apparent negative bias in samples with real PSA of ~0.001–0.002 μg/L. Human serum devoid of analyte is more effective than 60 g/L BSA reagent in reducing nonspecific binding of the detection antibody in heterogeneous, sandwich-type, one-step assays. Among 212 female sera from nonpregnant females, measured PSA was ≤0.002 μg/L in 50%, between 0.001 and 0.0019 μg/L in 17%, between 0.000 and 0.001 μg/L in 14%, and <0.000 in 18%—all concentrations derived by using 60 g/L BSA-based calibrators. If indeed the signal of the 60 g/L BSA solution was set at a higher value than an equivalent serum-based calibrator, this would be an advantage for monitoring postprostatectomy male patients because any detected increase in subsequent sera would be based on even more conservative criteria than the ones already mentioned and published *[13]*.

Is the immunoreactive PSA detected at concentrations (e.g.) ≤0.010 μg/L in female sera or sera from postprostatectomy patients indeed PSA and not cross-reacting substances or random variations? We now know that PSA can be produced by female tissues other than the periurethral glands. Normal, benign, and malignant breast tissues produce PSA *[15–21]*. PSA is secreted into the lumen of the ducts of the mammary glands and can be detected in breast milk *[19]* and discharge fluid (≤15 000 μg/L; unpublished data). It should thus not be surprising that some PSA from breast tissue, as with prostate tissue, may escape into the female blood circulation. Other potential sources of PSA in women may be the endometrium *[22]* or salivary glands *[23]*. During pregnancy, serum PSA increases in amniotic fluid according to gestational age (*[24]* and Melegos et al., ms. submitted). The pattern of PSA changes in maternal serum during gestation is statistically significant and cannot be attributed to random fluctuations. We also found nonrandom associations between serum PSA in nonpregnant females and their ages, and we provided preliminary evidence that the immunoreactive species in normal female serum is PSA bound to α_1-antichymotrypsin *[14]*. In male serum, the situation is clearer. In sequential patients' sera postradical prostatectomy, the PSA concentrations, although undetectable after surgery (<0.001 μg/L) in many patients, became detectable (0.001–0.005 μg/L) and then increased steadily, following an exponential model *[13, 25]*. Detailed descriptions of these data will be published elsewhere.

Ideally, quality-control specimens for PSA would be developed in authentic male serum, but this is not an absolute necessity. Male quality-control sera can be found from prostatectomy patients and may have serum PSA anywhere between 0.001 and 0.10 μg/L, but commercial availability would be very limited. PSA-supplemented female sera could serve as useful PSA controls to monitor precision even if the PSA subfractions did not match the subfractions in male serum. Animal serum, which is the most readily available, should also be evaluated carefully.

We believe that new PSA assays, designed to monitor patients after radical prostatectomy or for use in other applications such as measurement of PSA in breast tumor cytosolic extracts and female serum during pregnancy, should have FS values between 0.001 and 0.002 μg/L (day-to-day CV <20%). These assays should be monitored for imprecision at concentrations of ~0.003, 0.010, 0.030, and 0.100 μg/L and should include calibrators at 0, 0.002, 0.005, 0.020, 0.100, 0.500, and 2.00 μg/L.

Residual Cancer Detection Limit (RCDL)

Stamey suggests that, were we to follow the recommendations of Vessella et al., the BDL we calculated in our previous report *[2]* would have been much higher than 0.010 μg/L. This is not accurate. Our LLD, as calculated with the method described, was 0.002 μg/L. Our within-run CV at 0.016 μg/L, for a human serum sample, was 21.4%. We have thus conservatively calculated our BDL to be 0.010 μg/L. In our precision studies (Table 1 of reference *2*), we found little difference between within-run and day-to-day imprecisions, similar to the findings of others *[5]*. Consequently, our BDL would have probably been a little but not substantially higher than reported if we were to use day-to-day imprecision data instead of within-day imprecision data. Since publication of that report, we have developed another PSA assay that can detect PSA at 0.001 μg/L with within-run imprecision of <20% at PSA >0.002 μg/L *[25]*.

We disagree with the determination and use of RCDL as described by Stamey. First, he describes such determination as being simpler than the determination of BDL or FS. This does not seem to be the case. For FS studies, one needs a few sera from postradical prostatectomy subjects who were found to have PSA <0.10 μg/L by conventional PSA analysis. We have shown that, among these sera, ~50% have PSA of 0.005–0.10 μg/L; the rest have PSA of 0–0.005 μg/L. For RCDL determination, one needs sera from extremely well-defined patients: e.g., follow-up >5 years, Gleason score <5, prostatic tissue examination with 3-mm step-sections, and tumor size <0.5 mL. Such sera are not generally available, and unreliable selection will lead to wrong calculations.

The RCDL is based on the unproven premises that a patient cured of prostate cancer should not have any PSA in his serum and that, if a sensitive method detects something, this must be due to random assay variation. In our opinion, the possibility of extraprostatic sources of PSA cannot be excluded.

We anticipate that patients will be encountered with postsurgical PSA well below the RCDL of a method (as calculated by Stamey) who will show consistent and consecutive increases of PSA, suggestive of biochemical relapse, before crossing the RCDL. In one such patient, we found PSA to increase from <0.001 μg/L to 0.0028, 0.009, 0.032, 0.073, and 0.118 μg/L over a 19-month period. Other patients may have postsurgical

PSA above the RCDL but without consistent increases of PSA over time, suggestive of PSA originating from nontumor sources. Even by Stamey's definition of 99% confidence levels of RCDL, a few patients with no relapse will fall into the category of relapsed patients.

What, then, is our opinion? On the basis of limited published data and more-detailed unpublished data on >300 patients who underwent radical prostatectomy and then were monitored for at least 2 years with a new ultrasensitive assay that has a FS of about 0.001 μg/L, we came to the following conclusions: (*a*) PSA >0.10 μg/L at many weeks postradical prostatectomy is indicative of residual disease. This PSA concentration has also been proposed previously by the Stamey group as the RCDL *[26]* and by Vessella et al. as a clinical threshold *[5]*. (*b*) For patients with PSA <0.1 μg/L postsurgery, no PSA concentrations, even those below the RCDL, can assure freedom from disease. In our series (unpublished) we found many patients whose PSA decreased to <0.001 μg/L postsurgery and then relapsed. In contrast to the belief that a low postsurgical PSA (e.g., <0.01 μg/L) indicates a less-aggressive form of the disease *[10]*, we found patients with postsurgical PSA <0.001 μg/L who relapsed with short PSA-doubling times (<100 days), indicative of a rapid tumor-doubling time. We thus propose that the best way to detect biochemical relapse of prostate cancer patients after radical prostatectomy at the earliest possible time is by serially monitoring PSA and comparing any new PSA value with the previous PSA values of the same patient, not with the PSA values of any other group of patients (e.g., those who apparently were cured of cancer). If resources were not the issue, we would further propose to analyze each new patient's serum with the previous two sera in the same run, preferably in duplicate. This practice would minimize variability between runs and facilitate detection of PSA changes over time.

How sensitive must PSA assays be for relapse to be detected as early as possible? Our data suggest that PSA in ~60% of patients decreased to <0.005 μg/L postsurgery and in ~26% decreased to 0.001 μg/L. Thus, assays with FS around 0.001–0.002 μg/L should be desirable. Unquestionably, ultrasensitive PSA assays can detect relapse much earlier than conventional assays *[8, 11, 13]*. However, the detection of biochemical relapse at the earliest possible time depends not only on PSA assay sensitivity but also on other, fundamentally important factors, including:

(*a*) The frequency of serum sampling. More frequent sampling may facilitate earlier detection of relapse.

(*b*) The strategy of analyzing samples. Analyzing replicates and including in the same run three sequential sera (the current and previous two) will facilitate identification of PSA changes through minimization of assay variability.

(*c*) The interpretive criteria. These should be established on the basis of clinical outcomes *[13]* and should be conservative enough to avoid false positives.

(*d*) The doubling time of the tumor *[13]*. The time difference between relapse detection by an ultrasensitive assay and a less-sensitive assay will be shorter with rapidly growing tumors and longer with slow-growing tumors.

Should clinicians aggressively seek identification of biochemical cancer relapse of asymptomatic patients postradical prostatectomy? We answer yes, *if* something can be done about it; otherwise, the cancer patient will only suffer unfavorable psychological pressure. The current premise is that very small relapsing tumors are more sensitive to chemotherapy or radiation therapy than are large tumors. Small tumors may be amenable to treatment with lower doses of drugs. Also, early treatment of relapse may prevent or delay dissemination of the disease. Given that no effective therapies for prostate cancer relapse are as yet documented, we believe that the full potential of ultrasensitive PSA assays cannot, at present, be realized. Clearly, we need not only more-effective therapies but also clinical trials that include methods for early detection of relapse and institution of therapy to assess outcomes. We agree with Vessella et al. *[11]* that the current ultrasensitive assays should be used only in research settings until more clinical data become available.

References

1. Stamey TA. Lower limits of detection, biological detection limits, functional sensitivity, or residual cancer detection limit? Sensitivity reports on PSA assays mislead clinicians. Clin Chem 1996;42:849–52.
2. Yu H, Diamandis EP. Ultrasensitive time-resolved immunofluorometric assay of prostate-specific antigen in serum and preliminary clinical studies. Clin Chem 1993;39:2108–14.
3. Liedtke RJ, Kroon G, Batjer JD. Modified assay of prostate-specific antigen with a detection limit <0.01 μg/L. Clin Chem 1993;39:2150–4.
4. Khosravi MJ, Papanastasiou-Diamandi A, Mistry J. An ultrasensitive immunoassay for prostate specific antigen based on conventional colorimetric detection. Clin Biochem 1995;28:407–14.
5. Vessella RL, Noteboom J, Lange PH. Evaluation of the Abbott IMx® automated immunoassay of prostate-specific antigen. Clin Chem 1992;38:2044–54.
6. Klee GG, Dodge LA, Zincke H, Oesterling JE. Measurement of serum prostate specific antigen using IMx prostate specific antigen assay. J Urol 1994;151:94–8.
7. Stamey TA, Graves HCB, Wehner N, Ferrari M, Freiha FS. Early detection of residual prostate cancer after radical prostatectomy by an ultrasensitive assay for prostate specific antigen. J Urol 1993;149:787–92.
8. Prestigiacomo AF, Stamey TA. A comparison of 4 ultrasensitive prostate specific antigen assays for early detection or residual cancer after radical prostatectomy. J Urol 1994;152:1515–9.
9. Graves HCB, Wehner N, Stamey TA. Ultrasensitive radioimmunoassay of prostate-specific antigen. Clin Chem 1992;38:735–42.
10. Vessella RL. Trends in immunoassays of prostate-specific antigen: serum complexes and ultrasensitivity [Editorial]. Clin Chem 1993;39:2035–9.
11. Vessella RL, Noteboom J, Lange PH, Rittenhouse H, Liu R, Kusnierz E. Ultrasensitive PSA immunoassays. How low do we need to go? J Urol 1993;149:484A.
12. Diamandis EP. Immunoassays with time-resolved fluorescence spectroscopy. Principles and applications. Clin Biochem 1988;21:139–50.
13. Yu H, Diamandis EP, Prestigiacomo AF, Stamey TA. Ultrasensitive assay of prostate-specific antigen used for early detection of prostate cancer relapse and estimation of tumor doubling time after radical prostatectomy. Clin Chem 1995;41:430–4.
14. Yu H, Diamandis EP. Measurement of serum prostate specific antigen levels in females and in prostatectomized males with an

ultrasensitive immunoassay technique. J Urol 1995;153: 1004–8.

15. Diamandis EP, Yu H. New biological functions of prostate specific antigen? J Clin Endocrinol Metab 1995;80:1515–7.

16. Yu H, Diamandis EP, Sutherland DJA. Immunoreactive prostate specific antigen levels in female and male breast tumors and its association with steroid hormone receptors and patient age. Clin Biochem 1994;27:75–9.

17. Diamandis EP, Yu H, Sutherland DJA. Detection of prostate specific antigen immunoreactivity in breast tumors. Breast Cancer Res Treat 1994;32:301–10.

18. Monne M, Croce CM, Yu H, Diamandis EP. Molecular characterization of prostate specific antigen mRNA expressed in breast tumors. Cancer Res 1994;54:6344–7.

19. Yu H, Diamandis EP. Prostate-specific antigen in milk of lactating women. Clin Chem 1995;41:54–8.

20. Yu H, Diamandis EP, Monne M, Croce CM. Oral contraceptive-induced expression of prostate specific antigen in the female breast. J Biol Chem 1995;270:6615–8.

21. Yu H, Giai M, Diamandis EP, Katsaros D, Sutherland DJA, Levesque MA, et al. Prostate specific antigen is a new favorable prognostic indicator for women with breast cancer. Cancer Res 1995;55:2104–10.

22. Clements A, Mukhtar A. Glandular kallikreins and prostate specific antigen are expressed in the human endometrium. J Clin Endocrinol Metab 1994;78:1536–9.

23. Van Krieken JH. Prostate marker immunoreactivity in salivary gland neoplasms. A rare pitfall in immunohistochemistry. Am J Surg Pathol 1993;17:410–4.

24. Yu H, Diamandis EP. Prostate-specific antigen immunoreactivity in amniotic fluid. Clin Chem 1995;41:204–10.

25. Ferguson RA, Yu H, Kalyvas M, Zammit S, Diamandis EP. Ultrasensitive detection of prostate-specific antigen by a new time-resolved immunofluorometric assay and the Immulite immunochemiluminescent third-generation assay: potential applications in prostate and breast cancers. Clin Chem 1996;42:675–84.

26. Graves HCB. Prostate-specific antigen comes of age in diagnosis and management of prostate cancer [Editorial]. Clin Chem 1992; 38:1930–2.