Use of Analytical Goals by Health Care Manufacturers

Jan S. Krouwer, Ph.D.
Evaluations & Reliability
Ciba Corning
Medfield, Massachusetts

Abstract: Analytical goals are the *quantitative* requirements that the product must meet. Goals must be specific and have a clear success/failure criterion. Ideally, goals should include a protocol and data analysis method. There are three categories of analytical goals. First, clinical acceptability is the total analytical error and total analytical error sources. Second, there are a variety of regulatory goals depending on the approval required (510k, PMA, PLA). Finally, there is a list of marketing goals spawned by the competitive nature of business.

Setting goals involves: defining a metric, setting its target, and specifying a protocol and data analysis method. Whereas many of the metrics are defined, the target setting process is still difficult for manufacturers. Laboratorians may know what they need, but effective communication of these needs to manufacturers is lacking.

To set targets, manufacturers perform surveys (open ended questions, multiple choice questions, focus groups, and conjoint analysis). They also use performance data for similar, released assays (CAP data, internal data, published evaluation data).

Reviewing existing goals reveals inadequacies such as non existent goals, non quantitative goals, goals without a meaningful success/failure criterion, and unsupported goals. The goal setting process can be improved by deciding on and gaining experience with a metric, preparing cause and effect diagrams, and challenging existing goals.

Goals and claims are different. Manufacturers have internal goals. Upon product release, these goals are transformed into "claims", which may be different from the internal goals. Different manufacturers state claims differently, leading to confusion. Claims are: 1) The "typical data" claim - Half of the customers are expected to observe better performance, and the other half, worse performance. 2) The "warranty" claim - Here, all customers are guaranteed performance better than the limit. What is needed is a common vocabulary for the laboratorian and manufacturer.

Introduction

Manufacturers of diagnostic assays have a key milestone in the product development process: product release. The analytical performance goals they set are often quite different before and after product release. Before release, goals are often called *(internal) specifications* and, after release, *product claims*. Laboratorians never see the internal specifications. Sometimes there is confusion as to which goals are under consideration. This paper focuses on goals before product release. They can be divided into three conceptual categories: clinical acceptability, regulatory, and competitive goals.

Lab results have error. *Clinical acceptability* goals define how bad the error can be before it causes diagnostic problems. Laboratory assays are regulated. For

manufacturers, this means that assays must be FDA approved. *Regulatory* goals depend on the approval required (510k, PMA, PLA). For a laboratory, regulatory agencies require acceptable performance on proficiency surveys. Hence, assays must achieve a certain performance level with proficiency survey controls. Finally, companies must be *competitive* to remain in business. This spawns a list of marketing analytical performance goals.

Definition of an Analytical Goal

Analytical goals are a subset of the quantitative requirements that an assay must meet. The terms specifications, target values, and requirements are synonyms of analytical goals. Goals must be specific and have a clear success/failure criterion (e.g., there must be a metric). Ideally, goals should specify a protocol and data analysis method. This assures that not only the right type and amount of information will be collected but also describes how the data will be analyzed and reported. An analytical goal example is: the total precision CV should be less than 10% throughout the 50-500 mg/dL range as determined by the NCCLS protocol EP5.

A protocol and analysis method is recommended as part of a goal because the analytical performance of an assay differs from directly measurable assay properties such as the size and weight of an instrument. Analytical performance cannot be exactly determined - the true performance values (the "true state of nature") can only be estimated by experiments. Variation in the experimental results prevents their direct determination. The resulting data from these experiments has information in it that a properly designed analysis procedure will extract. Without a correct analysis and reporting procedure, interpretation of the data will be difficult if not impossible.

Without protocols and analysis methods it is unclear how to determine if a goal is being met. For example, for a glucose assay with a range of 5-1000 mg/dL: do we need to evaluate precision at 5 concentration levels, every 20 mg/dL, every mg/dL? Can we spike and dilute samples? If we dilute, what should be the diluent? To test interferences, what glucose level(s) should be used? If bias is evaluated, what is the criterion for meeting the goal, the point estimate? Its 95% confidence interval, its 99% confidence interval, every data point within the goal? NCCLS evaluation protocols help address some but not all of these issues. Experience shows that agreement for these issues helps to prevent questions after the data have been collected.

Constructing goals

Constructing goals involves:

- defining a metric (e.g., % CV precision)
- setting the metric's target (e.g., < 5% CV precision)
- defining a protocol to evaluate the metric (e.g., 2 observations per day for 20 days)
- defining an analysis and reporting procedure for the metric (e.g., ANOVA)

Consider that most assays are developed by manufacturers for sale to clinical laboratories. Laboratories run the assays and provide clinicians with results. Clinicians use the results to help answer the question, "Should I treat or not treat the patient?"

To a clinician, **total analytical error** is

the only parameter of importance. Total analytical error is defined as a percentage, (often 95% or 99%) of the distribution of differences in concentration between the test and reference method. Wrong results that cause misdiagnosis are just as harmful whether they are caused by random or systematic error.

Besides total analytical error, laboratories need to know **total analytical error sources** because these sources contribute to total analytical error and some error sources are specified by regulatory agencies. The manufacturer needs to know both total analytical error and total analytical error sources because he must satisfy all clinical and laboratory needs. Knowledge of error sources leads to improved assay performance which helps meet competitive goals.

Hence, there are two analytical goals for manufacturers:

1. total analytical error, used to validate the clinical use of an assay
2. total analytical error sources, used to improve assay quality

Setting Analytical Goal Targets

The reason that it is hard for manufacturers to set targets for goals is:

- manufacturers don't know how to ask for targets
- laboratorians don't know how to talk about targets

Surveys

Surveys would seem to be an easy way to set clinical acceptability targets. One simply asks clinicians. However, there are pitfalls. If one asks clinicians *open-ended* questions, such as "what are the clinically acceptable limits for a cholesterol assay?", one could get responses such as "no error" or "error that doesn't cause diagnostic problems." One remedy to this is to offer *multiple choice* questions. However, this is not foolproof either. Responses can be checked off without a guarantee that the question was correctly understood by the respondent. Moreover, respondents tend to want "the best." Thus, given a choice for total error for cholesterol to be 1%, 3%, or 5%, many respondents will simply check off 1%. Skenzel overcame many of these difficulties in a cleverly constructed survey.[1]

The problem is that in real life, one must make tradeoffs. One wants a car that is both luxurious and low priced. For laboratory assays, one wants low total error, low cost, high ease of use, high reliability, etc. For any of these situations, there will be acceptable compromises among the desired values of the goals. Conjoint analysis is a form of marketing research that provides a protocol and analysis method for estimating these tradeoffs.[2] Its idea is to present a clinician with a series of assays, each with different values for various attributes. The clinician ranks his preference for each of the assays. With several clinicians performing this ranking, the value of each attribute can be found by statistical analysis.

Theory

Studies have been made to set analytical performance goals by relating biological variation and analytical error to diagnostic decision making. Manufacturers keep track of these studies and try to ascertain to what extent the results are used by laboratorians.

Use of Current Performance Data

Given an assay that is in service, one can ask if the complaint rate is sufficiently low. A yes answer may signify that the assay's analytical performance is adequate. One can then measure this performance and use it as a

goal for a new version of the same or similar assay. Assay performance data sources can include: CAP or other proficiency survey data, published evaluations, or in-house studies. Of course, this method will not work for new analytes, for which there are no data. Moreover, the problem has not really been solved, it has been transferred. One must now decide what is a sufficiently low complaint rate.

Allocating Total Error into Goals for Total Error Sources

Setting goals for error sources that contribute to total analytical error creates a rather complicated problem. Goals for total analytical error sources have the constraint that the sum of combined values of the individual sources cannot exceed the total analytical error goal. Error modeling (propagation) can be used to achieve this.

Inadequate Goals

A **non existent goal**, while a rather obvious category, crops up surprisingly often. An example is lack of an outlier goal. Outliers are values that are so far away from the true values that they almost always cause problems. Yet, there is seldom a goal describing how far off a value must be to called an outlier and how many outliers are acceptable. Ideally, there should never be outliers, but an implied goal of zero outliers has its own problems. Realistically, there will always be a small but measurable frequency of outliers, even for the best assays. A goal of zero will be an **unrealistic goal** and not useful.

Aids to Improve Goal Setting

A goal requires a clear pass/fail criterion. This implies that a metric is in place which will be used to base the decision. **Deciding on a metric** is a step in agreeing on goals. A metric should be objective, easy to understand, and relevant to the goal. Not all metrics are appropriate. For example, one might use intuition[1] as a metric: "I've seen the data and feel the assay is OK." Metrics can be evaluated to assess their validity. This could be done with simulations (e.g., without actually running assays). For example, the correlation coefficient is sometimes used as a measure of agreement between two methods, in spite of reports that caution this metric's limitations. One can simulate what values of correlation coefficients would be observed with various expected datasets to see how well the correlation coefficient predicts agreement between methods. Alternatively, one could retrospectively analyze real data with the proposed metric, again testing its predictive value.

Preparing **cause and effect diagrams** (also called fishbone or Ishakawa diagrams) helps to highlight potential analytical problems and focus goal setting. The universe of potential goals must be narrowed down to those that have a reasonable likelihood of causing problems. Otherwise, the list of goals to be tested could be endless.

One can go further with a cause and effect diagram by **mathematically modeling** error sources such that sources combine to yield an estimate of total error. This method is also called propagation of errors. It allows goal limits to be assigned to total analytical error sources. One study showed that for a cholesterol assay, traditional analysis underestimated total analytical error

[1]This does not mean that one should not use intuition. This is an old statistical adage: "Beware of the following: *Statistics on - Brain off*"

compared with a method that estimates total analytical error and its sources.[3,4]

How goals are and should be used

Manufacturers, like mostly everyone, are faced with yes/no decisions. Should we release or not release the product (meaning has it met or not met its goals)? Consider two assays, however, where assay **A** is just inside and assay **B** is just outside of spec. From a manufacturer's standpoint, assay A has *full value* (i.e., identical to an assay that is perfect), whereas assay B has *zero value*. From a customer standpoint, the two assays have *similar* performance (and value). Yet, manufacturers must still treat these two assays that are similar to a customer, as either having full or zero value.

There is no easy way to deal with this problem. What happens in practice is that the yes/no region, while conceptually clear (i.e., accept if precision is less than or equal to 4.0% CV, reject if greater than 4.0% CV) becomes fuzzy: accept if precision less or equal to 4.0% CV; conduct further discussions if precision is between 4.0 and 5.0% CV; and reject if greater than 5.0% CV. Thus, targets set at the beginning of a project are revisited throughout the project development cycle and especially near product release. Since many specs are set close to the technical capability of a system due to competitive pressures, the situation occurs frequently.

Goals during the commercialization of the assay (the claims that are made to customers)

When the product is released, internal goals are transformed by the manufacturer into customer claims, which may or may not be the same as the internal goals. The claims represent a data source for customers, who can compare different manufacturers claims. Of course, as most consumers are aware, not all claims are always met! There is an additional problem. A claim can be stated in a way that is not clear, leading to confusion. Basically there are two types of claims:

1. The **"typical" data claim** - Half of the customers are expected to observe better performance, and the other half, worse performance. With statistical tests, one can determine whether performance is unreasonably far from the claim.

2. The **"warranty**" claim - Here, all customers are guaranteed performance better than the limit.

An NCCLS subcommittee is trying to address these problems by providing guidelines for standard language to be used for claims.

References

1. Skendzel LP Barnett RN and Plat R. Medically useful criteria for analytic performance of laboratory tests. *Amer J Clin Pathol.* 1985;83:200-5.

2. Green PE and Wind Y. New way to measure consumers' judgments. Harvard Business Review 1975;107-17.

3. Krouwer JS. Estimating total analytical error and its sources. Techniques to improve method evaluation. *Arch Pathol Lab Med.* 1992;116:726-31.

4. Krouwer JS and Monti KL. A simple, graphical method to evaluate laboratory assays. *Eur J Clin Chem Biochem.* 1995;33:525-7.