Significant Figures

***Tony Badrick,[1] Peter E Hickman[2]**
[1]Biochemistry Department, Sullivan Nicolaides Pathology, Taringa, Brisbane, Qld 4068, [2]Chemical Pathology, ACT Pathology, Canberra 2606, Australia.
*For correspondence: Dr Tony Badrick e-mail: Tony_Badrick@snp.com.au

Summary

- For consistency of reporting the same number of significant figures should be used for results and reference intervals.
- The choice of the reporting interval should be based on analytical imprecision (measurement uncertainty).

Introduction

The results of an analysis can be used for a number of different purposes by a clinician. Comparing a result against a reference interval or a clinical decision point can be used to confirm or reject a diagnosis. But often a result is compared with previous results from a particular patient to monitor treatment, be that result a drug level, a glycated haemoglobin level or a creatinine. In this latter case it is the difference between sequential results that contains the information needed, and this difference is conveyed to a large extent by the number of significant figures, or reporting interval, of an analyte. A further complicating factor for a clinician trying to interpret consecutive results is that they may come from different laboratories, so the clinician may be confronted with a pair of results such as 20 and 20.56!

When reporting a result, we may be faced with the problem "is an apparent change in result real, or simply a reflection of the 'noise' due to analytical imprecision and biological variation?" The choice of the reporting interval (incremental value chosen for reporting analyte concentration, e.g. for sodium this is 1 mmol/L, whilst for potassium, it is 0.1 mmol/L) should reflect this background noise.

As the measurement systems used in routine chemistry laboratories produce a continuous signal (spectrophotometric, voltage) what happens when a result is rounded is that a series of results are 'binned' together. Thus a result of 20 units actually represents the spread of measured results from 19.5 to 20.4. The 'rounding' could hide a significant change in result. For example, the range of results from 18.5 to 20.4 would be rounded to 19 and 20 units, a reported change of 1 unit, but a potential difference of up to 1.9. Often, it would appear that the reporting interval chosen is too small which can give a false impression that a change has occurred.

The reporting interval should be such that any result change is greater than the analytical imprecision. The number of significant figures reported should be dependent on analytical imprecision (standard deviation$_a$, SD_a) and, perhaps, biological variation (intra-individual SD, SD_i), but in practice often is quite arbitrary, and it is apparent that for many assays, analyte concentrations are reported to an excessive number of significant figures.[1]

Some Relevant Statistics

It is worth reviewing the relevant statistical basis of differences between two numbers which are the results of an analytical process. We assume that the measurements follow a Gaussian distribution, that is, if we repeatedly measured any sample, those results would follow a Gaussian distribution which is described by its mean and SD. That means that approximately 68% of all those repeated values would lie within the mean ±1 SD, that approximately 96% of those values would lie within the mean ±2 SD and 99.7% of values would lie within the mean ±3 SD. These multipliers of the SD, 1, 2 or 3, are called the standard normal deviates or the z-scores. Any measurement of a parameter has an associated analytical error which is described as the analytical standard deviation (SD_a), so any result will lie within the mean ± 2 SD 96% of the time.

Any biological parameter will also have some biological variation associated with it and this is described by a standard deviation (SD_i), so that a parameter will vary day to day around a mean value with an SD of SD_i.

Now we are interested in the question: when do two sequential measurements differ? Each of these measurements is subject to analytical error plus normal biological variation. This total

'normal' variation for serial results is equal to the sum of the variation about each measurement, that is total variation = variation in measurement 1 + measurement 2 =

$$z \times ((SD_a)^2 + (SD_i)^2)_1^{1/2} + z \times ((SD_a)^2 + (SD_i)^2)_2^{1/2}$$

We can simplify this to $\sqrt{2} \times z \times ((SD_a)^2 + (SD_i)^2)^{1/2}$

We are interested if the net result difference is greater than these combined effects or greater than

$\sqrt{2} \times z \times ((SDa)^2 + (SD_i)^2)^{1/2}$ which is called the critical difference or reference change value (RCV).

If we want a 95% confidence that two consecutive results truly differ, then their difference must be greater than

$2.77 \times ((SD_a)^2 + (SD_i)^2)^{1/2}$, the constant '2.77' being derived from $\sqrt{2}$ times the z statistic which is 1.96.

Different Approaches to Reporting Intervals

There have been various published approaches to determining an appropriate number of significant figures to report. For example, Hawkins and Johnson suggested that a simple rule of thumb to use for choosing the appropriate number of significant figures was that the SD should be ≤ 0.7 of the unit of reporting.[2]

For example, for serum Na, if the SD is $\leq$ 0.7 mmol/L but > 0.07 mmol/L, report to the nearest 0.1 mmol/L. Data on imprecision should be available at this level of detail to allow mathematical manipulation with rounding only at the final step. In the above example for serum Na a 95% significant change value would be 0.66 x 2.77 = 1.82 mmol/L and rounded up to 2 mmol/L for the user.

One can see the error if the SD is reported as "1 mmol/L" and then multiplied by 2.77 to give a significant change value of 3 mmol/L rather than the correct value of 2.

A second approach is to consider what information is 'lost' by the rounding process. As mentioned earlier, the effect of rounding can 'hide' an actual difference of 1.9 units. Thus, for a single point the appropriate reporting interval is $\sqrt{2} \times z \times \sigma/1.9$, where z is the z score at (1-p) confidence interval, and σ is the analytical SD.[1]

If we take into account the effect of sequential results, then there have been two different approaches reported in the literature. Badrick et al. calculated reporting intervals based only on the analytical imprecision of the method.[3] Thus reference intervals were determined without the SD_i term because we need to be able to report when two results are analytically different. The cause of this difference is significant in the interpretation of those results but biological variation may either increase or decrease in disease.

Another approach is to include biological variation in the calculation. Jones has reported the impact of using this calculation $(2.77 \times \sqrt{((SD_a)^2 + (SD_i)^2)^{1/2}})$[4] and has highlighted the impact of 'binning' results. Thus, at low concentrations, using a reference interval which is appropriate for the level of imprecision at high concentrations may lead to a loss of clinically useful information. Note that in the paper of Badrick et al., different concentrations with their respective SD_a were used rather than a constant imprecision value across the measurement range.

We have concentrated on the critical difference between two sequential results, but many results are singleton and are interpreted against a reference interval or clinical decision point (either empirical or based on evidence). Under these circumstances, the appropriate reference interval should be related to the imprecision of the measurement alone. Thus, the critical difference becomes $2.77 \times SD_a$, for a 95% confidence interval.[3]

This discussion assumes that the laboratory has complete control over the way it reports results and their significant figures. However, often this is not the case as there is an intervening Laboratory Information System which may or may not allow rounding or differential rounding based on a threshold concentration.

Results

In the table we have extracted some data from the papers of Hawkins et al.[1] and Badrick et al.[3] The table gives for a range of commonly measured analytes estimates of reporting interval based on different sources. The justifiable unit magnitude (using p=0.05) is calculated using SD for the 50th centile rankings of all laboratories enrolled in the General Serum Chemistry program from the Royal College of Pathologists of Australasia Quality Assurance Program Pty. Ltd. (RCPA QAP)[5] This represents current 'state of the art'. The table also gives the reporting unit intervals that are implicit from a well-known clinical chemistry textbook (Tietz Textbook of Clinical Chemistry[6]), a learned body (RCPA[7]), and from the data entry sheet for the General Serum Chemistry program from the RCPA QAP. The final column contains the theoretical reporting interval calculated from method of Badrick et al.[3]

It is apparent from the table that current laboratory reporting interval practice is inappropriate even when only analytical imprecision is considered. A more appropriate approach

Table. Justified reporting unit magnitude (p=0.05) based on performance in RCPA General Serum Chemistry QAP Cycle 68, together with unit size taken from various authorities.

Analyte	Units	Justifiable Unit magnitude from 50% QAP SD	RCPA Unit Size	Tietz Unit Size	QAP Unit Size	RI *
Ca	mmol/L	0.07	0.01	0.01	0.01	0.05
Creatinine	µmol/L	10.83	10	1	1	1
Ferritin	µg/L	30.86	1	1	1	5 @ 56 20 @ 415
Glucose	mmol/L	0.45	0.1	0.1	0.1	0.1 @ 3.4 0.5 @ 15.0
LD	U/L	21.84	1	1	1	5@ 115 10 @ 436
Urea	mmol/L	0.48	0.1	0.1	0.1	0.2 @ 5.5 0.5 @ 18.5
Cortisol	nmol/L	47.61	1	1	1	10@81 40@550

* RI according to Badrick et al. (Reference 3). Reporting Interval dependent on concentration or activity.

taking into account laboratory imprecision is seen from the final column.

Even when only analytical imprecision is considered it has been reported that many laboratories use an inappropriate number of significant figures or reporting interval.[1] There is a great need for portability of results and therefore agreement between laboratory information systems in reference intervals and reporting intervals. One of the purposes of the uncertainty of measurement exercise in laboratories should be to critically review the current number of significant figures reported by laboratories and to amend these based on the imprecision of the assay and the biological variation of the analyte. However, we suspect that many laboratories have not taken the opportunity to revise their reporting intervals. When a laboratory does report a result it must be aware that the number of significant figures reported should be carefully considered and be small in comparison to the imprecision and biological variation. We have summarised the current literature and strongly suggest that laboratories ensure that their reporting intervals are fit for the purpose of adding value and not confusion to the differential diagnosis.

Competing Interests: None declared.

References

1. Hawkins RC, Badrick T, Hickman PE. Over-reporting significant figures - a significant problem? Clin Chim Acta 2007;375:158-61.
2. Hawkins RC, Johnson RN. The significance of significant figures. Clin Chem 1990;36:824.
3. Badrick T, Wilson SR, Dimeski G, Hickman PE. Objective determination of appropriate reporting intervals. Ann Clin Biochem 2004;41:385-90.
4. Jones GR. Effect of the reporting-interval size on critical difference estimation: beyond "2.77". Clin Chem 2006;52:880-5.
5. Royal College of Pathologists of Australasia Quality Assurance Program Pty.Ltd (RCPA QAP) http://www.rcpaqap.com.au (Accessed 1 May 2008).
6. Burtis CA, Ashwood ER, editors. Tietz Textbook of Clinical Chemistry. 3rd ed. Philadelphia, USA: W.B. Saunders; 1999.
7. Royal College of Pathologists of Australasia. RCPA manual. Version 4.0. http://www.rcpamanual.edu.au (Accessed 26 August 2005).