Report from the first EFLM Strategic Conference

By Maria Stella Graziani
Chair of the EFLM Communication Committee

The first EFLM Strategic Conference "Defining analytical performance goals 15 years after the Stockholm Conference" held in Milan (Italy) last November (24th - 25th) was very successful, with 241 participants (speakers included) from 41 different Countries, embracing the five Continents. The delegates came from the clinical laboratories, from EQAs provider, from professional organizations and from the manufacturers as well, demonstrating that the topic was of interest for the whole world of Laboratory Medicine.

The Stockholm Conference in 1999 was a landmark in trying to achieve a consensus on how quality requirements should be set and a hierarchy of models was established. Time has come to revisit this hierarchy (see Table 1), investigating to what extent it is still valid or if it should be modified or expanded, incorporating performance goals for qualitative tests and for the whole testing process (pre- and post - analytical aspects included) .

Table 1 *The Consensus Hierarchy from Stockholm Conference*

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. Organizers of External Quality Assessment (EQA) schemes
5. Goals based on the current state of the art:
a. As demonstrated by data from EQA or Proficiency Testing scheme
b. As found in current publications on methodology

According to Sverre Sandberg, Co-Chair with Mauro Panteghini of the Conference, the ultimate goal of the meeting was to propose models to set analytical performance goals and to present these models in a new Consensus Document .

The Conference comprised 5 Sessions. The first day Sessions N 1, 2 and 3 examined the possibility and the pros and cons to base the performance criteria on clinical needs, on biological variation data, on state of the art, respectively. The second day, Session N 4 discussed the setting of performance criteria in different situations, like the performance criteria needed for reference measurement procedures and reference materials preparation, or which criteria should be established in the implementation of metrological traceability (uncertainty budget). Furthermore, performance criteria for internal QC and EQA schemes and for qualitative tests have been examined. Session 4 included also a summary on which models should be used to derive performance criteria for laboratory tests.

The last session was dedicated to the performance criteria and quality indicators to be adopted for the pre- and post-analytical phases.

At the end of the Conference, Sverre Sandberg summarized the content of the different presentations and illustrated the Consensus Document that was elaborated and approved by the Scientific Committee before the Conference and that was distributed to all the participants at the end of day 1, for further discussion.

Input to this document was given during and after the conference, and the document will be amended according to this before being published in Clinical Chemistry and Laboratory Medicine.

After two days of brilliant presentations and fruitful discussions, a number of issues are secured :

- The essence of Stockholm hierarchy is still valid, although new perspectives were forwarded inducing cautious modifications and explanatory additions.
- Three main models are available to set performances goals: some of these are better suited for certain measurands than for others:
  - o **Model 1.** Based on the effect of analytical performance on clinical outcomes. This model is the most rationale since it is based on the actual clinical outcome; however, in practice it is applicable only to a few tests since it is difficult to show the direct effect of laboratory tests on medical outcome. Different options are available:
    - ▪ Use the results of the outcome studies: how the analytical performance influences the clinical outcome
    - ▪ Investigate by a simulation study the impact of the analytical performance on the probability of clinical outcomes
    - ▪ Use surveys of clinicians or "experts" opinions to check the impact of analytical performance on medical decisions
  - o **Model 2.** Based on components of biological variation of the measurand. This model seeks to minimize the ratio of the analytical noise to the biological signal. Its applicability can however be limited by the validity and robustness of the data on biological variation.
  - o **Model 3.** Based on the state of the art. This model is the one where data is most easily available. It is linked to the highest level of analytical quality achievable with the currently available techniques. If the best laboratories can achieve only a "certain" analytical quality not at the level required by Models 1 and 2, then manufacturers have to strive to develop better assays. On the other hand, if the majority of laboratories can achieve the analytical quality required by Model 1 and 2, then the laboratories that cannot, have to change their practice. With this model, however, there is no link or only a weak one between what is technically achievable at present and what is needed to obtain a better outcome for the patient (Model 1) or to minimize the ratio of the analytical noise to the biological signal (Model 2). For this reason, Model 3 is the least preferable model.
- The three models use different principles; the hierarchy applies only when high quality studies are available for each model. If the studies are of insufficient quality, then we have to be prepared to shift to a model where better data are available.
- Some measurands could have different performance goals dependent on the different clinical applications of the test. A typical example is the blood glucose that can be used in critical care units, for self-monitoring, for the diagnosis of glucose intolerance in outpatients.
- Concerning the performance goals for the pre- and post-analytical phases, it is time to go further and to include performance goals for these extra-analytical phases. The criteria should ideally follow the same models as those of the analytical phase.

To reinforce and continue the activities arisen from the Conference, EFLM is now launching a Task Force on "Performance goals in Laboratory Medicine" (TF-PG). The Task Force comprises 5 groups (see Fig 1); these will have the structure of Task and Finishing Groups (TFG) and preferentially will end their work within two years. The names and the terms of reference of the TFG are the following:

Performance criteria models for specific laboratory tests : Aim is to allocate different tests to the different models described above, and to give an overview and a reason for why tests are allocated to different models.

Harmonization of allowable limits in EQAS : Aim is to define performance criteria for the most common analytes that can be used by EQAS organizers .

Measurement total error : Aim is to define how to use the total error concept or if it should be used at all (how can performance criteria for bias and imprecision be combined into performance criteria for total error?).

Performance criteria for pre- and post-analytical phases: Aim is to generate performance criteria for the pre- and post-analytical phases (and the total measurement process).

Biological variation database: Aim is to use a critical appraisal list to evaluate literature on biological variation and to generate a database containing, for each analyte, essential summary information from selected papers, so that it can be used for setting performance criteria based on biological variation.

Figure 1. *Structure of the EFLM Task Force*