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.
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.

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