

Opinion Paper

Callum G. Fraser*

The 1999 Stockholm Consensus Conference on quality specifications in laboratory medicine

DOI 10.1515/cclm-2014-0914

Received September 17, 2014; accepted January 1, 2015; previously published online February 7, 2015

Abstract: The setting of analytical quality specifications in laboratory medicine has been a topic of discussion and debate for over 50 years: 15 years ago, as the subject matured and a profusion of recommendations appeared, many of them from expert groups, it was realised by a number of leading professionals that there was a need for a global consensus on the setting of such specifications. The Stockholm Conference held in 1999 on “Strategies to set global analytical quality specifications in laboratory medicine” achieved this and advocated the ubiquitous application of a hierarchical structure of approaches. The hierarchy has five levels, namely: 1) evaluation of the effect of analytical performance on clinical outcomes in specific clinical settings; 2) evaluation of the effect of analytical performance on clinical decisions in general using a) data based on components of biological variation, or b) analysis of clinicians’ opinions; 3) published professional recommendations from a) national and international expert bodies, or b) expert local groups or individuals; 4) performance goals set by a) regulatory bodies, or b) organisers of external quality assessment (EQA) schemes; and 5) goals based on the current state of the art as a) demonstrated by data from EQA or proficiency testing scheme, or b) found in current publications on methodology. This approach has been much used since its wide promulgation, but there have been ongoing criticisms and new developments. The time seems right for an objective reappraisal of recommended strategies to set analytical performance goals.

Keywords: analytical goals; analytical performance goals; analytical quality specifications; biological variation; clinical opinions; professional bodies; quality assessment; state of the art.

Background

It has been widely realised for many years that it is impossible, and rather non-productive, to discuss quality in laboratory medicine unless analytical quality specifications (analytical goals, or analytical performance goals) are set *a priori* [1]. In addition, it is recognised that objective analytical quality specifications must be attained so that adequate patient care is provided [2]. Such specifications are required for many purposes, including: 1) when drawing up specification documents for new analytical methodology or equipment and short listing potential candidates, in evaluation studies to decide whether the found performance characteristics of a method are satisfactory and to set up quality control, assurance and planning in individual laboratories; 2) to assist the organisers of external quality assessment schemes (EQAS) and proficiency testing (PT) programmes to design and deliver appropriate monitoring of performance; 3) to help the *in vitro* diagnostics industry, the manufacturers of instruments and reagents, in design, construction and marketing; and 4) to encourage laboratories to decide which particular examinations are less than satisfactory and require expenditure of scarce resources on improvement [3].

Towards the end of the 1990s, the ongoing and growing discussion of the advantages and disadvantages of strategies to set analytical quality specifications became much more focussed. At that time, it was widely recognised that much work on the documentation of strategies to set analytical quality specifications had been performed. Early proposals had been made based on a quarter of the reference interval [4], set from what were stated to be opinions of clinicians [5] and derived from the relevant components of biological variation [6]. Since these proposals and the considerable body of other early work that has been reviewed [7] appeared, a flow of publications ensued, which have also been critically summarised [8, 9]. In addition, expert conferences had addressed the topic, including those organised by the College of American Pathologists [10] and the then American Association for Clinical Chemistry [11]. Moreover, national and international professional groups had addressed analytical quality specifications for a large number of individual

*Corresponding author: Callum G. Fraser, PhD, Centre for Research into Cancer Prevention and Screening, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, Scotland, UK, Phone: +44 1382553799, Fax: +44 1382425679, E-mail: callum.fraser@nhs.net

quantities of, at the time, topical clinical interest, such as cholesterol [12]. Further, more general approaches had been documented, including analytical quality specifications for use in evaluation of instruments [13], routine laboratory medicine [14] and reference methodology [15]. Finally, in early 1999, the need for ubiquitous application of analytical quality specifications was emphasised yet again and a proposal published for a new concept based on a hierarchical classification of available strategies [16].

However, in spite of this large body of work, widely held agreement had not been achieved on the best strategies to use to set analytical quality specifications. Thus, it seemed to a number of leading professionals in laboratory medicine that the time was ripe to agree on global strategies to set analytical quality specifications which could be applied ubiquitously. This view was much stimulated by the somewhat contentious proposals in drafts generated by Working Group 3 of ISO Technical Committee (ISO/TC) 212, Clinical Laboratory Testing and In Vitro Diagnostic Test Systems, which were undertaking the preparation of a standard or guide on how to define analytical quality specifications. When a draft was presented external to ISO/TC 212 in 1998, discussion and input from a wide range of professionals occurred. However, in view of the further controversial work done in the plans to create what might become ISO 15196, the International Union of Pure and Applied Chemistry (IUPAC) Clinical Chemistry Section, International Federation of Clinical Chemistry (IFCC) and World Health Organization (WHO) then initiated a “consensus conference” in an attempt to proclaim strategies to set global quality specifications in laboratory medicine: the outcomes of that conference are described here.

The consensus conference

The conference was held in the Nobel Forum, Stockholm, Sweden, 25–26 April 1999, and was considered at the time to be most successful. Over 100 participants from 27 countries actively contributed to the discussions on the 22 formal presentations, which were given by professionals in laboratory medicine who had contributed to the peer-reviewed literature on the setting of analytical quality specifications. Along with a foreword and introduction, the formal contributions and the consensus statement were documented in a special issue of the *Scandinavian Journal of Clinical and Laboratory Investigation* [17]. The main aim of the organisers of the conference was to provide a forum for reaching consensus on the setting of global analytical quality specifications in laboratory medicine. The organisers considered that the objective had

been achieved, since agreement seemed complete on the principles laid down in the following widely reproduced consensus statement [18].

The consensus statement

The main outcome of the consensus conference was agreement that the following hierarchy of models should be applied to set analytical quality specifications.

1. Evaluation of the effect of analytical performance on clinical outcomes in specific clinical settings.
2. Evaluation of the effect of analytical performance on clinical decisions in general:
 - a. Data based on components of biological variation,
 - b. Data based on analysis of clinicians’ opinions.
3. Published professional recommendations:
 - a. From national and international expert bodies,
 - b. From expert local groups or individuals.
4. Performance goals set by:
 - a. Regulatory bodies,
 - b. Organisers of EQA schemes.
5. Goals based on the current state of the art:
 - a. As demonstrated by data from EQA or Proficiency Testing schemes,
 - b. As found in current publications on methodology.

Where available, and when appropriate for the intended purpose, models higher in the hierarchy are to be preferred to those at lower levels.

The following matters were also discussed and agreed.

The above hierarchy includes currently available models; however, new useful concepts will undoubtedly evolve. Implementation of any of the models should use well-defined and described procedures.

To facilitate the future debate on the setting of analytical quality specifications, there is a need for agreement on concepts, definitions and terms.

There is a need for continuous improvement in the exchange of information on quality issues: between clinical laboratory professionals and the diagnostics industry, and between clinical laboratory professionals and the users of the laboratory service.

Application, caveats, progress and the future

As described earlier, analytical quality specifications have many uses, including in: 1) evaluation and introduction of

new analytical methodology or equipment; 2) setting limits for acceptable performance in EQAS and PT; 3) helping the in vitro diagnostics industry; and 4) assisting laboratories to decide which examinations require improvement [3]. The germane question is whether publication of the consensus statement has led to outcomes which have improved laboratory medicine. There are anecdotal indications that laboratories do detail objective analytical quality specifications in their procurement documents. There is evidence that the hierarchy has become more used in EQAS over time, e.g., the Royal College of Pathologists of Australasia Quality Assurance Programme Chemical Pathology states that: the allowable limits of performance (ALP) have been set using the Stockholm criteria hierarchy [19]. Regrettably, many EQAS and PT still use the state of the art (Level 5), the lowest level in the hierarchy, as criteria for acceptable analytical performance. Although there have been very few detailed studies on effect of analytical performance on clinical outcomes, many recommendations disseminated recently on analytical quality specifications from professional bodies (Level 3) are based on biological variation (Level 2a) rather than on subjective opinions. Unfortunately, most major equipment and reagent manufactures do not seem to include clinically-based analytical quality specifications for their products. In addition, most regulatory authorities generally do not require that analytical systems meet a priori set specified quality requirements. It is hoped that laboratories do use objective criteria based on the consensus statement, not only to decide which of their repertoire of examinations requires improvement or change, but also in the quality planning strategies that require analytical quality specifications to enable the number of controls to be analysed and the control rules used to be set objectively [20].

The consensus statement has been cited more than 130 times to date (according to Google scholar – <http://scholar.google.co.uk/>) and wide acceptance of the principles has actually occurred as shown, e.g., in the detailed discussions held at three recent convocations of experts on quality in laboratory medicine [21–23]. There seems a particular acceptance of the use of Level 2a of the hierarchy, namely the use of numerical estimates of the components of within-subject and between-subject biological variation. This is in spite of criticisms, certain of which do have some validity, as shown in an excellent recent systematic review of data on biological variation for alanine aminotransferase, aspartate aminotransferase and γ -glutamyl transferase [24]. This is probably in large part due to the establishment and ongoing updates of a database on biological variation that includes desirable analytical quality specifications for imprecision, bias

and total allowable error, now in its eighth edition [25]. In addition, the proposals made for setting analytical quality specifications solely based on biology, presenting three categories for both imprecision and bias of minimum, desirable and maximum [26], seem to have gained considerable acceptance. Interestingly, in spite of the fact that the views of TC212 seem to have matured into support for the hierarchical approach to setting analytical quality specifications [27], at a meeting of ISO/TC 212 in Australia in 2003, it was stated (sic): One project, ISO 15196 on performance goals, has been cancelled, with the expectation that WG3 will reconsider the need for the project and reaffirm its scope; if deemed appropriate by the TC, a new work item proposal will be circulated for vote [28]. This does not seem to have occurred.

There was a clear realisation that the hierarchy of strategies to set analytical quality specifications was not finality and this was restated in a review of progress made in the 10 years following the consensus conference [29]. Indeed, work has not ceased and new models, or developments of more traditional models, have been developed, including those by Haeckel and Wosniok [30] and Klee [31]. These have been summarised and evaluated [32]: it was considered they were certainly in keeping with the views of the consensus conference that the hierarchy could be modified in the future if significant new evidence-based ideas were developed and used in practice.

Laboratory medicine has changed significantly over the last 15 years. Since there has been wide application, but with some caveats, and some progress, a reappraisal of the 1999 consensus statement might be considered an essential prerequisite to the evolution of quality management in laboratory medicine. In consequence, the first European Federation for Clinical Chemistry and Laboratory Medicine (EFLM) Strategic Conference “Defining Analytical Performance Goals – 15 years after the Stockholm Conference” was held in Milan on 24–25 November, 2014. The contributions from experts and the consensus statement are documented in this issue of the journal. I shall wait with great interest, as should others, to assess if the new consensus approaches becomes widely used and the hierarchy of choice to set analytical quality specifications.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Financial support: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Competing interests: The funding organisation(s) played no role in the study design; in the collection, analysis, and

interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

- Kallner A, McQueen M, Heuck C. Foreword. The Stockholm Consensus Conference on quality specifications in laboratory medicine, 25–26 April 1999. *Scand J Clin Lab Invest* 1999;59:475–6.
- Fraser CG, Kenny D, Kallner A, Hyltoft Petersen P. Introduction: strategies to set global quality specifications in laboratory medicine. *Scand J Clin Lab Invest* 1999;59:477–8.
- Fraser CG. Biological variation: from principles to practice. Washington, DC: AACC Press, 2001.
- Tonks DB. A study of the accuracy and precision of clinical chemistry determinations in 170 Canadian laboratories. *Clin Chem* 1963;9:217–33.
- Barnett RN. Medical significance of laboratory results. *Am J Clin Pathol* 1968;50:671–6.
- Cotlove E, Harris EK, Williams GZ. Biological and analytic components of variation in long-term studies of serum constituents in normal subjects. III. Physiological and medical implications. *Clin Chem* 1970;16:1028–32.
- Fraser CG. Desirable performance standards for clinical chemistry tests. *Adv Clin Chem* 1983;23:299–339.
- Fraser CG, Hyltoft Petersen P. Desirable standards for laboratory tests if they are to fulfil medical needs. *Clin Chem* 1993;39:1447–55.
- Fraser CG. Quality specifications in laboratory medicine. *Clin Biochem Rev* 1996;17:109–14.
- Elevitch FR, editor. College of American Pathologists Conference II (1976): analytical goals in clinical chemistry. Skokie, IL: College of American Pathologists, 1977.
- Rej R, editor. 1992 Clinical Chemistry Forum. Accuracy and precision goals in clinical chemistry testing: can they be defined by medical relevance. *Clin Chem* 1993;39:1446–544.
- National Cholesterol Education Program Laboratory Standardization Panel. Current status of blood cholesterol measurement in clinical laboratories in the United States. *Clin Chem* 1988;34:193–201.
- Fraser CG, Hyltoft Petersen P, Ricos C, Haeckel R. Proposed quality specifications for the imprecision and inaccuracy of analytical systems in clinical chemistry. *Eur J Clin Chem Clin Biochem* 1992;30:311–7.
- Stockl D, Baadenhuisjen H, Fraser CG, Libeer JC, Hyltoft Petersen P, Ricos C. Desirable routine analytical goals for quantities assayed in serum. *Eur J Clin Chem Clin Biochem* 1995;33:157–69.
- Thienpont L, Franzini C, Kratochvila J, Middle J, Ricos C, Siekmann L, et al. Analytical quality specifications for reference methods and operating specifications for networks of reference laboratories. *Eur J Clin Chem Clin Biochem* 1995;33:949–57.
- Fraser CG, Hyltoft Petersen P. Analytical performance characteristics should be judged against objective quality specifications. *Clin Chem* 1999;45:321–3.
- Hyltoft Petersen P, Fraser CG, Kallner A, Kenny D, editors. Strategies to set global analytical quality specifications in laboratory medicine. *Scand J Clin Lab Invest* 1999;59:475–585.
- Kenny D, Fraser CG, Hyltoft Petersen P, Kallner A. Consensus agreement. *Scand J Clin Lab Invest* 1999;59:585.
- RCPAQAP Chemical Pathology. Assessment of performance. Available from: www.rcpaqap.com.au/chempath/assessment-of-performance/. Accessed on 22 October, 2014.
- Westgard JO, Westgard S. Basic quality management systems. Madison, WI: Westgard QC, Inc., 2014.
- 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;48:41–52.
- Cooper G, DeJonge N, Ehrmeyer S, Yundt-Pacheco J, Jansen R, Ricós C, et al. Collective opinion paper on findings of the 2010 convocation of experts on laboratory quality. *Clin Chem Lab Med* 2011;49:793–802.
- Adams O, Cooper G, Fraser C, Hubmann M, Jones G, Plebani M, et al. Collective opinion paper on findings of the 2011 convocation of experts on laboratory quality. *Clin Chem Lab Med* 2012;50:1547–58.
- Carobene A, Braga F, Roraas T, Sandberg S, Bartlett WA. A systematic review of data on biological variation for alanine aminotransferase, aspartate aminotransferase and γ -glutamyl transferase. *Clin Chem Lab Med* 2013;51:1997–2007.
- Desirable biological variation database specifications. Available from: <http://www.westgard.com/biodatabase1.htm>. Accessed on 1 September, 2014.
- 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.
- Kaplan LA. Determination and application of desirable performance goals: the ISO/TC 212 approach. *Scand J Clin Lab Invest* 1999;59:479–82.
- ISO/TC 212 N116. Meeting Summary. ISO/TC 212, Clinical laboratory testing and in vitro diagnostic test systems. Sydney, Australia, 19 and 21 May 2003.
- Hyltoft Petersen P, 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.
- Haeckel R, Wosniok W. A new concept to derive permissible limits for analytical imprecision and bias considering diagnostic requirements and technical state-of-the-art. *Clin Chem Lab Med* 2011;49:623–35.
- Klee GG. Establishment of outcome-related analytic performance goals. *Clin Chem* 2010;56:714–22.
- Hyltoft Petersen P, 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* 2011;49:637–40.