Commentary | **Open Access**

About Estimating the Limit of Detection by the Signal to Noise Approach

Desimoni E*and Brunetti B

Department of Food, Environmental and Nutritional Sciences, University of Milan, Via Celoria, Milan, Italy

Abstract

The limit of detection (LOD) is an important figure of merit estimated in validation studies or when reporting the performances of an analytical method. While the very meaning of LOD is clear to anyone, its estimation is still matter of discussion. This explains the plethora of available approaches. However, the most popular are generally the simplest ones. Between these, it is worth mentioning the approach based on the signal to noise (S/N) ratio, included among those suggested by international guidelines such asthe United States Pharmacopeia, the European Pharmacopoeia and others. This contribution attempts evaluating if this approach could be replaced by that based on the standard error of estimate, $s_{y/x}$.

Introduction

According to IUPAC, the limit-of-detection (LOD) is a measured quantity value, obtained by a given measurement procedure, for which the probability of falsely claiming the absence of a component in a material is β, given a probability α of falsely claiming its presence [1]. Recommended default value for α and β is 0.05 (i.e. 5%). According to ISO, the LOD is the true net concentration or amount of the analyte in the material to be analyzed which will lead, with a probability (1-β), to the conclusion that the concentration or amount of the analyte in the analyzed material is larger than in the blank material [2]. Both definitions require the choice of the acceptable values, α and β, of the probabilities of false positive (type I) and false negative (type II) errors, respectively. When dealing with LODs, type I error is detecting the analyte while it is not present, while type II error is failing to detect the analyte presence.

ISO and IUPAC definitions are unquestionable, but a lively debate remains open about how estimating the LOD [3-18]. First, the very meaning of LOD itself needs a comment. In fact, whatever the adopted approach, the LOD is only a point-estimate test-statistic for the true LOD composite population parameter [19]. Obviously, any subsequent evaluation of LOD, by using a new, independent set of experimental data, always gives a more or less different estimate. Moreover, the true probability density function of the blank signal (which is at the basis of some popular LOD approaches) is likely skewed, since lying close to a physical limit, the absence of the analyte [20]. In this context, even parametric statistics are questionable. It follows that any LOD estimate represents only an approximated indication of the analyte being present/absent. These aspects, together with the many approximations involved in the practical LOD estimation (see below), justify the acceptance of the most approximated approaches.

Presenting or discussing pros and cons of the very many procedures developed for estimating LOD values is outside the aims of this contribution. Here below, the attention focuses to the Signal-to-Noise (S/N) approach. The International Conference on Harmonization (ICH) [21], the United States Pharmacopeia (USP) [22], the European Pharmacopoeia (Ph. Eur.) [23] and the International Organization for Standardization (ISO) [24] includes this simple approach among those suggested. This explains its popularity, for example in chromatography, spectroscopy and electroanalysis. The following discussion attempts evaluating if the approach based on the standard error of estimate, $s_{y/x}$, can replace the S/N one.

The Signal-to-Noise Approach

The idea at the basis of the *S/N* approach is defining the LOD as that analyte concentration which is large enough to produce a signal (peak or plateau) enough larger than the noise, the signal recorded in the absence of the analyte, also called the blank signal. This approach is usually adopted in the belief that noise can be easily estimated. Choosing as LOD a *S/N* ratio equal to 3.0 allows proving the presence of the analyte in the test sample with a probability larger than 99%. The S/N ratio is often evaluated manually. Its main advantage is that modern analytical instruments make available a large number of signal and noise values. In the example presented in Figure 1, a signal Y of about 0.31 a.u. (a.u.: arbitrary units) overlaps to a noise bandwidth of about 0.17 a.u. This peak is low enough to pertain to a concentration range close to the LOD. However, several points must be considered before estimating the S/N ratio:

1. Deciding the value of the noise bandwidth can be problematic. Often, it is less regular than in Figure 1. Specific information describe test statistics for reliably testing the population signal-to-noise ratio [25].

2. Must the noise bandwidth exclude eventual spikes (abnormal values of the background signal) or not?

3. Some international bodies define the signal-to-noise ratio as $2S/N = 3.0$ in place of $S/N = 3.0$: see for example reference [10,22,23,26,27]. This means considering the half-width of the noise band. In the case of Figure 1, and accepting the noise bandwidth equal to 0.17 a.u., the *S/N* ratio at the peak maximum can be estimated as $2S/N = (2 \cdot 0.31)/0.17 = 3.6$ a.u. and, respectively, as $S/N = 0.31/0.17 = 1.8$ a.u. If the LOD is fixed at 3.0, the result 3.6 means that the test sample does contain the analyte, while the results 1.8 means that analyte is absent (or lower than LOD).

*Corresponding author: Desimoni E, Department of Food, Environmental and Nutritional Sciences, University of Milan, Via Celoria, Milan, Italy, Tel: 0039 02 50319223, Fax: 0039 02 50319061; E-mail: elio.desimoni@unimi.it
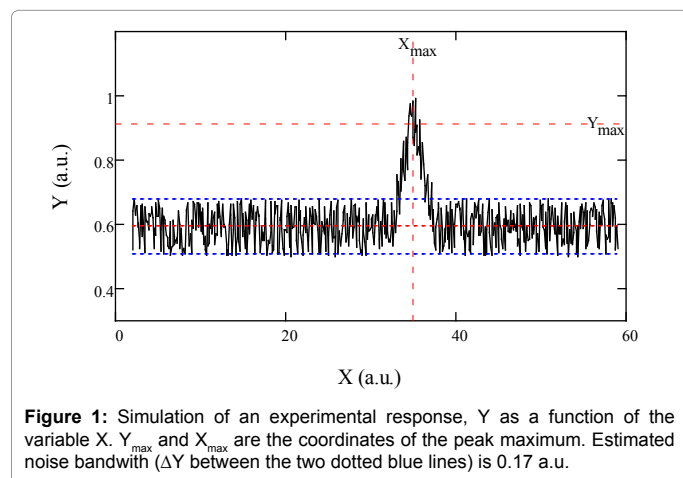
Figure 1: Simulation of an experimental response, Y as a function of the variable X. Y_{max} and X_{max} are the coordinates of the peak maximum. Estimated noise bandwith (ΔY between the two dotted blue lines) is 0.17 a.u.

4. Noise may change even over limited periods (as during the recording of a chromatogram or a voltammogram).

5. The *S/N* approach consider only peak height measurements, not peak areas (how managing asymmetric/skewed peaks?).

6. The *S/N* value includes mainly the instrumental noise and does not include the "chemical" noise originated by variations of the signal arising from sample in homogeneity and sample preparations along the entire measurement process [28].

7. The *S/N* ratio can dramatically change (improve) by eventual smoothing or thresholding treatments of the raw data.

8. The *S/N* approach does not conform to the indications of ISO and IUPAC definitions [1,2]. It was in fact underlined that "the signal-to-noise school explicitly recognize only the false positives, which in effect makes the probability of the false negatives equal to 50%" [4]. This value of β equals the probability of observing heads or tails by coin tossing. Neglecting all the above listed points in routine analyses allows an easier estimation of the LOD.

Using the Standard Error of the Estimate

The approach based on the standard error of the estimate derives from that based on the population standard deviation of the signals of the blank, σ_B. Even if it is well known [4,11,13,16,17,19,21,28-31], it is shortly recalled here below for the reader convenience. According to this approach, the LOD can be estimated by the equation

$$LOD = k . \sigma_B / b \qquad (1)$$

where σ_B is the population standard deviation of the blank signals and *b* is the slope of the signal/concentration functional relationship, usually obtained by ordinary least squares regression (OLS). *k* is the expansion factor chosen according to the analyst preferences about the acceptable α and β values. Using $k = z_{1-\alpha} = 3.0$ ($z_{1-\alpha}$ is the one-tail, standardized normal variable), as chosen by several Authors, indicates the choice of a probability of false positive errors (α) of 0.135%. In this case the probability of false negative errors, (β) is 50% [31]. Such a β value implies no control of false negative errors, as in the case of the S/N approach. If Authors chose controlling both kinds of errors, for example at 5% (α = β= 5%), *k* in equation 1 changes to [31]

$$k = \left(z_{1-\alpha} + z_{1-\beta} \right) = 2 \cdot z_{1-\alpha} = 3.3 \qquad (2)$$

This approach also relies on quite severe theoretical assumptions, concerning the estimation of σ_B and of *b*. Those relevant to σ_B are the following [28,32]:

1. Random errors in the blank signal must be normally distributed;

2. The population parameters of the signal distribution of the blank must be known;

3. Systematic errors must be negligible or absent;

4. The analyte concentration in the blank is effectively equal to zero;

5. The variance of the blank signal is equal to that of samples with very low analyte concentrations;

6. A consistent number of independent measurements is necessary, since σ_B is the standard deviation of the population of the blank signal [29].

If condition *6* is not verified, σ_B must be replaced by its estimate (s_B), the standard deviation of the sample of the blank signal, and z must be replaced by the *t*-values of the *t*-Student distribution. In this case, calculations become more complex [33].

If, as above mentioned, the slope *b* is estimated by ordinary least square regression, the theoretical assumptions at the basis of OLS [9,28,31,34] add to conditions 1-6. They are the following:

7. Random errors must occur only in the *y*-direction within the explored concentration range;

8. All random errors must be normally distributed within the explored concentration range;

9. The matrix of all the examined samples must be identical;

10. Variance must not change within the explored concentration range (that is the analytical system is homoscedastic: this corresponds to point *5*.);

11. Signals must be linearly related to concentration in the explored concentration range;

12. Good estimates of the slope, *b*, and intercept, *a*, of the calibration line must be available;

13. The intercept *a*, must not be significantly different from the mean blank signal, μ_B.

Conditions 1-13 are hardly satisfied in real work but they are usually accepted as valid *a* priori, since allowing an easier and simplified estimation of the LOD.

The main objection against this approach is the eventual unavailability of an actual blank, or the impossibility to measure the signal of the blank, such as when using instrumentations which automatically subtract the background from the responses. In this last case, it is possible spiking the blank with the lowest analyte concentration allowing the measurement of the minimum signal different from zero [29]. However, on considering that experimental precision exponentially decreases on decreasing the analyte concentration, even a limited spiking of the blank with the analyte can lead to LOD estimates appreciably more optimistic (lower) than that based on a true blank. Moreover, as above underlined, the approach needs a consistent number of measurements of blank or fortified blanks.

The approach based on the standard error of the estimate, the statistic $s_{y/x}$ allows avoiding some of the problems above highlighted. The value of $s_{y/x}$ is calculated when performing OLS, since allows

evaluating the uncertainty of slope and intercept of the regression line [31]. It was underlined that $s_{y/x}$ can replace σ_B because assumption *x.* implies the homoscedasticity of the analytical system [31]. This means that each signal used to estimate the regression line, including the signal of the blank, has a normally distributed y-variation with a standard deviation estimated by $s_{y/x}$ [31,34]. This replacement seems very convenient, since $s_{y/x}$ is already known from OLS calculations when reporting the equation of the regression line as an analytical performance of a given method. In this way, no additional work is necessary for estimating σ_B (to apply equation 1). Of course, some care is also necessary when using $s_{y/x}$ in place of σ_B. In fact, it is evident that different values of $s_{y/x}$ and *b* are obtained by repeating the calibration or by changing the number of data points. When estimating the LOD, the best way should be performing an OLS on few data acquired in a narrow concentration region close to the lowest limit of the linear range. This is not an additional workload, since these calibration points add to the others when performing the regression over the whole explored linear concentration range. It follows that the LOD can be estimated as

$$LOD = k \cdot \frac{s_{y/x}}{b} \qquad (3)$$

Even in this case, the choice of *k* depends on the selection of proper values of α and β. Using $k = 3.3$ (see equation 2), in which case α=β=0.05, meets the requirements of ISO and IUPAC [1,2]. The convenience of using this approach was already underlined in the recent literature [35,36].

Conclusion

The LOD approaches based on the S/N ratio and on $s_{y/x}$ rely both on severe limitations and on approximations, which can be ignored in order of ensuring the largest acceptance by the users. These simplifications are at least in part justified by considering that the very meaning of any LOD estimate is only a point-estimate test-statistics for the true LOD value. However, the approach based on $s_{y/x}$ seems preferable, since allowing the control of both types of errors without the need of acquiring repetitions of the blank signals. May be that only few additional data points are necessary when performing the calibration for measuring $s_{y/x}$ in a concentration region close to the lowest limit of the linear range. In fact, slope and $s_{y/x}$ may significantly change with the explored concentration range.

At last, users of the LOD estimate have the responsibility of properly understanding and justifying the reported value. This suggests that a detailed description of the adopted approach is mandatory. Comparing LOD values from different sources without considering the differences between the experimental conditions (approach, analytical technique, analyte, matrix, number of measurements, etc.) is meaningless.

Acknowledgements

References

1. McNaught AD, Wilkinson A (1997) IUPAC Compendium of Chemical Terminology (2nd edn) The Gold Book, Blackwell Scientific Publications.

2. (2008) International Vocabulary of Metrology – Basic and General Concepts and Associated Terms, The News Magazine of the International Union of Pure and Applied Chemistry (IUPAC).

3. Huber W (2003) Basic calculations about the limit of detection and its optimal determination. Accred Qual Assur 8: 213-217.

4. Currie LA (1997) Detection: International update, and some emerging di-lemmas involving calibration, the blank, and multiple detection decisions. Chemom Intell Lab Syst 37: 151-181.

5. Voigtman E (2008) Limits of detection and decision. Part 1, Spectrochim Acta Part B 63: 115-128.

6. Voigtman E (2008) Limits of detection and decision. Part 2, Spectrochim Acta Part B 63: 129-141.

7. Voigtman E (2008) Limits of detection and decision. Part 3, Spectrochim Acta Part B 63: 142-153.

8. Voigtman E (2008) Limits of detection and decision. Part 4, Spectrochim Acta Part B 63: 154-165.

9. Desimoni E, Brunetti B (2009) About estimating the limit of detection of heteroscedastic analytical systems. Anal Chim Acta 655: 30-37.

10. Boque R, Vander Heyden Y (2009) The limit of detection, LC-GC Eur 22: 82-85.

11. Eksperiandova LP, Belikov KN, Kimchenko SV, Blank TA (2010) Once again about determination and detection limits. J Anal Chem 65: 223-228.

12. Jasielec JJ, Sokalski T, Filipek R, Lewenstam A (2010) Comparison of different approaches to the description of the detection limit of ion-selective electrodes. Electrochim Acta 55: 6836-6848.

13. Shrivastava A, Gupta VB (2011) Methods for the determination of limit of detection and limit of quantification of the analytical methods. Chron Young Sci 2: 21-25.

14. Voigtman E, Abraham KT (2011) True detection limits in an experimental linearly heteroschedastic system. Part 1, Spectrochim Acta Part B 66: 822-827.

15. Voigtman E, Abraham KT (2011) True detection limits in an experimental linearly heteroschedastic system. Part 2, Spectrochim Acta Part B66: 828-833.

16. Rajakovic LV, Markovic DD, Rajakovic-Ognjanovic VN, Antanasijevic DZ (2012) Review: The Approaches for estimation of limit of detection for ICP-MS trace analysis of arsenic. Talanta 10: 79-87.

17. Thompson M, Ellison SRL (2013) Towards an uncertain paradigm of detection capability. Anal Methods 5: 5857-5861.

18. Fonollosa J, Vergara A, Huerta R, Marco S (2014) Estimation of the limit of detection using information theory measures. Anal Chim Acta 810: 1-9.

19. Montville D, Voightman E (2003) Statistical properties of limit of detection test statistics. Talanta 59: 461-476.

20. Currie LA (2001) Some case studies of skewed (and other ab-normal) data distributions arising in low-level environmental research. Fresenjus J Anal Chem 370: 705-718.

21. (2005) ICH Q2(R1) Validation of Analytical Procedures: Text and Methodology. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use.

22. (2007) Validation of compendial methods, The United States Pharmacopeial Convention, United States Pharmacopeia 30, National Formulary 25, Rockville, Maryland, USA.

23. (2008) European Pharmacopoeia 6.2, Council of Europe, Strasbourg, France.

24. (2006) ISO/CD 13530: Water quality – Guide to analytical quality control for water analysis.

25. Golam Kibria BM, George F (2014) Methods for testing Population Signal-to-Noise ratio. Commun Stat-Simul C 43: 443-461.

26. Dolan J, HPLC Solutions #124: How to Determine Signal-to-Noise Ratio. Part 3.

27. Saadati N, Md. Pauzi A, Zakaria Z, Tavakoli Sany SB, Rezayi M et al. (2013) Limit of detection and limit of quantification development procedures for organochlorine pesticides analysis in water and sediment matrice, Chem Cent J 7: 63-71.

28. Danzer K (2007) Analytical chemistry. Theoretical and metrological fundamentals, Springer, Berlin.

29. (2014) The Fitness for Purpose of Analytical Methods. A Laboratory Guide to Method Validation and Related Topics (2nd edn) Eurachem.

30. Otto M (2007) Chemometrics, Wiley-VCH, Weinheim.

31. Miller JC, Miller JN (2005) Statistics and chemometrics for Analytical Chemistry (6th edn) Pearson Prentice Hall, Edinburgh.

32. (1987) Recommendations for the definition, estimation and use of the detection limit. Analyst 112: 199-204.

33. Mocak J, Bond AM, Mitchell S, Scollary G (1997) A statistical overview of standard (IUPAC and ACS) and new procedures for determining the limits of detection and quantification; application to voltammetric and stripping techniques. Pure &Appl Chem 69: 297-328.

34. Miller JN (1991) Basic statistical methods for analytical chemistry. Part 2. Calibration and regression methods. A review. Analyst 116: 3-14.

35. Desimoni E, Brunetti B (2012) About literature information on functional relationships and limits of detection: may authors improve their statements? Accred Qual Assur 17: 635-637.

36. Desimoni E, Brunetti B (2013) Presenting Analytical Performances of Electrochemical Sensors. Some Suggestions a Electroanalysis 23: 1645-1651.