Analytical Quality Specifications in the Clinical Laboratory – Limiting Analytical Error with Allowable Error Limits

Francisco Freitas[1,2]

[1]Departmento de Medicina Laboratorial, Centro Hospitalar Tondela-Viseu, Viseu, [2]Escola Superior de Saúde Dr. Lopes Dias, Instituto Politécnico de Castelo Branco

Resumo

A função dos laboratórios clínicos é fornecer informações para decisões médicas no tratamento de doentes. Essa informação deve ser tão correta quanto possível, e o erro associado limitado a valores aceitáveis tendo em conta o uso médico pretendido. O estabelecimento de limites de erro ou especificações de qualidade analítica baseia-se na hierarquia de consenso definida em Estocolmo (1999). Os laboratórios devem ter como meta a definição de especificações a partir de modelos mais elevados na hierarquia, sempre que existam dados disponíveis, e o desempenho atual dos métodos o permita. A qualidade dos processos de medição pode ser avaliada através da métrica Sigma.

Palavras-chave: Controlo de Qualidade • Especificações de Qualidade Analítica • Avaliação da Performance

Aceite para publicação: 12 de Novembro de 2013

Abstract

Clinical laboratories work to provide information for medical decisions on patient management. That information should be as correct as possible, and the associated error limited to a certain amount that is tolerable for medical purposes. A consensus hierarchy of models has been established in Stockholm (1999) to define such allowable error limits or analytical quality specifications. Laboratories should aim for higher models in hierarchy whenever data is available, and current method performance allows achieving the goal. Sigma-metrics can be used to assess the quality of measurement procedures.

Key-words: Quality Control • Analitycal Quality Specifications • Performance Evaluation

Accepted for publication: 12 November 2013

*** Correspondência:**

Francisco Freitas, Departamento de Medicina Laboratorial, Centro Hospitalar Tondela-Viseu, Av. Rei D. Duarte, 3504-509 Viseu, E-mail: francyske@gmail.com

Introduction

The current practice of Quality Control in laboratories has significantly changed in the past decade, and far are the days when only the so called Westgard Rules (1) were used as criteria to judge analytical runs and detect out of control events and methods instability. Today, it is important that laboratory professionals have up to date skills, and comprehend their importance in limiting analytical error and what boundaries should be established for unacceptable results based on method performance. The utmost priority of laboratories is to use the best samples to produce the most correct results with a limited amount of error.

Limiting Analytical Error

Great attention is being paid to Quality and Safety issues in healthcare (2), either by the regulatory bodies or the public opinion, and healthcare professionals, in particular laboratory personnel should be aware of that (3). Medical decisions often rely on laboratory tests information in about two-thirds and three-quarters of the cases (4), and because a fraction of diagnostic er-

Copyright © 2013, Sociedade Portuguesa de BioAnalistas Clínicos

rors still comes from laboratory testing (5) every effort to reduce them is worthwhile.

Two studies from Plebani and Carraro (6, 7) in a stat laboratory from Padua, Italy, have shown that in the total testing process the errors are more prone to occur in the pre-analytical phase (~65%), being followed by the post-analytical (~20%) and analytical phase (~15%). Strikingly, 73% of those were preventable, leaving room for improvement in laboratory error management.

Although currently the lowest error percentage lies on the analytical phase, we must be aware that the assessment of those errors is more difficult, and in instances somewhat neglected by professionals. Indeed, the International Federation of Clinical Chemistry (IFCC) Working Group on "Laboratory Errors and Patient Safety" developed a series of Quality Indicators (QI's) and specifications for the laboratory total testing process, and those for the analytical phase were the least standardized at first (8), but that model has evolved to include seven QI's (9), from which independent quality specifications can be derived.

Having that in mind, we must also highlight that the Clinical Laboratory is one of the leading health services in addressing Quality issues (10), and the industry has greatly contributed with improved instruments and technology, that limits result variation by reducing human tasks and associated errors, and now tend to meet the highest performance standards on allowable bias (or inaccuracy), imprecision and total error (11).

Nowadays, Quality of care can be intended as "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes, and are consistent with current professional knowledge, meeting the expectations of healthcare users" (12). So, being the focus on the client or user, the requirements (patient and medical ones) should be determined by the laboratory (13), and a normalized process based on a management system should be in place in order to consistently achieve the desired results (14).

What Allowable Error Limits?

The Internal Quality Control process, which is time and skilful demanding, and staff training is a key factor to success due to the amount of tests performed, is the moment where the quality that is required (numerical analytical quality specification – mostly defined as a %TEa – Allowable Total Error) is usually determined, and that ultimately guides the establishment of a statistical quality control procedure (number of measurements and control rules) (4). In general, TEa sets a limit for both the imprecision (random error) and bias (systematic error) that are tolerable in a single measurement or single test result (4). It can

Table 1 – Stockholm Consensus hierarchy of models to set analytical quality specifications.

Level	Model	Source examples
1.	*Evaluation of the effect of analytical performance on clinical outcomes in specific clinical situations*	Diabetes Control and Complications Trial – DCCT
2.	*Evaluation of the effect of analytical performance on clinical decisions in general*	
2.A	Data based on the components of biological variation	Biological variation database
2.B	Data based on analysis of clinicians opinions	Surveys on patient treatment by doctors
3.	*Published professional recommendations*	
3.A	From national and international expert bodies	NCEP[a]; ADA[b]; NACB[c]
3.B	From expert local groups or individuals	Best practice, good laboratory practice
4.	*Performance goals*	
4.A	Set by Regulatory bodies	CLIA[d] – USA, Rilibäk[e] – Germany
4.B	Set by Quality Assessment (EQA) schemes	NEQAS[f], RIQAS[g], RCPA[h]
5.	*Goals based on the current state of the art*	
5.A	As demonstrated by data from EQA or Proficiency Testing Schemes	MQS Spain[i], Unity Inter-laboratory Program
5.B	As found in current publications on methodology	Evaluation of methods

[a]National Cholesterol Education Program; [b]American Diabetes Association; National Academy of Clinical Biochemistry[c]; [d]Clinical Laboratory Improvement Ammendment; [e]Guidelines ("Rili") of the German Federal Medical Council (BÄK); [f]United Kingdom National External Quality Assessment Service; [g]Randox International Quality Assessment Scheme; [h]Royal College of Pathologists of Australasia; [i]Minimum quality specifications.

Copyright © 2013, Sociedade Portuguesa de BioAnalistas Clínicos

also be viewed as the magnitude of error that causes a result to be considered incorrect or unreliable (14).

These analytical quality specifications or analytical goals as usually named (15), are to be set based on the "Stockholm Consensus Hierarchy of models" (Table 1) established in the 1999 Stockholm Consensus Conference on Quality Specifications in Laboratory Medicine (16).

The hierarchy recommends that higher level models should be preferred to those at lower levels, providing that goals are available and appropriate to a specific analyte (15). These models can be split into different approaches (17), such as:

1. Clinical-driven goals – This means that the actual clinicians behavior using test results should be studied, to derive the size of an analytical error (clinical decision interval) that could change a medical decision on a specific therapy (i. e. Diabetes Control);
2. Biology-driven goals – This means ignoring current method performance, but establishing how good test methods performance should be, based on within and between-subject biologic variation;
3. Consensus-driven goals – This means establishing a goal that the majority of labs and methods can achieve, based on current method performance.

In current practice, laboratories should aim first at consensus goals, which are best attained, though less demanding, but mandatory in some countries (i.e. CLIA – USA, RILIBÄK – Germany) for Accreditation purposes, and then establish more demanding ones (in general biology and clinical use goals) if they are known and documented (17, 18) as shown in the examples of Table 2. Apart from %TEa, concentration based limits with different specifications given at different medical decision levels, and imprecision based limits (i.e. 3SD) specifications can also be set.

In Czech Republic EQA program Friedecky *et al* (19) found that 98% of the labs could meet Rilibäk and SEKK (Czech Republic) acceptance limits, 87% the CLIA limits, 72% the RCPA limits, and only 22% the biological variation limits.

Despite the above results, the biological variation model (Level 2A) is now widely used (20) and considered the best for practical purposes (monitoring individual patients, and diagnosis using reference intervals) (21), and the database that is updated every two years, currently comprises more than 320 analytes (18, 22). Due to the large or small biological variation of some analytes, the desirable analytical goal can be easily attained (CPK) or hardly met (Sodium) with current technology, so optimum and minimum specifications (Table 2) can also be defined by laboratories for bias, imprecision and TEa (22), based on the formulas proposed by Fraser *et al* (23).

In Spain, EQA results have shown that the BV total error specification for albumin, HDL-cholesterol, sodium and calcium was attained by close to 10%, 20%, 50% and 70% of labs respectively (22), somewhat demystifying its unreachable character. It is also of note that analytical goals for levels 3 and 4 of the hierarchy are also being set by the organizations involved based on Level 2A (15, 24).

Performance Evaluation

The way labs attain the analytical goals of a specific analyte will characterize its performance. One way is by estimating the Total Error of the method and verifying that it is smaller than the defined requirement as TEa (25):

Table 2 – Analytical quality specifications for cholesterol, glucose, creatine kinase and sodium.

Specification Source	Analytical Quality Specification			
	Cholesterol	Glucose	Creatine kinase	Sodium
Rilibäk (Germany)	13%	15%	20%	5%
Minimum QS (Spain)	11%	11%	24%	5%
BV Minimum goal	12.73%	10.78%	45.5%	1.33%
CLIA (USA)	10%	10%	± 30%	± 4 mmol/L
BV Desirable goal	8.5%	7.2%	± 30.3%	0.9%
RCPA (Australia)	6%	8%	12%	± 3 mmol/L ≤ 150 mmol/L ± 2% > 150 mmol/L
BV Optimum goal	4.24%	3.60%	15.2%	0.45%

Copyright © 2013, Sociedade Portuguesa de BioAnalistas Clínicos

$$TE_{95\%CI} = Bias\% + (1.65xCV\%) < TEa$$

Then, the lab can also monitor method improvement by calculating TE budget percent:

[TE / TEa] x 100

Another approach is Sigma-metrics (six sigma analytical performance) which is calculated as follows (4):

$$\sigma = (\%TEa - \%BIAS_a)/ \%CV_a$$

The resulting Sigma value, characterizes method performance in terms of the number of standard deviations or sigmas that fit within the tolerance limit or quality requirement of a test, and also gives us a notion of the number of defects (possible errors) per million tests we are delivering in the analytical process (26).

As shown in Table 3 the higher the sigma level, the better the test performance, being 3-sigma quality usually recommended as the minimum for a production process (26), but in many cases this are problematic tests requiring maximum statistical quality control procedures (13).

The achieved performance is greatly influenced by the defined tolerance limit (TEa), and here there's a need for harmonization of practice since significant differences exist between and within countries (14), making comparison of performance data unreliable. Sigma performance is also different at different medical decision levels, being usually better at higher concentrations, so at lower levels is where we should find space for method improvement.

By knowing the estimates of method bias and precision and having defined TEa, the most appropriate statistical quality control procedure (number of measurements and control rules) can be selected that provides at least 90% probability to detect clinically significant errors, and a maximum of 5% chance of false rejections (4). Not being the scope of this article to explain such a process, a general picture of QC frequency based on Sigma value has been proposed (13):

>6 σ (excellent tests) – evaluate with one QC per day (alternating levels between days) and a 1:3.5 s rule.

4s–6 σ (suited for purpose) – evaluate with two levels of QC per day and the 1:2.5 s rule.

3s–4 σ (poor performers) – use a combination of rules with two levels of QC twice per day.

<3 σ (problematic tests) – maximum QC, three levels, three times a day. Consider testing specimens in duplicate.

Several software tools like UnityRealTime® (Bio-Rad, Hercules, USA), Accusera 24.7® (Randox, Crumlin, UK), are now in the market to assist the Clinical Laboratory in Quality control design and evaluation. These tools which are part of inter-laboratory programs allow real-time peer comparison of QC data from thousand of Labs around the world, faster detection of errors in the test systems and increased confidence on selected comparison targets. It is also a step for meeting the requirements of medical laboratories accreditation, which currently is only mandatory in France at the European level (14).

Table 3 – Sigma conversion chart.

Sigma Level	Defects/million tests	% Defect	Sigma Level	Defects/million tests	% Defect
6	3	0,00034%	3	66,807	6,7%
5.5	32	0,0032%	2.5	158,655	15,86%
5	233	0,023%	2.0	308,538	31%
4.5	1,350	0,135%	1.5	500,000	50%
4	6,210	0,62%	1.0	691,462	69%
3.5	22,750	2,275%	0.5	841,345	84.1%

Final Considerations

Every lab result always includes error. Our job as laboratory professionals is to limit that error to a tolerable amount that is not clinically significant for patient management. Strictly adherence to good laboratory practices, from specimen collection, calibration procedures, planning and implementation of proper internal quality control for each analyte (including definition of analytical quality specifications), timely and accurate result delivery will certainly impact on the magnitude of those errors, and provide doctors best clinical decisions and patient improvement.

Copyright © 2013, Sociedade Portuguesa de BioAnalistas Clínicos

References

1. Westgard JO, Barry PL, Hunt MR. A Multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem* 1981; 27:493-501.
2. Kohn KT, Corrigan JM, Donaldson MS. To Err Is Human: Building a Safer Health System. Washington, DC: National Academy Press, 1999:287pp.
3. Portugal. Plano estratégico de Saúde 2012-2016 – 3.3 Eixo Estratégico – Qualidade em Saúde. Lisboa: DGS, 2012.
4. Westgard JO. Internal quality control: planning and implementation strategies. *Ann Clin Biochem* 2003; 40(Pt 6):593-611.
5. Plebani M. Exploring the iceberg of errors in laboratory medicine. *Clin Chim Acta* 2009; 404(1):16-23.
6. Plebani M, Carraro P. Mistakes in a stat laboratory: types and frequency. *Clin Chem*.1997; 43(8 Pt 1):1348-51.
7. Carraro P, Plebani M. Errors in a stat laboratory: types and frequencies 10 years later. *Clin Chem* 2007; 53(7):1338-42.
8. Sciacovelli L, O'Kane M, Skaik YA, Caciagli P, Pellegrini C, Da Rin G, Ivanov A, Ghys T, Plebani M; IFCC WG-LEPS. Quality Indicators in Laboratory Medicine: from theory to practice. Preliminary data from the IFCC Working Group Project "Laboratory Errors and Patient Safety". *Clin Chem Lab Med* 2011; 49(5):835-44.
9. Plebani M, Chiozza ML, Sciacovelli L. Towards harmonization of quality indicators in laboratory medicine. *Clin Chem Lab Med* 2013; 51(1):187-95.
10. Lippi G, Plebani M, Šimundić AM. Quality in laboratory diagnostics: from theory to practice. *Biochemia Medica* 2010; 20(2):126-30.
11. Zaninotto M, Plebani M. The "hospital central laboratory": automation, integration and clinical usefulness. *Clin Chem Lab Med* 2010; 48(7):911-7.
12. Buttell P, Hendler R, Daley J. Quality in healthcare: concepts and practice. Westport: Praeger, 2008:62-83.
13. Cooper G, DeJonge N, Ehrmeyer S, Yundt-Pacheco J, Jansen R, Ricós C, Plebani M. Collective opinion paper on findings of the 2010 convocation of experts on laboratory quality. *Clin Chem Lab Med* 2011; 49(5):793–802.
14. Burnett D, Ceriotti F, Cooper G, Parvin C, Plebani M, Westgard J. Collective opinion paper on findings of the 2009 convocation of experts on quality control. *Clin Chem Lab Med* 2010; 48(1):41-52.
15. Petersen H, Fraser CG. Strategies to set global analytical quality specifications in laboratory medicine: 10 years on from the Stockholm consensus conference. *Accred Qual Assur* 2010; 15:323–30.
16. Petersen PH, Fraser CG, Kallner A, Kenny D (eds). Strategies to set global analytical quality specifications in laboratory medicine. *Scan J Clin Lab Invest* 1999; 59:475–585.
17. Westgard S. What's the right goal? 2011; www.westgard.com/right-quality-goal.htm.
18. Ricos C, Alvarez V, Cava F, Garcia-Lario JV, Hernandez A, Jimenez CV, *et al*. Desirable analytical quality specifications for imprecision, bias and total error upon biological variation. The 2012 update. www.westgard.com/biodatabase1.htm.
19. Friedecky B, Kratochvila J, Budina M. Why do different EQA schemes have apparently different limits of acceptability? *Clin Chem Lab Med* 2011; 49(4):743-5.
20. Petersen PH, Sandberg S, Fraser CG. Do new concepts for deriving permissible limits for analytical imprecision and bias have any advantages over existing consensus? *Clin Chem Lab Med 2*011; 49(4):637–40.
21. Fraser CG, Petersen PH. Analytical performance characteristics should be judged against objective quality specifications. *Clin Chem* 1999; 45(3):321-3.
22. Ricos C, Perich C, Minchinela J, Alvarez V, Simon M, Biosca C *et al*. Application of biological variation - a review. *Biochemia Medica* 2009; 19(3):250-9.
23. Fraser CG, Hyltoft Petersen P, Libeer JC, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. *Ann Clin Biochem* 1997; 34:8-12.
24. Jones GRD, Sikaris K, Gill J. Allowable limits of performance for External Quality Assurance Programs - an Approach to application of the Stockholm criteria by the RCPA Quality Assurance Programs. *Clin Biochem Rev* 2012; 33(4):133-9.
25. Westgard JO, Carey RN, Wold S. Criteria for judging precision and accuracy in method development and evaluation. *Clin Chem* 1974; 20(7):825-33.
26. Westgard JO, Westgard SA. The quality of laboratory testing today - An assessment of σ metrics for analytic quality using performance data from proficiency testing surveys and the CLIA criteria for acceptable performance. *Am J Clin Pathol* 2006; 125:343-54.

Copyright © 2013, Sociedade Portuguesa de BioAnalistas Clínicos