This 1992 Clinical Chemistry Forum asks if accuracy and precision goals for the laboratory can be specified by reference to medical requirements, which are a function of the total laboratory testing process as well as the needs of patients, clinicians, and societal institutions. Furthermore, medical decisions and decisions about medical requirements must be made in situations of uncertainty and thus are subject to predictable cognitive errors. The interaction of all these factors must be considered by patients, laboratorians, and clinicians to identify practical and effective performance goals. Pathologists and clinical chemists are uniquely trained to identify thoughtful clinicians who are knowledgeable in quantitative judgment to participate in this goal-setting endeavor. It is time for these parties to accumulate the available data and engage in the synthesis of effective performance goals.

Indexing Terms: cognition • decision making • uncertainty • analytical goals

This 1992 Clinical Chemistry Forum poses the question, "Can laboratory performance goals for precision and accuracy be determined by medical relevance?" The question implies that the medical requirements for precision and accuracy can be specified on a test-by-test basis. For some tests, this may be possible in the general statistical sense; for others, it may require setting goals for specific clinical situations. Specific examples have been addressed earlier in the Forum, e.g., "What sensitivity, precision, and accuracy for measurements of serum thyrotropin are necessary to adequately demonstrate pituitary suppression in a thyroid-cancer patient receiving exogenous thyroxine?" (1). Here, I outline various nonanalytical factors to be considered in goal setting.

The 1992 Clinical Chemistry Forum underscores the importance of establishing analytical goals for precision and accuracy that are medically relevant. If these goals are specified, manufacturers can design instruments to meet them, laboratorians can manage their laboratories to meet them, regulators can use them as guides, and writers of practice guidelines can use them in formulating reasonable expectations from the laboratory testing process.

Testing Process

Quality laboratory medicine means doing the right tests, doing the tests right, and facilitating the right medical action. This has always been the goal of quality-oriented medical laboratorians and is illustrated by Figure 1, which shows the total testing process. Figures similar to Figure 1 have appeared in previous symposia (2). It is the goal of the present Forum to focus on the portion of Figure 1 from sample analysis to patient and provider agreement on an action. It is the laboratorian's hypothesis that the quality of the action is linked to both the quality of the analysis and the quality of reasoning. Further, it is the goal of the conference to devise a system whereby the quality requirements in the analysis could be specified such that they meet all the expectations of the reasoning and action-decision portion of the process. Elsewhere, other Forum participants (3, 4) have indicated a paucity of data to illustrate an alteration of action decisions attributable to either inaccuracy or imprecision of the analysis. Furthermore, most problems in laboratory testing are attributable to portions of the testing cycle other than the analysis. In fact, as much as 90% of the problems have been attributable to the pre- and postanalytical portions (4).

The quality requirements for precision and accuracy in analysis will differ for different clinical situations, and will also differ for diagnostic and monitoring decisions.

The Decision Milieu

Data become more useful when they are collated into carefully organized information, further collated, and then synthesized into knowledge. The application of this knowledge into decision-making and policy-making requires a further synthesis of the data into wisdom (5). The use of laboratory data for making wise decisions regarding individual patients as well as wise population-health policy requires synthesis of data into information and knowledge. This synthesis of data into wisdom occurs through the process of medical cognition. Simultaneously, the same data are acted upon by patients and family members as well as societal institutions. It is the coalescence of these three viewpoints (patients, caregivers, and society) that can be expected to yield wise medical decisions, as illustrated in Figure 2. This conceptual analysis is similar to that of Eddy's discussion of the "anatomy of a decision" (6). As Eddy points out, a part of these policy and health-care decisions are scientific and require quantitative, logical thought. A second part of the decision process involves individual preferences and yields judgments that are used to make individual patient-care decisions as well as set medical policy.

It is important for laboratorians to realize that laboratory data are acted upon by patients, family members, physicians, and societal bodies to generate decisions.
Therefore, any quality specification for the precision and accuracy of the laboratory data must include a consideration of the thoughts and preferences of those who are using these data.

Cognition and Common Cognitive Errors

In their classic paper, "Judgment under Uncertainty," Tversky and Kahneman (7) outline the most frequent errors in individual judgments regarding uncertain probabilities. These concepts are important, because the interpretation of laboratory data requires an appreciation for the prior probability of disease and the probability of normal or abnormal results in specific clinical situations. The most common biases are representativeness, availability, and anchoring and adjustments.

"Representativeness" refers to misrepresenting the probability that a given individual is a member of a given class of individuals based on the degree to which that individual's description is representative of the observer's preconceived notion of that group. Tversky and Kahneman indicate representativeness bias is illustrated in medical decision-making by an insensitivity to sample size and an insensitivity to the predictability of the evidence. Also, there is an unwarranted confidence in correlative and redundant data. The representative bias also causes observers to discount the regression towards the mean when individuals are retested.

The second cognitive bias relates to "availability." This involves the false notion that estimation of the prior probability of the presence of a disease in a given individual is related to how easy it is to remember a similar individual with that disease. The data they (7) present illustrate that frequencies of abstract events are often overestimated, whereas frequencies of concrete events are underestimated.

A third cognitive bias is attributed to "anchoring and adjustment." When asked to estimate a prior probability of disease in a given individual, most observers have a first impression. This would be the anchoring point. Usually, when individuals realize their anchoring point is incorrect, they make an insufficient adjustment of that anchoring point. This causes individuals both to be overconfident of their probability estimates and to overestimate the probability of consecutive events (the optimism of planning) and underestimate the probability of failure when the failure rate for each step in the process is relatively low. In addition to Tversky and Kahneman (7) the reader is referred to Dawson (3) for further elucidation of these concepts.

Tversky and Kahneman have also shown (8) that choices are influenced by the form in which the choice is presented. This is the cognitive error attributable to "framing." Most individuals, when given a choice regarding a potential gain, are risk avoiding. In contrast, most individuals, when given a choice regarding potential losses, are risk taking. This is illustrated in Tversky and Kahneman's paper by the situation in which a vignette is given to individuals stating that an unusual disease is likely to kill 600 people if nothing is done. In the first pair of foils, the individuals are given a choice between Program A, where there is a 100% probability 200 people will be saved, and Program B, where there is a one-third probability that all 600 can be saved. Respondents chose Program A over B 72% to 28%, even
though the probability of saving lives is equal for A and B. Conversely, if the respondents were told that if Program C is instituted there is a 100% probability 400 people will die, but for Program D there is a one-third chance of there being no deaths, they chose Program D over C, 78% to 22%. Once again, the probability of Programs A, B, C, and D are identical for saving lives. It is important to remember that the way a question is asked may alter its answer.

Medical Decision-Making

Kassirer (9) and Moskowitz et al. (10) have recently reviewed the mental processes involved in medical decision-making. They classify them as probabilistic, causal, or deterministic.

Probabilistic reasoning involves the use of rigorous data and Bayes' theorem. Laboratorians are familiar with Bayes' theorem as well as the concepts of sensitivity, specificity, positive and negative predictive values, and likelihood ratio (11). Kassirer indicates that probabilistic Bayesian reasoning is strong if used carefully and if adequate data are available to support the prior probability and the sensitivity and specificity estimates.

Causal reasoning involves the use of fundamental pathophysiological principles. That is, a given sign or symptom is known to be associated with a specific cause. This form of reasoning is also powerful, but occasionally there are difficulties with initiating (triggering) the concept of the association. This form of reasoning favors the experienced clinician.

Deterministic reasoning involves use of well-defined rules or algorithms. Deterministic reasoning can be very reproducible; however, it may perform poorly in clinical situations where a patient has multiple problems.

These three types of medical decision-making contrast the use of formulas and rigorous nonjudgmental mathematics with the use of clinical intuition. Dawson and Arkes (3) provide examples of the use of intuition vs mathematics—a well-studied contrast (12, 13). In most situations, nonjudgmental mathematics is shown to be superior to clinical intuition. However, the experienced clinician may have an advantage in some situations where recognition of a rare event would provide an advantage over the rigorous mathematic approach. Unfortunately, few personal examples of successful clinical intuition have been studied prospectively with objective measurement of outcomes for both intuitive and mathematical decision-making. References 12 and 13 further elucidate these cognitive biases.

Confidence in decision-making is important for clinicians. It is also important to ascertain whether those who appear to know more, know more about how much they know (14). Cognitive psychologists studying this question have discovered that knowing more increases performance, as would be expected—but knowing just a little more leads to overconfidence and therefore to poorer performance. This overconfidence decreases with an increasing knowledge base, and those with the greatest knowledge base tend to have a slight underconfidence. This is the concept of calibration of intuitive probability estimates (3, 14).

Patient Cognition

The needs of the patient must be paramount in any medical decision-making process. Likewise, the psychological profile of the patient is important. Patients differ in their desires to take risks or avoid risks and in their evaluation of potential gains and potential losses. These concepts have been studied (15). It is important to remember that patient cognition is subject to the same biases as all cognition: availability, representativeness, and anchoring and adjustment. Patient outcomes must be optimized, and the value of the outcome must be considered for the patient first, society and health-care providers second.

The use of the patient's preferences in making wise medical decisions as to treatment choices for benign prostatic hypertrophy has been studied by Wennberg et
al. (16, 17). Surgical vs medical therapy for this condition have significantly different outcomes and side effects. Individual patients will value these side effects much differently, and these preferences are important when therapeutic decisions are being made.

Society Cognition

Nowadays all medical decisions have an overlay of societal expectations. These include government, health-care institutions, employers, and payors. Eddy (18) has discussed a specific aspect of society's needs regarding medical policy-making. He indicates that one serious problem is a lack of connection between the value of a health-care service provided and the cost of providing that service. He states: "An essential condition for achieving an equilibrium between cost and value is that the two must be connected through decisions. When people decide what products and services (goods) they want, they must not only see the value they will receive, but they must also be responsible for the costs . . ." Wise decisions require a careful scientific judgment about health outcomes, i.e., their benefits, risks, and harms. Then the patients must weigh these benefits and harms against their individual value set. Clinicians will help and advise patients in making these decisions. Secondly, the costs of any health-care intervention must be estimated. It is important to note that, from a societal point of view, charges for health care are frequently a poor proxy for costs (19). It is unfortunate that most health-care decisions made by individual patients as well as individual physicians tend to be charge-benefit analyses rather than cost-benefit analyses. Nevertheless, these charge-benefit analyses are the micro-economic decisions to be made by patients. It would be beneficial if larger societal bodies would think in terms of cost-benefit rather than charge-benefit. Clinicians must help patients make the individual charge-benefit analysis and should become well enough informed to help societal bodies make the cost-benefit analyses. Finally, it is most important that, once the connection has been made between cost or charge and service, the decision be adhered to, no matter what future events occur. For laboratorians this means that we must design our analytical processes with the right goals for precision and accuracy, and that the final health-care and financial outcomes must be planned in advance.

Clinical Cognition

The medical decision-making process by individual clinicians has been discussed elsewhere in this Forum. It is important to remember that judgments by individual physicians are subject to the same systematic cognitive errors as judgments by any other individuals. However, it is the goal of medical education to minimize adverse effects from these cognitive errors (listed in Table 1).

Laboratory Cognition

Laboratorians have debated the necessary accuracy and precision goals to support clinical reasoning and medical decision-making for many years. Acland and Lipton (20) and Tonks (21) have related the precision goals to the breadth of the normal "reference" range for a given analyte. In this Forum and elsewhere, Fraser (22), Harris (23), and Ross (24) have attached goals to the within-subject biological variability. There is considerable agreement among investigators that a limiting factor in test interpretation is this source of variability. Therefore, it is attractive to relate both precision and accuracy goals to some fraction of that within-subject biological variability.

Skendzel et al. (25) have used clinical vignettes to assess individual physicians' goals regarding laboratory accuracy and precision, and the use of such vignettes (25, 26) has provided some data regarding individual clinician opinion. Does physician behavior in a theoretical clinical vignette match with physician behavior in real patient-care situations? This has not been studied with a view to precision and accuracy goals; however, for resource utilization there has been a clear discrepancy between test utilization in vignettes and test utilization in practice (27). Although the responses to questionnaires have been useful in formulating goals for precision and accuracy, monitoring of performance by means of chart review would seem to yield more rigorous information about physician behavior in real situations.

There is a perception, at least among industrial clinical chemists, that the German Medical Association has significantly addressed the medical relevance issue regarding accuracy and precision goals in clinical chemistry (28). A review of the German Medical Association's quality assurance and medical relevance guidelines (29) reveals that the allowable inaccuracy and imprecision for the German system virtually matches the CLIA-USA system for bilirubin, calcium, chloride, iron, aspartate and alanine aminotransferases, uric acid, creatinine, lactate dehydrogenase, phenobarbital, phenytoin, primidone, theophylline, and valproic acid. However, there are some differences, as shown in Table 2. The linkage in the German system between medical relevance and accuracy and precision goals needs to be further delineated.

---

Table 1. Systematic Errors in Medical Decision-Making*

<table>
<thead>
<tr>
<th>Error Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inaccurate probability estimation</td>
<td>Availability</td>
</tr>
<tr>
<td>Representativeness</td>
<td>Ego</td>
</tr>
<tr>
<td>Hindsight</td>
<td>Regret (avoidance of regret, risk aversion)</td>
</tr>
<tr>
<td>Impediments to information synthesis</td>
<td>Confirmatory bias</td>
</tr>
<tr>
<td></td>
<td>Discounting negative evidence</td>
</tr>
<tr>
<td></td>
<td>Framing</td>
</tr>
</tbody>
</table>

* Adapted from Table 1 in Dawson and Arkes (3).
Hyltoft Petersen and Herder (30) described their mathematical system for investigating the alteration of diagnostic classification by changes in test accuracy (bias) and test imprecision at the 1987 CAP Conference XIII (Evaluation of Proficiency Testing Results for Quantitative Methods in Relation to Clinical Usefulness), but unfortunately did not present at the 1992 Clinical Chemistry Forum. This model has been applied to screening for hemochromatosis (31). In this model, Wiggers et al. used the clinical data of Olsson et al. (32). Hemochromatosis is an autosomal recessive iron-overload disease that is variably expressed in homozygotes only. It is well known that the morbidity and mortality can be greatly reduced by detecting homozygous individuals before clinical disease is present (33). The mathematical modeling of Wiggers et al. (31) was interpreted to be not supportive of the idea of screening the general population for hemochromatosis by using data on percent transferrin saturation. However, our review of the data of Olsson et al. (32) for 10,512 samples from 4098 inpatients, 3340 outpatients, and 1311 blood donors revealed that 84 individuals had repeatedly measurable increases of transferrin saturation, suggesting hemochromatosis. Many of these increases were explained by careful history and physical examination, but 13 individuals were biopsied and all 13 were positive for the disease. A study (34) not cited by Wiggers et al. involved 11,165 blood donors screened by using transferrin saturation. Individuals with reproducible values >62% for transferrin saturation were biopsied. Sixty-nine percent of those biopsied were patients with homozygous hemochromatosis; the other 31% of those biopsied had heterozygous hemochromatosis. As an additional benefit, 23 additional homozygous siblings of these identified probands were identified. This latter study was interpreted to suggest clear benefit from screening. In addition, Guzman et al. (35) have modeled the financial and medical aspects of hemochromatosis and suggested that there are financial as well as health-care benefits from screening the asymptomatic population. Is this difference in judgment by the statistical modelers an example of risk aversion when seeking gains? Is it an example where statistical analysis must be carefully balanced with results of actual practice? These are significant health-care policy judgments made in uncertain conditions and subject to the cognitive biases as discussed above.

Wiggers et al. (31) further modeled the utility of transferrin saturation to identify individuals with hemochromatosis if imprecision or inaccuracy were present in the laboratory method. It is important to note that the model of Hyltoft Petersen and Herder (30) and Wiggers et al. (31) differs from that of Ross (24) and Fraser (22), in that it involves modeling a specific clinical situation rather than relating goals to biological variation. Some analytes probably will require modeling and evaluation of specific clinical situations for the determination of appropriate analytical goals.

Many laboratorians have received calls from concerned physicians observing a high frequency of abnormal results in patients they expected to be normal. This loss of specificity could be attributable to analytical bias or increased imprecision. This concept has been modeled by Klee in this Forum (1) and elsewhere (36), offering additional insight for choosing goals for accuracy and precision.

A Call to Action

Previous efforts by CAP and AACC to define analytical goals based on medical relevance have been primarily statistical and analytical (2, 11, 22-24). The 1992 Clinical Chemistry Forum has also provided additional discussion of statistical and analytical modeling. However, this Forum has added a specific clinical focus and a recognition of the predictable biases of medical cognition. Unfortunately, many years have passed since the previous conferences. Data have been collected, but little consensus has been developed regarding analytical goals. Now is the time for the CAP and AACC to act on this important subject.

Many tools are available. We understand the analytical precision, analytical bias, and within-subject biological variability for many analytes. The pretest probabilities are known for many clinical situations. Bayesian statistical analysis is well developed in the medical laboratory and medical decision-making arena. The utilities and judgments of individuals, clinicians, and society are becoming more widely known. Data are readily available for many common analytes, but still are needed for some less-common analytes. The task at hand is synthesis of these data into information and knowledge to generate wise decisions.

Collaborative effort is needed at the interface between clinical science and analytical science. Pathologists (CAP) and clinical chemists (AACC) are uniquely trained to work at this interface, but we must not work there alone. Expectations of our clinical colleagues,
patients’ preferences, and society’s needs must be considered. The analytical goals for accuracy and precision are quality specifications. Quality planning requires an understanding of the needs of patients, physicians, and society. We must enlist these groups as partners in this planning process. The outcome of this Forum is to be the beginning of a plan for defining the medical requirements for precision and accuracy on a test-by-test basis and, where needed, on a clinical situation basis—an important goal for our collective futures. With these goals defined, manufacturers can design instruments to meet them, laboratory managers can manage laboratories to meet them, regulators can use them as guides, and practice-guideline formulators can understand the laboratory’s ability to perform.

We must be ready to commit the expertise of both the CAP and AACC to this goal. Thoughtful quantitative-minded clinicians need to be enlisted as partners to formulate realistic and usable goals. These goals can be submitted to the National Committee for Clinical Laboratory Standards for the consensus-gathering process. We must begin by taking the first step and the time to take it is now.

It is important to end on a note of humility. Even though we are ready to embark on making a list of very important specifications, we should remember the words of Henry David Thoreau: “If you look over a list of medicinal recipes in vogue in the last century, how foolish and useless they all seem to be! And yet, we use equally absurd ones with faith today” (37). We must begin immediately to improve the present.

References
14. Lichtenstein S, Fischhoff B. Do those who know more also know more about how much they know? Organ Behav Hum Perform 1977;20:159–83.