NORDIC BIOMARKER

2012-08-16

Performance Evaluation of D-dimer applications

Introduction

Many of our customers are experienced and choose to design and evaluate their own applications. Others are wizards on instruments, but have not considered the performance of reagents before. Yet others, are more focused on the medical side of the matter, and like to have the application ready from the supplier before they get started on their own. Whoever you are, we hope this text can help you.

We have seen in practice, that there are many different ways to evaluate or verify the performance characteristics of an application; which features that are tested, the number of measurements, etc., can wary considerably between different laboratories around the world. The European Standard EN 13612, *Performance evaluation of in vitro diagnostic medical devices*, does not suggest specific experiments to perform, or limits to adhere to, for a successful evaluation. Rather, this standard is more a general guideline with a list of topics to consider for the evaluation of an IVDD. CLSI in the USA, however, are more specific as to exactly which experiments to conduct. Therefore, the set of performance tests suggested below in this document has been strongly influence by the CLSI guidelines. That way, we hope that the data can be presented in a format that is useful for most markets.

Before the evaluation can start, we must first come up with an application; otherwise there is nothing to evaluate. An instrument application is a number of predefined settings in an instrument, and the application should be designed to give the reagent optimal performance on that instrument. In another document of ours, *Design of D-dimer Applications*, we describe how to design an application in a systematic manner.

In the end, what we ultimately strive for is an application that is reporting correct and consistent D-dimer results to the end-user of the system. The performance evaluation of the application can be viewed as a number of stress tests, to ensure that reported D-dimer values are in fact correct. Most importantly, with the performance data at hand, the limitations of the application can be reported in application sheets and directional inserts, and clearly communicate the boundaries within which correct results can be expected.

NORDIC BIOMARKER

2012-08-16

Linearity

Linearity is the ability to provide results that are directly proportional to the concentration of the analyte in the test sample. Thus, the calibration curve of the method need not necessarily be linear, but the over system response (i.e. analytical results) should be! For these measurements, individual patient samples should be avoided, as they may contain interfering substances that will contribute to non-linearity. The matrix should be similar to real specimens, and dilution buffers should not be used. Instead, plasma pools from several donors should be used. In practice, a plasma pool in the upper linear range of the method is mixed with a plasma pool in the lower range to get at least 7 (preferably 11) equally spaced concentrations within the claimed linear range.

- Reconstitute 6 x 1 mL D-dimer Calibrator as per IFU (or some other normal plasma pool that has been spiked with D-dimer into the range 2880 – 3520 ng/mL)
- Pool the contents of all 6 vials into only one of the vials.
- Pipette 1000 µL, 900 µL, and so on, of the pooled calibrator into tubes 1 - 10.
- Pipette 100 µL, 200 µL, and so on, of the Normal Plasma Pool into tubes 2 – 11.
- Close all tubes and mix by inverting several times.
- Analyze all 11 samples 4 times in a random order.

The table is an example only, and the actually pipetted volumes (if recorded by a balance) should be used for the calculations. 3200 and 50 ng/mL are also examples, and should be replaced by actually measured values or previously assigned values.

Sample	Calib. (µL)	NPP (µL)	Conc.(ng/mL)
1	1000	0	3200
2	900	100	2885
3	800	200	2570
4	700	300	2255
5	600	400	1940
6	500	500	1625
7	400	600	1310
8	300	700	995
9	200	800	680
10	100	900	365
11	0	1000	50

Limit of Detection & Quantification

Both the limit of detection and the limit of quantification are easiest to analyze together.

- Prepare a D-dimer stock sample holding approximately 400 ng/mL.
- From this stock sample, prepare further samples with D-dimer concentrations 200, 150, 100 and 50 ng/mL using Dilution Buffer as the diluent.
- Analyze these dilutions, including the 400 ng/mL stock samples, 6 times with the application. This is only a preliminary analysis to 1) establish the concentration of the stock sample and 2) to asses roughly the LoQ.
- From these results, prepare 5 samples each of 2 different D-dimer concentrations that you deem to be close and slightly above, respectively, the LoQ.
- Analyze the 2 D-dimer concentrations, as well as the Dilution Buffer itself, at least 12 times. That is a total of 60 points for these 3 samples.

Before the experimental part is over, make sure that both the precision and accuracy are in fact, for at least one of the two different D-dimer levels, better then the preset thresholds for the LoQ.

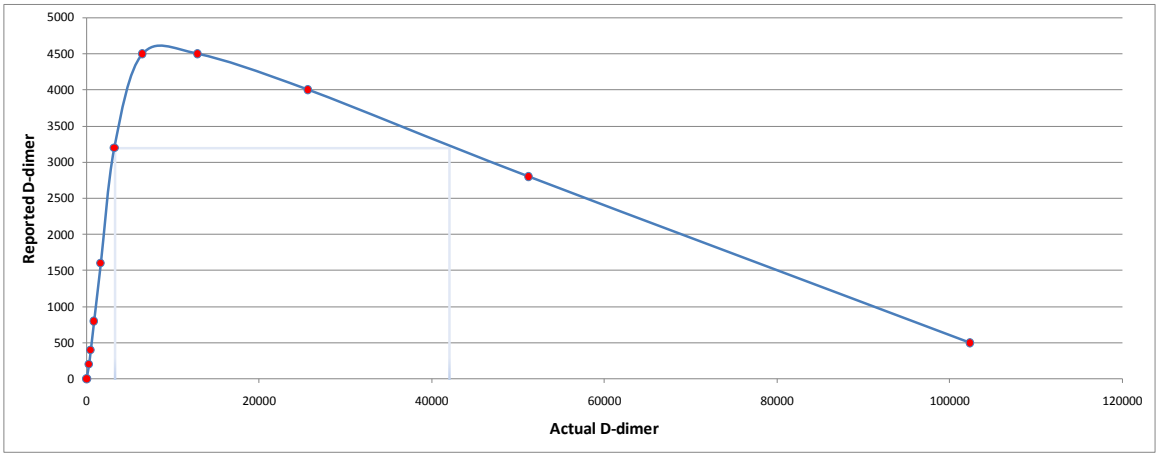
Precision

Three D-dimer levels within the linear range 20 days x 2 runs x 2 duplicates.

- Prepare three D-dimer samples with 250, 1500, and 3000 ng/mL in large amounts (analyze them to ascertain that they are in fact close to the desired levels). Make several aliquots of each sample (>40) and freeze them all at -20°C.
- In the morning, thaw a sample and analyze it 2 times. Before the end of your work shift, thaw a fresh sample and analyze it 2 times.
- Repeat the procedure 19 more days.

Upper reportable range

Even if the calibration curve is nice and linear all the way up to the value of the un-diluted calibrator, there will be higher D-dimer levels at which the latex particles becomes more or less saturated (e.g. figure below).



This effect is commonly called the “prozone” effect, or the “high-dose hook” effect. Often there is a claim the directional insert saying something like “there is no prozone effect below X ng/mL”; where X is a relatively high number. What this means, as illustrated in the constructed example in the figure above, is that if you claim a reportable range up to 3200 ng/mL, you must also measure higher concentrations to find if there are any other concentration that can give the same apparent result. In the example above, both 3200 and 42000 are reported as 3200 ng/mL.

The shape of the curve you see in the figure above, can be altered significantly just by changing a few parameters in the application settings of the instrument. Therefore, the prozone effect and the upper reportable range, is tightly linked to the instrument and the instrument application that the reagent is running on. Actually, a prozone effect can be *more* due to a poorly designed instrument application, than due to a poor reagent. This latter statement is particularly true for newer instrument that typically have good software that can analyze reaction curves and warn the operator for high-dose hook effects.

So, a prozone test ought to be conducted for all new instrument applications. Here is how we typically do this. First, construct a dilution series by diluting a D-dimer bulk with normal plasma.

Make the first sample to be approximately 100,000 ng/mL D-dimer. Then serial dilute this first sample with normal plasma to give decreasing concentration down below 1000 ng/mL. For instance, 100000, 50000, 25000, 12500, 6250, 3125, 1562, and 781 ng/mL. These 8 samples, and the normal plasma, should be assayed in duplicate.

Now, you don't really know that the final concentration in the series really is 781 ng/mL; the D-dimer bulk may not have been perfectly assigned, or your initial dilution down to 100000 may not have been perfect. But, provided that your dilutions from 100000, and down, are well done, you can back calculate the concentration of the series by using the assays results that you obtain for the last two sample in the series, that are surely within the reportable range. Don't forget to take into account the assayed D-dimer value of the normal plasma that you used for the dilutions.

Reference of normality

Analyze at least 150 normal patient samples in duplicate.

Method comparison

Correlate against a recognized reference method, preferably one that runs on the same instrument and/or that the end-user of the new application has used before. If the reference method runs on a different instrument placed in another laboratory, and it is not possible to run the patient samples on both side-by-side, it is better to divide each sample into aliquots and freeze them. Then, a sample can be thawed at each location where both have been treated equally. Even if D-dimer levels have been shown to be unaffected by several freeze-thaw cycles, it takes away the argument of differently treated samples, when there are deviations from the correlation plot (and samples containing lipids can in fact become more opaque after a freeze-thaw cycle, which can affect the results) .

Analyze at least 50 patient samples in duplicate with both methods over at least 5 days. The results from the 50 samples should be spread over the entire linear range (if not, complement with more patient samples). If any sample deviates drastically from a correlation plot, first rerun this sample again. Second, manually dilute the sample (e.g. 1:5) and rerun (i.e. one method could be more sensitive than the other to an interfering substance present in that sample).

Interferences

We typically test interferences from Hemoglobin, Billirubin, Lipids (intralipid), and Heparin for each method. As interferences from such substances are dependent on both optics and sample amount used, it is relevant to test these characteristics for each new instrument application.

- Make up several samples with an intermediate D-dimer concentration.
- Make up a dilution series of the interfering substance to be tested, having a 10-fold higher concentration than the interval you wish to test.
- Add 1/10 Vol of the dilution series to the D-dimer samples, including only buffer to the zero/reference sample.
- Analyze each sample in duplicate.

General remarks

For each set of experiment above, review all data before you leave the instrument for the day/session. There may be outliers/errors amongst the data, and the best time to find reasonable explanations for these (if at all possible), is at the instrument as soon as possible after the measurement as possible. Plotting and viewing the data is usually the quickest way to see any problems. If a sample needs to be rerun, the time to find this out is not at the desk in the office a week later!

Whenever possible, use analytical balance to corroborate all pipetting results.

Keep records of each experiment, when it was made, and where the raw data can be found. Take note of part numbers, lot numbers, and expiry dates for all the reagents used.

Use just one kit lot number for all evaluations. Lot-to-lot variations of the reagent are not tested here (nor are instrument-to-instrument variations, so you may use only one instrument as well).

Tests that are not instrument specific

For a complete product, we do not only need to evaluate the instrument application, but also other characteristics of the reagent. However, as these characteristics are general for the reagent, and not instrument specific, they are not within the scope of this document.

Take for instance the shelf-life or open-vial stability of the reagent. Although we would need one instrument application to asses these characteristics, e.g. to see any drift in performance over time, we would hold the results general for the reagent and not test them over and over again on each platform used.

Nor is the traceability or the lot-to-lot consistency of the calibrator within the scope of the performance testing of a single instrument application. Clearly, this factor is not instrument specific. Yet, it is perhaps the single most important factor to ensure that correct and consistent assay results are delivered to the end-user. That is also why the decision cut-off should not be validated for each new instrument application. The numerical value of the cut-off has more to do with the assignment of the calibrator, and has very little to do with which instrument application that is used.