Assessing the Impact of Analytical Error on Perceived Disease Severity

Martin H. Kroll, MD; Carl C. Garber, PhD; Caixia Bi, PhD; Stephen C. Suffin, MD

• Context.—The perception of the severity of disease from laboratory results assumes that the results are free of analytical error; however, analytical error creates a spread of results into a band and thus a range of perceived disease severity.

Objective.—To assess the impact of analytical errors by calculating the change in perceived disease severity, represented by the hazard ratio, using non–high-density lipoprotein (nonHDL) cholesterol as an example.

Design.—We transformed nonHDL values into ranges using the assumed total allowable errors for total cholesterol (9%) and high-density lipoprotein cholesterol (13%). Using a previously determined relationship between the hazard ratio and nonHDL, we calculated a range of hazard ratios for specified nonHDL concentrations affected by analytical error.

Results.—Analytical error, within allowable limits, created a band of values of nonHDL, with a width spanning 30 to 70 mg/dL (0.78–1.81 mmol/L), depending on the cholesterol and high-density lipoprotein cholesterol concentrations. Hazard ratios ranged from 1.0 to 2.9, a 16% to 50% error. Increased bias widens this range and decreased bias narrows it.

Conclusions.—Error-transformed results produce a spread of values that straddle the various cutoffs for nonHDL. The range of the hazard ratio obscures the meaning of results, because the spread of ratios at different cutoffs overlap. The magnitude of the perceived hazard ratio error exceeds that for the allowable analytical error, and significantly impacts the perceived cardiovascular disease risk. Evaluating the error in the perceived severity (eg, hazard ratio) provides a new way to assess the impact of analytical error.


Quantitative results play an important role in diagnosing and assessing the severity of disease and in influencing treatment plans.1 Even though diagnostic errors have been linked to undesirable outcomes, the mechanisms relating clinical outcomes to analytical error are unclear.2 The perceived disease severity related by quantitative results is proportional to the disease spectrum, which is often quantified using hazard ratios, relative risk, odds ratios, and morbidity and mortality statistics. Biochemical test values may directly and continuously correspond to disease spectrum, and thus to disease severity, and clinicians may use this continuous scale to perceive the severity of disease and modulate therapy.3 To our knowledge, few studies or metrics have been published linking analytic errors with perceived severity for the individual patient.4 Diseases follow a natural history divided into 4 stages: initial, subclinical, clinical, and final.5 The pathologic changes may be structural or biochemical and follow the natural progression of disease over time.5 Biochemical testing can provide important information during the subclinical stage, prior to the appearance of signs or symptoms.5 Clinical studies linking disease severity, such as risk of morbidity or mortality, with laboratory values create a continuous scale, and the relationship between biochemical markers and disease progression can be represented using an empirical function.3,6 Test values provide a measure of perceived disease severity; however, bias shifts test values away from error-free results, and random error (imprecision) generates a distribution that can be quantified as a band of results. Together, bias and imprecision produce a spread of values around an error-free result. Although it is difficult to observe the impact of analytical error on the outcome for any individual patient, one can calculate a range of error-transformed values and use that range to calculate an error-transformed perceived severity range; the size of that range stipulates the impact of analytical error.4 The hazard ratio (a measure of disease severity) relates to the laboratory values in a continuous fashion and provides a reliable mechanism to calculate hazard ratios altered by error-transformed values.6 Our objective was to quantify the change in the perceived disease severity imparted by bias or imprecision. In this study we developed a metric that quantifies the impact of bias and imprecision on the perceived disease severity, using non–high-density lipoprotein cholesterol (nonHDL) as an example. The metric does not depend on any special characteristics of cholesterol, high-density lipoprotein cholesterol (HDL-C), or coronary heart disease and could be applied to any continuous test.
MATERIALS AND METHODS

We used results transformed by analytical error to calculate the error-transformed disease severity. The method applies the allowable bias (both positive and negative analytical error) to the test result, creating a range or band of bias-transformed results. Then, imprecision (both positive and negative analytical error) is applied to the extreme values of the bias band to create a band of error-transformed results. Alternatively, one could apply the total allowable error to the test result to create a band of error-transformed results. We use a measure of disease severity, such as a hazard ratio, relative risk, odds ratio, etc, to correlate the test results to the disease severity, and apply the band of error-transformed results to create the range of the error-transformed disease severity.

We present the above method applied to nonHDL. The concentrations of cholesterol and HDL-C are used to calculate nonHDL:

\[
\text{nonHDL} = \frac{\text{cholesterol}}{\text{HDL-C}}
\]

We combined the individual allowable errors for cholesterol and HDL-C to produce error-transformed values for nonHDL. According to the National Cholesterol Education Program, for cholesterol the allowed bias is ±3% and imprecision is ±6%.\(^7\) For HDL-C the allowed bias is ±5%, whereas imprecision is ±8% for concentrations 42 mg/dL (1.09 mmol/L) or greater and ±3.4 mg/dL (0.044 mmol/L) for concentrations less than 42 mg/dL (1.09 mmol/L).\(^7\)

For each concentration of cholesterol or HDL-C we created a band of bias-error-transformed values by adding or subtracting the allowable error related to bias. We created a band of error-transformed values by adding or subtracting the allowable imprecision to the upper and lower ends of the bias-transformed values and using the most extreme values.

\[
\text{error – transformed value} = (1 \pm \text{bias})(1 \pm \text{imprecision})\times(\text{error-free value})
\]

To calculate the limits of allowable error:

\[
\begin{align*}
\text{upper limit nonHDL} &= (1 + \text{bias})(1 + \text{imprecision})(\text{cholesterol}) - (1 - \text{bias})(1 - \text{imprecision})(\text{HDL-C}) \\
\text{lower limit nonHDL} &= (1 - \text{bias})(1 - \text{imprecision})(\text{cholesterol}) - (1 + \text{bias})(1 + \text{imprecision})(\text{HDL-C})
\end{align*}
\]

Based on the analysis of 301,430 participants without any known history of coronary heart disease after adjustment for age, sex, systolic blood pressure, smoking status, history of diabetes, and body mass index;\(^8\) the hazard ratio for coronary heart disease related to nonHDL can be represented by a simple linear relationship using linear regression analysis.\(^6\) We calculated the range of the hazard ratio using the error-transformed bands of nonHDL.

The percentage difference is calculated as

\[
100 \times \left( \frac{\text{hazard ratio}(\text{error-transformed nonHDL})}{\text{hazard ratio}(\text{error-free nonHDL})} - 1 \right)
\]

RESULTS

Error-transformed values at the extreme limits for the allowable errors for bias and imprecision for nonHDL demonstrate a large constant bias and a small proportional bias in relationship with the error-free nonHDL values (Figure 1). From lowest to highest values the error range spans approximately 40 mg/dL (1.04 mmol/L) for an HDL-C concentration of 30 mg/dL (0.78 mmol/L) and approximately 60 mg/dL (1.55 mmol/L) for an HDL-C concentration of 70 mg/dL (1.81 mmol/L; Figure 1). The ranges show a wider range with either increasing values of error-free nonHDL or increasing values of HDL-C.
The regression analysis described the relationship between the hazard ratio and non-HDL can be expressed as

\[
\text{hazard ratio} = 0.021 \times \text{nonHDL} - 1.77
\]

(with standard error of the slope = 0.0003, standard error for the intercept = 0.06, \( r^2 = 0.99 \), and standard error of the regression = 0.03). The low standard error of the regression implies that any difference in hazard ratio less than 0.06 is statistically significant at the 95% confidence level.

The difference between the hazard ratios for a 3% positive or negative bias is about 0.6 for a constant imprecision (Figure 2). From a negative 6% to a positive 6% imprecision, the hazard ratio varies from a low of 0.5 to 1.45 for 130 mg/dL (3.37 mmol/L) nonHDL (Figure 2, green shading), 1.05 to 2.15 for 160 mg/dL (4.14 mmol/L) nonHDL (Figure 2, yellow shading) and 1.65 to 2.8 for 190 mg/dL (4.92 mmol/L) nonHDL (Figure 2, red shading). For nonHDL of 175 mg/dL (4.53 mmol/L), a concentration of nonHDL that falls between the middle and upper cutoffs, the hazard ratio varies from 1.35 to 2.5.

Doubling the bias of cholesterol or HDL-C increases the absolute value of the difference in the hazard ratio, but that for cholesterol has more of an effect (Figure 3). Doubling the imprecision for HDL-C increases the absolute value of the

![Figure 2](image2.png)

**Figure 2.** Hazard ratio error bands caused by error-transformed non–high-density lipoprotein (nonHDL) values. Positive and negative analytical errors impart error bands around bias-free cholesterol for 130 mg/dL (3.37 mmol/L) nonHDL (bottom, shaded green), 160 mg/dL (4.14 mmol/L) nonHDL (middle, shaded yellow), and 190 mg/dL (4.92 mmol/L) nonHDL (top, shaded red). The bands average 0.5 hazard ratio units at each percentage of imprecision; dashed line indicates negative cholesterol bias and solid line positive cholesterol bias. Within the allowed cholesterol imprecision, the green bands range from 0.5 to 1.45, the yellow bands from 1.1 to 2.15, and the red bands from 1.65 to 2.85 hazard ratio units.

![Figure 3](image3.png)

**Figure 3.** Impact of various values for bias and imprecision on the difference (percentage) of hazard ratios. The absolute values of the difference in hazard ratio between bias-free cholesterol (Chol) and bias-limited Chol increases with increasing error and decreases with decreasing error. The dotted line represents the allowable error: Chol Bias = 3%, high-density lipoprotein (HDL) Bias = 5%, and HDL imprecision = 8% (dotted line). The variations in error, in order from the top, are Chol Bias = 6%; HDL coefficient of variation (CV) (HDL imprecision) = 16%; HDL Bias = 10%; HDL Bias = 1.5%; Chol Bias = 1%; and Low Error (cholesterol bias = 1%, HDL Bias = 2.5%, and HDL imprecision = 4%). The percentage difference is calculated by dividing the difference of the hazard ratio by 1.60.
difference in the hazard ratio, more than the effect experienced by doubling the bias of HDL-C but less than that of doubling the bias of cholesterol (Figure 3). Conversely, decreasing the bias for HDL-C decreases the absolute value of the hazard ratio, but less so than for decreased bias for cholesterol (Figure 3). Decreasing the biases for cholesterol and HDL-C as well as the imprecision for HDL-C decreases the absolute value of the difference in the hazard ratio, as far down as 0.1 in the absence of cholesterol imprecision.

**COMMENT**

Quantitative test values continuously relate to the spectrum of disease, that is, disease severity. Test values offer a continuous perception of the severity of disease, influencing communication of disease prognosis and treatment. Complexity of patient care complicates linking impact of analytical error directly to patient outcomes; however, a metric that calculates the change in perceived disease severity related to analytical error quantifies this impact.

The metric combines the changes in analyte values imparted by analytical error with a relationship that maps a measure of disease severity, such as the hazard ratio, relative risk, odds ratio, or other prognostic outcome, to the analyte values. Analytical error transforms error-free values to error-transformed values. In a perfect world, our methods would be error free; however, in the real world some bias and imprecision must always be allowed (allowable error). Transforming the error-free values creates a spread or band of values. The allowable error dictates the acceptable extremes of this band.

Evaluation of the extremes of the band quantifies the impact of analytical error. The extremes of acceptability are governed by the partitioning of the total allowable error into bias and imprecision. The size of the range provides guidance on how the results may be perceived and potentially misinterpreted. In addition to misclassification, a value associated with a hazard ratio of 2.0 demonstrates a much worse degree of disease severity than one associated with 1.3.

We demonstrated the use of the metric with nonHDL and the hazard ratio for coronary heart disease. NonHDL is calculated by subtracting the HDL-C concentration from that of cholesterol. The nonHDL value provides an estimate of the total number of circulating atherogenic particles and has been recommended in primary risk detection in predicting cardiovascular disease, especially in patients with diabetes.

The target goal for patients with diabetes is nonHDL less than 130 mg/dL (3.37 mmol/L). The general target cutoffs for nonHDL are less than 130 mg/dL (3.37 mmol/L) for high-risk (20% risk) patients and optional for very-high-risk patients, less than 160 mg/dL (4.14 mmol/L) for moderate-risk patients (2 risk factors with <20% chance for an cardiovascular event), and less than 190 mg/dL (4.92 mmol/L) for low-risk (0–1 risk factors) patients, these regions corresponding to the desirable, moderate-risk, and high-risk zones, respectively. Hazard ratios for nonHDL are stronger than those for low-density lipoprotein cholesterol or triglycerides.

Bias error transforms a single result into a band of results, resulting from multiplying the single result by the positive and negative percentages of bias error. The allowable bias error determines an allowable band, which imprecision widens. Because nonHDL is calculated by subtracting the HDL-C from the total cholesterol, the analytical error for both analytes must be considered. The positive (upper) limit for nonHDL is calculated by applying the positive analytical error to cholesterol and the negative analytical error to HDL-C and subtracting the negative error-transformed HDL-C from the positive error-transformed cholesterol (upper limit of cholesterol minus lower limit of HDL-C). The negative (lower) limit of nonHDL is calculated by the converse approach.

At a nonHDL concentration of 160 mg/dL (4.14 mmol/L), the error-transformed range extends from 130 mg/dL (3.37 mmol/L) to 190 mg/dL (4.92 mmol/L; Figure 1). The error-transformed range extends from the low-risk to the high-risk group. An original nonHDL of 140 mg/dL (3.63 mmol/L) would be considered to be at a sufficient target for a patient in the moderate-risk group, but could appear to have missed the target because the error-transformed value is greater than 160 mg/dL (4.14 mmol/L). Likewise, if the patient started in the high-risk group, he or she might appear to have reached the low-risk target, because the error-transformed value is less than 130 mg/dL (3.37 mmol/L).

Likewise, if a patient’s target was below 190 mg/dL (4.92 mmol/L), the patient’s error-transformed values could fall below the target, even though nonHDL was greater than 190 mg/dL (4.92 mmol/L; Figure 1). Midway in the low-risk region, at a value of 175 mg/dL (4.53 mg/dL), error-transformed values could appear as low as 140 mg/dL (3.63 mmol/L) and as high as 205 mg/dL (5.31 mmol/L; Figure 1). The patient’s severity of disease could appear as favorable for the lower values and unfavorable for the higher values.

The hazard ratio, a measure of the patient’s severity of disease, illuminates this problem. At a nonHDL of 130 mg/dL (3.37 mmol/L), the patient’s hazard ratio should be about 1.0. A positive error could transform the apparent hazard ratio to 1.45, which suggests that the patient is at a higher risk and indicates a greater disease severity (Figure 2, green shading). A nonHDL concentration of 160 mg/dL (4.14 mmol/L) has a hazard ratio of 1.60, but with negative precision error would appear at 1.05, that is, low risk, and with positive imprecision error, at 2.15, that is, high risk (Figure 2, yellow shading). The patient’s values may imply low to high risk, all based on the allowable error for the methods. A nonHDL of 175 mg/dL (4.53 mmol/L) falls in an intermediate position between the higher-severity (190 mg/dL [4.92 mmol/L]) and moderate-severity cutoffs (160 mg/dL [4.14 mmol/L]). Here a patient might appear to have a hazard ratio as high as 2.5 and as low as 1.4, which represents a large error in the apparent severity of disease. At a cutoff of 190 mg/dL (4.92 mmol/L) and hazard ratio of 2.25, the hazard ratio ranges from a high value of 2.8 to a lower value of 1.65, covering an area of very high risk to one of moderate risk (Figure 2, red shading). The range of possible hazard ratios indicates how the error-transformed values may result in misinterpretation of the severity of the patient’s disease status.

So far, we have assumed that the analytical errors are confined to the allowable error. In the event that the assays are not maintained within the allowable error, which might occur if the statistical process control fails to detect increased bias or imprecision, then the perceived hazard ratio beyond the ±6% imprecision and the examples above is exacerbated (Figure 2).

For the desirable zone, the hazard ratio ranges from 0.4 to 1.3; for the moderate zone, from 1.0 to 2.2; and for the high-
risk zone, from 1.5 to 2.9. The wide ranges diminish the meaning of the test results. The difference in the hazard ratio increases with increasing imprecision and can be visualized using the absolute value of the difference (Figure 3). The curves in Figure 3 to the right of 0% cholesterol imprecision refer to the increased upper limit, and those to the left to the decreased lower limit. The difference in hazard ratios between cutoffs is about 0.65. Because the errors represent a summation of both positive and negative errors, ideally the error should be kept below 0.325. Figure 3 shows the difference in the hazard ratio (and the percentage error in the hazard ratio) for cases with increased and decreased analytical error. The error in the hazard ratio is about 16% for an allowable bias error but without imprecision in the cholesterol measurement (Figure 3). Increased cholesterol bias has a greater percentage impact than the imprecision or bias in HDL-C. As the imprecision in cholesterol measurement increases, the percentage increase in the hazard ratio increases markedly, much greater than the analytical error (at 6% imprecision, the percentage error in the perceived hazard ratio is 32% whereas the analytical error is 13% + 9% = 22%).

Non–high-density lipoprotein cholesterol determined in different laboratories or even on different analyzers within the same laboratory could result in failure to discriminate between baseline (1.0) and twice baseline (2.0) risk. If one assumes the probability that the cholesterol imprecision exceeds the limit of 6% only 5% of the time, then the expected number of defects would be 25 000 per million.

Reducing the bias or random error reduces the impact of the analytical error. Assessing the impact on the perception of severity by examining the hazard ratio at various values for bias and imprecision provides a new way to determine the degree of impact for these types of errors. In addition, one could calculate an analytical limit from the perceived disease severity. The problem is not the test, nonHDL here, but the allowed bias and imprecision.

Distortion of the disease severity due to error hampers effective treatment, biasing therapy intensity and increasing health care costs. Classification of disease severity guides clinical decision making for the individual patient, driving the strength and timing of therapeutic interventions. Inaccuracy in the severity message may expose the patient to poor decision making, affecting the strength, dose, and timing of medications and recommended lifestyle changes. Interpretation of disease severity for the individual patient is directly affected by test accuracy and distorted by analytical bias and imprecision. One may measure the clinical effect of analytical errors by observing the change it induces in the perceived severity of the patient’s disease.

Limitations of this method include the examination of only one analyte, as an example; however, the metric could be applied to many quantitative tests. In addition, analytes with large biological variations may complicate this approach and the proposed investigations should be limited, at first, to those analytes with minimal biological variations.

Current metrics for evaluating the impact of analytical errors are inadequate because the ultimate outcome is not solely determined by the laboratory testing; there is usually a prolonged interval between test results and outcomes, and the process of diagnosis is complex.

One cannot establish causality by simply looking at inputs and outputs. Simple cause-and-effect relationships single out one entity as cause and another as outcome, assuming all factors combine through simple addition. The relationships among physiology, biochemistry, and pathology are complex, so application of the simple cause-and-effect assumption may lead to erroneous conclusions. In evaluating the impact of analytical errors on clinical outcomes, it is preferred to link the altered test range to distortions in the meaning of the message.
Studies in quality measures have directed most of their attention to large-scale testing errors (those that impact multiple patients).16,17 Systematic errors regress toward the mean and thereby provide only an average picture of the impact of error; however, the bias and imprecision of a particular method on an analyzer may vary from run to run (run-limited errors) because of variations in calibration, reagents, maintenance, and other aspects of the instrument environment. Statistical quality control provides feedback on bias and random error, but the feedback is mainly systematic if reviewed only on a monthly basis. Even though the bias and imprecision, on average, may be small, run-limited errors may be larger than systemic values and their impact should be considered as well, because they affect individual patients. Real-time statistical quality control provides a mechanism to evaluate and provide feedback on run-limited errors.

Our results suggest that analytical error should be kept below the current total allowable amount. Our analysis of nonHDL agrees with the recommendations by the Joint Committee for Guides in Metrology.18

The effect of imprecision may be improved by obtaining additional samples.19 Additional samples may be incorporated into our metric by dividing the imprecision by the square root of the number of samples (the bias is unaffected). Figures 4 and 5 show the effect of using 2 samples, where the imprecision is divided by $\sqrt{2}$. If we were dealing with only one test, one could simply examine the hazard ratio at the reduced imprecision along the x-axis of Figure 2, but the reduced imprecision of HDL-C needs to be incorporated into the lines drawn to indicate the hazard ratios now shown in Figure 5. Comparing Figure 4 with Figure 1, it is apparent that adding an additional determination decreases the spread of values. When examining Figure 5, the vertical line at 4.2% imprecision should be used for a method with 6% imprecision, and the values for the hazard ratios have a smaller range than in Figure 2. For a nonHDL concentration of 130 mg/dL (3.37 mmol/L) the range of hazard ratios from the lowest to highest allowable error is 0.77 (compared with 0.91 for one determination), for a nonHDL concentration of 160 mg/dL (4.14 mmol/L) the range is 0.86 (compared with 1.02 for one determination), and for a nonHDL concentration of 190 mg/dL (4.92 mmol/L) the range is 0.95 (compared with 1.14 for one determination). Additional determinations do not affect the bias error but do decrease the imprecision error; thus, the effect is improved. Additional determinations should be considered, as a way to decrease the effect of analytical error on the perception of disease severity, along with developing pathways to reducing bias and imprecision. The metric relating the error to the hazard ratio deftly performs the corrections imparted by additional determinations.

Analytic errors transform error-free values into error-dependent ranges. By linking disease morbidity and mortality with quantitative results, one can calculate a perceived disease severity. The error-dependent range can be incorporated into the perceived disease severity, creating a range. The perceived disease severity range demonstrates the distortion level under the influence of analytical error. Because patient outcomes are multifactorial, laboratory results cannot be directly linked to outcome in simple cause and effect. Excessive perceived disease severity may lead to unnecessary treatment and patient anxiety, whereas diminished perceived severity may lead to inadequate treatment and impaired patient motivation for lifestyle changes. We should seek to minimize distorted perceived disease severity. The impact should be considered when quantifying the validity of statistical quality control systems, both systemic and real time. Ultimately, the quality of the laboratory information provided has a great impact on medical decision making. Better quality leads to better outcomes, and better outcomes build value in laboratory information. The metrics developed here offer a sound way to calculate defined limits and circumvent the need to collect large quantities of empirical data. The results shown here for nonHDL could be extended to other analytes as well.

References


