



# STATE OF NEW YORK DEPARTMENT OF HEALTH

Wadsworth Center The Governor Nelson A. Rockefeller Empire State Plaza P.O. Box 509 Albany, New York 12201-0509

Richard F. Daines, M.D.  
*Commissioner*

Wendy E. Saunders  
*Chief of Staff*

October 26, 2007

## **New York State Tumor Marker Proficiency Test 9/2007 Evaluation <sup>1</sup>**

Dear Laboratory Director,

Attached is a summary and evaluation of the New York State Proficiency Testing mail-out from September 11, 2007 for Tumor Markers AFP, CA125, CA15-3, CA27.29 and CA19-9, CEA, PSA, free PSA and complexed PSA.

Please note that questions regarding the electronic proficiency testing reporting system (EPTRS) account application process and the entry and submission of proficiency test results can be directed to [clepeptrs@health.state.ny.us](mailto:clepeptrs@health.state.ny.us), or directly to Kathi Wagner at (518) 402-4266 or by e-mail at [klw05@health.state.ny.us](mailto:klw05@health.state.ny.us).

We would like to comment on some difficulties that were encountered with electronic submission of the PT results. There is still some confusion about the PSA2 line in the event menu. The PSA2 line was added to allow entry of results from a second PSA assay **only** for those labs that use a different method for total PSA for the determination of the free/total PSA ratio. If only one PSA test was done, then these results should have been entered in the first PSA line. Most labs should have selected “test not performed” for PSA2 since only a few actually do a second assay. For labs that entered two PSA tests, the primary PSA test should have been entered on the first PSA line and the secondary assay for determination of the free/total PSA ratio on the PSA2 line. Some labs entered the same PSA results twice for this PT, possibly because they did not realize they had to select “test not offered” for PSA2 and thus were not able to move off the event page. Other required fields that caused problems were those for the range of total PSA for measuring free PSA and calculating the free/total PSA ratio. Values for a quantitative range or text, such as “all levels”, “NA” (not applicable) or “see comments”, could be entered here. If the test was performed, then something had to be entered in the range field to go forward to the results page. One cautionary note: please be sure to apply the stated ranges to all of your PT samples, as a failure to apply the range correctly to all can result in sample failure.

On the results pages, the absence of data in the required fields for upper limit of the normal reference range and sample interpretation led to problems. Furthermore, some labs appear to be confusing the limits of the normal reference range for the test interpretation with the assay’s lower and upper limits of detection.

Lastly, a few labs encountered problems with using the drop-down menus. The first selection at the top of the menu is the blank and a few labs had trouble finding it to use when appropriate. Some labs selected qualifiers apparently in error and should have used the “blank” selection, and several others selected the incorrect instrument/reagent codes, for example, a reagent mis-matched with the incorrect instrumentation. Please

---

<sup>1</sup> The use of brand and/or trade names in this report does not constitute an endorsement of the products on the part of the Wadsworth Center or the New York State Department of Health.

check these entries carefully in the future to make sure your selection is correct as this could impact your PT evaluation. Also, please be aware that in the next event, these fields will be pre-populated based on what you entered this time. **Therefore, make sure that the selected instruments and reagents are correct, whether this is pre-populated from the last event or newly entered information.** This is important and in your interest since we need this information to properly evaluate your results and compare them to those of your peers. **No changes can be made for incorrect or missing information once the submission deadline has past.**

Samples:

Laboratories were challenged with five (5) different coded specimens prepared by Wadsworth Center personnel. Purified analyte preparations were added in various amounts to a protein-based matrix, sterile filtered, aseptically dispensed into sample vials and stored at 4°C until mail-out. Analyte levels were pre-assayed and stability tested in our laboratory. All laboratories received the same samples, regardless of whether they tested for one or all of the analytes.

Result evaluation:

Your laboratory's results, scores and grades are printed on a separate page. Also included are the grades from the previous two PT events and your performance status. **Please review and sign your evaluation. Keep the signed result sheet in your files.** You will need it for your next laboratory survey to demonstrate successful participation in the NYS PT program.

For your information, we also included a tabular summary of all the results with high/low cut-off values (mean +/- 3 SD) for each analyte and a graphical comparison of the results obtained with the different assay methods/kits. In order to compare results between different kits more easily across all five samples, figures for CA125, CA15-3, CA19-9, CA27.29 and CEA were prepared from normalized values that were calculated by dividing the mean values for each method by the median of the means for all kits (all kit median) for each sample. The all kit median is used instead of the all lab mean to eliminate some of the bias toward a method used by a large number of labs. For AFP, PSA, free PSA and % free PSA, the figures show the ratio of the peer group means to the assigned target value (see below), instead of the all kit median. When comparing the results, please keep in mind that for some kits the number of results (i.e. N, the number of labs measuring a particular analyte with a specific kit) was small. However, the fact that the relative performance for almost all kits has been very constant over the last several years indicates that the results shown reflect the true behavior of each method compared to its peers, at least under the conditions of the NYS PT. Note that all means were calculated from results that fell within +/- 3SD of the corresponding mean after exclusion of outliers. The tabular summary and the figures include the results from kits used by at least two labs.

Results for **AFP, CA125, CA15-3, CA19-9, CA27.29, CEA, PSA, free PSA, % free PSA and complexed PSA** were evaluated based on their respective peer group mean. In order for you to more easily compare your results to those of your peer group, we calculated a D/Dmax value and displayed it directly under your individual results. D/Dmax is a measure of how much your result (x) deviates from your peer group mean,  $D/D_{max} = (x - \text{mean}) / 3SD$ , with D being the difference of your result from the mean, and Dmax being the maximal allowable deviation, i.e. 3SD. Thus, D/Dmax needs to be between -1 and +1 for a result to be scored. **Note: If your D/Dmax is not within +/- 0.66 (equivalent to 2SD), especially for more than one or two samples, you should carefully check your assay/result(s) since this indicates that your result(s) are significantly different from the mean(s) of your peer group.** While this could be an isolated incident, it could also potentially indicate that your assay may not be performing as well as it should. Possible causes could include a calibration drift, reagents that are close to their expiration date, or subtle malfunction of your instrument. We strongly encourage you to take a close look at the run in question and others performed around that time and/or with the same reagent lots.

Results were reported by 120 labs using 11 methods to measure **CA125** (Fig.1). The figure shows a relatively consistent gradient from high to low results across all five samples. The highest measurements by TOSOH

ST A1A (TOM/TO2) were 33% higher, and the lowest results by Fujirebio/Centocor (GAA/CE1) were 34% lower, on average, than the all kit medians. Consequently, the TOSOH ST A1A results were more than two-fold higher than those from the Fujirebio/Centocor method. Results from Beckman Unicel and Access (BCU and BCX/BC1), TOSOH A1A (TOM/TO1) and Abbott Axsym, Imx and Architect (ABB, ABM & ABH/AB1) were consistently higher than the medians on average by 12-15%, while, on the other side, those from Roche Elecsys and E170 (BME & BMR/BM1) and Ortho Clinical Vitros Eci (JJC/JJ1) were consistently lower by 15-16%. Finally, four methods, including Siemens/DPC Immulite 1000, 2000 and 2500 (DPB, DPD, & DPF/DP5) and Siemens/Bayer ADVIA Centaur (COB/BA1) were essentially identical and within 6% of the medians. Thus, consistent with previous results, there are large differences in how CA125 is measured.

The MUC1 breast cancer antigen was measured by 106 labs, evenly split between those that used one of seven different **CA15-3** assays (Fig. 2), and those that used one of four different **CA27.29** assays (Fig. 3). Results for CA15-3 from the Abbott Axsym & Architect (ABB & ABH/AB1), Beckman Unicel and Access (BCU & BCX/BC1) and Siemens/Bayer ADVIA Centaur (COB/BA1) methods were similar and within 8% of the medians, whereas those from the Ortho Clinical Vitros Eci (JJC/JJ1) and Roche Elecsys and E170 (BME & BMR/BM1) methods were consistently lower than the medians on average by about 17-20%. In contrast, measurements by the Siemens/DPC Immulite 1000/2000 and 2500 (DPB, DPD & DPF/DP5) methods stood out with measured concentrations for CA15-3 more than twice as high as the medians. Consequently, the Siemens/DPC results were not included in the calculations for the all kit means and medians because of the impact that these large differences would have had on those values. It is noteworthy that a similar result was also reported in the latest CAP survey (TM-B 2007). For CA27.29, results (Fig. 3) from the two Siemens/Bayer ADVIA-Centaur and ACS-180 (COB & COS/BA1) and the TOSOH A1A (TOM/TO1) methods were close to the medians (within 5%). In contrast, those from the TOSOH ST A1A (TOM/TO2) were from 10-35% higher than the medians, with the highest positive bias observed at the lowest CA27.29 concentration. Overall, the CA27.29 results were about 20% higher than the corresponding CA15-3 results.

Results for **CA19-9** (Fig. 4) were reported from only 58 labs. More than half of these labs (30 or 52%) used Siemens/Bayer ADVIA Centaur (COB/BA1), 11 labs (19%) used Beckman Unicel (BCU/BC1) or Access (BCX/BC1), and 8 labs (14%) used Roche Elecsys or E170 (BME & BMR/BM1). All the other methods were used by 3 or fewer labs. The results from the Beckman Access (BCX/BC1) and Unicel (BCU/BC1), Siemens/DPC Immulite 2000 (DPD/DP5) and Roche Elecsys and E170 (BME & BMR/BM1) methods were similar and within 3% of the medians. In contrast, those from Fujirebio/Centocor (GAA/CE1) and Siemens/Bayer ADVIA Centaur (COB/BA1) were, on average, 50-70% higher, while, on the other side, those from TOSOH A1A (TOM/TO1) and TOSOH ST A1A (TOM/TO2) were 16 and 25% lower than the medians. Interestingly, for sample TM172 only, results from the two TOSOH methods were essentially the same as those from the other methods. The reason for this difference in relative performance between the samples is not known. There seems to be a clear difference between measurements from the high Siemens/Bayer and Fujirebio/Centocor and the low TOSOH methods compared to those made by the other four methods.

Results for **CEA** (Fig. 5) were reported by 176 labs using 14 different methods. The figure shows a relatively consistent gradient from high to low results across all five samples. The highest measurements by TOSOH ST A1A (TOM/TO2) were 24-29% higher than those by any of the other methods, whereas, Roche E170 and Elecsys (BMR & BME/BM1) gave consistently lower measurements ranging from 23-26% lower than the medians. The other (11/14) methods gave measurements that were on average within 13% of the medians. The CEA results from this PT, shown in Figure 5, as well as those from previous PTs, suggest that differences exist among the methods used to measure CEA.

As in the last several PT events, target values were assigned using traceable International Standards for **AFP, free PSA and PSA**. Although results for AFP, PSA and free PSA were evaluated based on their respective peer group means for grading purposes, information on the performance of individual methods relative to the

target values for these analytes is provided in the discussion below, as well as in the summary tables and graphs.

Absolute target values for AFP, PSA and free PSA were established based on the following International Standard preparations that were obtained from NIBSC (National Institute for Biological Standards and Control, A WHO International Laboratory for Biological Standards, Blanche Lane, South Mimms, Poters Bar, Hertfordshire EN6 3QG, UK, <http://www.nibsc.ac.uk>): PSA (free), 96/668, 1 µg per vial; PSA (90:10), 96/670, 1 µg per vial; and AFP, 72/225, 100,000 IU per vial with a conversion factor of 1.21 ng/IU. Each vial was resuspended as recommended by NIBSC, followed by serial dilution to obtain six different concentrations. Each dilution was measured in duplicate on a Beckman Access and a Roche Elecsys 2010 instrument and the measurements repeated later, if possible, with a different reagent lot, and in collaboration with Siemens/Bayer Diagnostics, on an ADVIA Centaur (only AFP and total PSA). The raw data from each measurement were used to construct separate standard curves, which were then used to assign the respective analyte concentrations (assigned target values) to the TM171-175 samples that had been measured in the same run as the standards. Thus, two sets of target values were obtained from the Beckman Access and Roche Elecsys 2010 instruments for AFP, total and free PSA and one set of target values was obtained for AFP and total PSA from the Siemens/Bayer ADVIA Centaur. These were then averaged to obtain the target values for each sample and analyte. The respective target values with their standard deviations can be found in the summary tables.

Results for **AFP** (Fig. 6) were reported from 101 labs using 11 different methods. All results were evaluated according to traditional peer group statistics and received a passing score if they fell within the mean  $\pm$ 3SD. In addition to the peer group statistics, the ratio of the group mean/target value is given for each sample to compare measurement biases between the different methods. Most results were within 10% of the target, with the exception of the Ortho Clinical Vitros Eci (JJC/JJ1), which gave results on average 17% lower than the targets. These results suggest that overall the methods for AFP are well standardized.

Results were reported by 271 labs using 15 different methods to measure total **PSA** (Fig. 7). The samples were prepared as a mixture of 10% free and 90% ACT-complexed PSA for TM172, as a mixture of 20% free and 80% ACT-complexed PSA for TM171, 174 and 175, and as a mixture of 30% free and 70% ACT-complexed PSA for TM173. All results were evaluated according to traditional peer group statistics and received a passing score if they fell within the mean  $\pm$ 3SD. In addition to the peer group statistics, the ratio of the group mean/target value is given for each sample to compare measurement biases between the different methods. The average bias for all methods in this PT was 1.05. The majority (ten) of the methods, including Dade Behring Dimension (DUD/DA1), Abbott Architect (ABH/AB1), Siemens/Bayer ADVIA-Centaur or ACS-180 (COB or COS/BA1), Roche E170 (BMR/BM1), Ortho Clinical Vitros Eci (JJC/JJ1), Siemens/DPC Immulite 3rd generation PSA (DPB, DPD or DPF/DP6), Siemens/DPC Immulite 1000, 2000, and 2500 (DPB, DPD & DPF/DP5) and Elecsys (BME/BM1) gave results for PSA with an average positive bias of 1.8  $\pm$  4.2%. However, five methods, including Beckman Access and Unicel (BCX & BCU/BC1), TOSOH A1A (TOM/TO1), TOSOH ST A1A (TOM/TO2), and Abbott AxSYM or Imx consistently gave higher PSA results with an average positive bias of 11.8  $\pm$  2.7%. This difference is statistically significant with  $p < 0.001$ . There are currently two different calibration methods for the various PSA assays as described in an article by Julie McDowell in the June 9, 2005, online newsletter of the American Association of Clinical Chemistry (AACC): [Clinical Laboratory Strategies](#). One of these is the original Hybritech Tandem-R assay (or “traditional” method), and the second is the WHO standard based on the First International Reference Standard for PSA available from NIBSC (PSA 90:10, 96/670). The latter is the same standard that was used to determine the target values for our PT samples. Since the two calibration standards can result in differences in the PSA measurements obtained from different assays, it is likely that the four methods with substantially higher PSA results than those observed from the majority of the other methods were calibrated against the original Hybritech standard, while the other methods used the WHO calibration standard. As noted previously, the standard used for calibration can have clinical implications when the result measured is close to a decision point, such as the 4 ng/ml cut-off, and thus it may be important for the physician to know what calibrator you used.

Seventy-three labs measured **free PSA** (Fig. 8) with slightly less than half (47%) of the results reported with the Beckman Hybritech Access or Unicel (BCX or BCU/BC1) methods. All results were evaluated according to traditional peer group statistics and received a passing score if they fell within the mean  $\pm$ 3SD. However, in addition to the peer group statistics, the ratio of the group mean/target value is given for each sample to compare measurement biases between the different methods. As observed in previous PTs, Beckman Unicel and Access (BCU and BCX/BC1) results were consistently higher (on average 40% for this PT) than the target values for these samples, whereas most of the results from the other methods were within 10% of the targets.

As in prior surveys, the figure for **% free PSA** (Fig. 9) is meant as a qualitative comparison only since there is a large number of method combinations used for its determination. The figure shows the method mean %free PSA/target %free PSA (generated from the ratios of the free PSA target values to the total PSA target values). As usual, since the %free PSA is derived from the ratio of free to total PSA, the differences in free PSA and total PSA measurements are reflected, or possibly even exaggerated, in the ratio. As could be expected, the higher free PSA values measured with the Beckman Hybritech Unicel and Access assays resulted in %free PSA values that were between 16-23% higher than the targets. The remaining four method combinations, including Roche Elecsys and E170 (BME & BMR/BM1), Siemens/DPC Immulite and Dade Behring Dimension (DUD/DA1) gave ratios that on average were within 10% of the target. It should be noted that the %free PSA that is calculated from measurements of complexed PSA does not always give results that are comparable to those calculated directly from measurements of free PSA in the same samples, especially if non-homogeneous methods are used. In the latter case, the results are usually substantially different from those obtained either directly from free PSA measurements or from a homogeneous combination of assays for complexed and total PSA.

**Note:** Several labs measured free PSA even though the total PSA was outside the range for measuring free PSA given by the lab. This appears a violation of these labs' policy, and indicates that they did not treat the PT specimen exactly like a patient sample. Labs are expected to calculate the %free PSA if they perform the free and total PSA assays and would do so for a similar patient sample. However, if a lab's policy is not to measure and calculate %free PSA outside a certain range of total PSA, then this rule should also be applied to the PT samples. In that case, please indicate this on the result sheet, so we know that the failure to provide a result was deliberate, or the absence of the %free PSA calculation without an acceptable explanation for its omission will be counted as a failure. Furthermore, results must be given as percent free PSA, not as a fraction.

Only 11 labs measured **complexed PSA**, and all of these used the Siemens/Bayer ADVIA Centaur or ACS-180 method. Furthermore, the mean % complexed PSA calculated from these values of 77.4% compared well with the mean of 18.9% free PSA for TM171, TM174 and TM175, the mean % complexed PSA of 88.7% compared well with the mean of 10.0% free PSA for samples TM172, and the mean % complexed PSA of 74.2% compared well with the mean of 26.3% free PSA for samples TM173.

**Cut-off values:** As explained previously, the result we intended to get for cut-off values was the upper limit of your normal or reference range for each analyte, above which you (or your computer) would flag a result as elevated or abnormal. We also asked you to classify each result as either normal, i.e. within the normal or reference range, or abnormal or elevated, i.e. above the reference range. We will continue to ask for this and expect it to be filled in the result form. As recommended in the instructions included with the samples, where there is a range of reference values (for example, age-specific reference values, or smokers versus non-smokers), please enter that information in the comments on the form. Also, if there are two or more reference values, e.g. smoking versus non-smoking populations, please use the non-smoking reference for your normal versus abnormal evaluation, but enter a note in the comment section that there are two or more reference values and list the other values if possible.

In conclusion, there can be significant differences between results obtained with various methods, especially for CA125, CA15-3, CA19-9 and CEA, as observed previously. While some of these may be due to the

artificial nature of the PT samples, others are probably due to inherent differences in the assays themselves. We will continue to try to minimize the differences that can be attributed to the sample composition. Nevertheless, despite the admittedly somewhat artificial nature of the PT samples, we would like to suggest that the differences between results obtained by various methods might also be reflected in patient serum samples. Therefore, caution needs to be used when comparing the results from the same patient obtained with different methods, since clearly not all methods are equal. For this reason, we require that the method used must be clearly indicated on the patient report (Oncology Standard OC 3b). We would also like to encourage you to educate your physician clients about this potential problem. Furthermore, the comparison of method means to target values set by traceable International Standards for AFP, PSA, and free PSA clearly shows that there are a few methods that appear to be calibrated differently from the rest of the methods used to measure these analytes as discussed in the respective paragraphs.

Finally, we would like to raise the usual cautionary notes when interpreting these results which are 1) since some of the assays were done by a small number of labs, the results might be skewed due to a lack of statistical power; 2) it is difficult to make an accurate comparison of results when the %CVs are large; and 3) the analyses for PT purposes are done with artificially prepared mixtures of proteins which may or may not accurately reflect patient derived samples.

If you have any questions or wish to discuss some of the issues alluded to you may contact us at the address below. Also, this discussion with the tables and figures (in color) will eventually be posted on our website at <http://www.wadsworth.org/labcert/lep/PT/oncology/index.html>.

For your information, the schedule for the 2008 Tumor Marker Proficiency Test mail-out is:

**Mail-out date:**

January 29, 2008  
May 6, 2008  
September 9, 2008

**Due date:**

February 13, 2008  
May 21, 2008  
September 24, 2008

Erasmus Schneider, Ph.D.  
Director, Oncology Section  
Clinical Laboratory Evaluation Program  
Wadsworth Center  
Empire State Plaza  
Albany, NY 12201-0509

Ph: (518) 474-2088  
FAX: (518) 474-1850  
Email: [schneid@wadsworth.org](mailto:schneid@wadsworth.org)