Clinica Chimica Acta 440 (2015) 23–30

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/clinchim

A mathematical procedure to estimate the impact of a change in method on discordance or misclassification at a decision limit in laboratory method comparison studies

Robert L. Schmidt [a,*], Michal A. Kordy [b], Jonathan R. Genzen [a], Joely A. Straseski [a], Dina N. Greene [c], Christopher M. Lehman [a]

[a] Department of Pathology and ARUP Laboratories, University of Utah School of Medicine, Salt Lake City, UT, United States
[b] Department of Mathematics, University of Utah, Salt Lake City, UT, United States
[c] Kaiser Permanente, TPMG Northern California Regional Laboratory, Berkeley, CA, United States

ARTICLE INFO

Article history:
Received 28 August 2014
Received in revised form 28 October 2014
Accepted 30 October 2014
Available online 7 November 2014

Keywords:
Discordance
Decision limit
Analytical specificity
Precision
Cost-effectiveness

ABSTRACT

Background: Laboratories often adopt new methods. It would be useful to have a statistical procedure to estimate the incremental impact of a change in assay.
Methods: Mathematical modeling, statistical analysis, and case example.
Results: We derived equations to estimate the proportion of discordant results that can be attributed to the new laboratory method. The calculations were demonstrated by comparing eGFR values based on creatinine values determined using the enzymatic method (existing method) and Jaffe method (new method). The discordance rate at the 60 ml/min eGFR decision limit was 3.15%. In this example, we estimated that 60% of the discordant results could be attributed to the Jaffe method.
Conclusion: The sources of discordance in a laboratory method comparison study can be divided into three categories: The baseline discordance due to imprecision in the established method, the incremental discordance due to imprecision in the new method, and lack of analytical specificity. Discordance due to imprecision can be attributed to each individual method. Discordance due to bias can be attributed to individual methods if information is available to estimate the rate of biased observations in either method. Such information can be used to estimate the incremental cost effectiveness associated with the adoption of a new method.

1. Introduction

Laboratory method comparison studies evaluate the agreement between 2 different laboratory assays. In many cases, laboratories compare agreement between an established method and a new method to evaluate the new assay. In such studies, the objective is to determine the *incremental* change that would result from adoption of a new assay. What is the incremental impact on accuracy? What is the incremental impact on cost? A procedure to estimate the incremental impact of a change in assay would be useful.

Clinical decisions are often made relative to specific decision limits. Method comparison studies often give rise to observations in which measurements from the 2 methods fall on either side of a decision limit. Such pairs of observations are called discrepant or discordant results. Such results can arise from imprecision and/or bias in either or both methods. Discordant results near decision limits are particularly significant when they affect the classification of patients, potentially affecting outcomes. For that reason, it is often helpful to assess the diagnostic agreement between 2 assays with respect to specific decision limits. Such information could be used to predict the impact of a change in assay. Unfortunately, agreement studies only determine the overall discordance rate and do not attempt to assign discordance to either assay.

Discordance can arise from both bias and imprecision. Bias is a systematic measurement error; imprecision is due to random error. Systematic measurement errors can occur in a single sample (LAS) or across all samples (method bias). We will focus on discordance due to imprecision and lack of analytical specificity (LAS) which might be caused by interference in a single sample. It would be useful to have a mathematical procedure that could estimate the relative contribution of LAS and imprecision to discordance and to estimate the impact of a change in method on misclassification. To our knowledge, no such method exists. Haeckel et al. developed a method to predict the overall discordance rate between 2 assays but the Haeckel method does not determine the source of discordance [1]. Thus, a method to estimate the sources of discordance in method comparison studies would be helpful.

* Corresponding author at: Department of Pathology, 15 N Medical Drive East, University of Utah, Salt Lake City, UT, United States. Tel.: +1 801 585 2010; fax: +1 801 585 2805.
E-mail address: Robert.Schmidt@hsc.utah.edu (R.L. Schmidt).

http://dx.doi.org/10.1016/j.cca.2014.10.043
0009-8981/

2. Methods

Our objective was to demonstrate the practical application of a mathematical procedure using data for a typical analyte. To that end, we compared eGFRs obtained from serum creatinine (SCr) measurements obtained by the Jaffe and enzymatic methods.

2.1. Patient population

We collected routine workflow samples obtained at outpatient clinics and submitted for serum creatinine analysis. Samples were stored at 5 °C for 24 h prior to analysis. Five hundred-forty unique outpatient samples were randomly selected from the sample population submitted to the University of Utah hospital laboratory over a 45-day period (7/31/2013 to 9/13/2013). Daily sample sizes were approximately 10% of the eligible outpatient samples. We obtained 247 samples from outpatient clinics located at the University of Utah Hospital and 293 samples from outpatient clinics located outside the hospital. The study was approved by the University of Utah IRB.

2.2. Serum creatinine measurement

Sample analysis was performed on the Abbott Architect c8000 analyzer. Each sample was tested by both Jaffe (kinetic alkaline picrate, Abbott Laboratories) and enzymatic (creatininase, Abbott) methods [2,3]. After centrifugation, samples were loaded onto Architect sample trays and batch ordered to perform both Jaffe and enzymatic testing. eGFR measurement: eGFR was calculated using the Chronic Kidney Disease Epidemiology (CKD-EPI) [4] Eq. (4)

2.3. Serum creatinine and eGFR precision profiles

eGFR is a function of serum creatinine (SCr), age, sex and race. The precision of the eGFR was based upon the precision profile for serum creatinine and calculated using mathematical formulas derived from the CKD-EPI equation (see Appendix 1).

Patient specimens were pooled to samples with concentrations of 0.28, 0.79, 1.21, 2.73 and 5.08 mg/dl. The 5 samples were run in duplicate daily for 20 days. The total SD (within-run and between-run) was determined from the 40 observations at each concentration.

Precision studies provide estimates of precision at discrete concentrations; however, our calculations require precision estimates for all intermediate concentrations (i.e., a precision profile). We used linear interpolation to estimate the precision at intermediate SCr concentrations. The relationship between the CV and SCr was determined by regression analysis. The linear relationship was then used to estimate the SD at each SCr concentration: $\mathrm{sd(SCr)} = \mathrm{CV} * \mathrm{SCr}$.

2.4. Distribution of eGFR (historical values)

The distribution of eGFR was based on 45,911 serum creatinine values reported at the University of Utah clinical laboratory during 2013. Serum creatinine was determined by the enzymatic method. Age, gender and race were recorded and used to calculate the eGFR using the CKD-EPI equation.

2.5. Statistical analysis

Statistical calculations were performed using Stata 13 (Stata Corp, College Station, TX). Confidence intervals for the estimated expected conditional discordance rate were determined by bootstrapping. A thousand bootstrap samples of 540 were drawn from a historical population of 45,911 serum creatinine values.

3. Results

We present the results in 2 sections: theoretical framework and clinical application. In the theoretical section, we derive general formulas to calculate the incremental discordance rate (IDR). We then illustrate the practical application of these formulas with example calculations based on clinical data.

3.1. Theoretical framework

We estimated the discordance that can be attributed to the established method (Method *E*) and the new method (Method *N*). We also wish to estimate the discordance that arises from LAS and imprecision.

The incremental discordance rate can be determined by comparing the discordance rate obtained in a method comparison study (Method *N* vs Method *E*) against the discordance rate obtained when method *E* is compared against itself:

$$\overline{IDR}_{NE} = DR_{NE}^{obs} - \overline{DR}_{EE}^{prec} \tag{1}$$

where the subscripts *N* and *E* designate the new and established methods, respectively. $\overline{IDR}_{NE}$ is the expected incremental discordance rate attributable to the new method. DR_{NE}^{obs} is the observed discordant rate — the number of discordant observations divided by the total number of observations. $\overline{DR}_{EE}^{prec}$ is the discordance rate that would be expected if repeated measurements were conducted using the established method, Method *E*. In effect, $\overline{DR}_{EE}^{prec}$ is the discordance rate that one would expect to observe if Method *E* was compared against itself. For example, one could take a large number of patient samples, obtain duplicate measurements using Method *E*, and determine the rate of discordant observations at a chosen decision limit. In such a study, discordances would only arise from imprecision in Method *E*. We will refer to this quantity as the incremental discordance rate (IDR). Given DR_{NE}^{obs}, $\overline{IDR}_{NE}$ can be calculated if we can calculate $\overline{DR}_{EE}^{prec}$.

In principle, $\overline{DR}_{EE}^{prec}$ could be determined by conducting repeated measurements on patient samples with Method *E*. The accuracy of the estimate, $\overline{DR}_{EE}^{prec}$, depends on the sample size. Method comparison studies generally have relatively small samples sizes. A large study might have up to 200 comparisons which would not provide an accurate estimate of $\overline{DR}_{EE}^{prec}$. On the other hand, laboratories would generally have large volumes of historical observations (X_E) for an established (incumbent) method. For that reason, we describe a method to estimate $\overline{DR}_{EE}^{prec}$ using historical data from the established method. We now describe the calculation of $\overline{DR}_{EE}^{prec}$.

Consider 2 repeated measurements of an analyte by Method *E* close to a decision limit, *L*. Let $X_{E,i}$ be the *i*th observation by Method *E*. Let $D_{E,E}$ be the event that the measurements disagree with respect to the decision limit. Given 2 observations on the same sample, there are 2 ways that the 2 observations can be discordant relative to the decision limit:

$$D_{E,E} = \begin{cases} 1 \;\; if \left[\left(X_{E,1}<L\right) \text{ AND } \left(X_{E,2}>L\right)\right] \text{ OR } \left[\left(X_{E,1}>L\right) \text{ AND } \left(X_{E,2}<L\right)\right] \\ 0 \;\; otherwise \end{cases}. \tag{2}$$

Given a true value, *x*, the probability of discordance between 2 observations with respect to the decision limit is (see Fig. 1):

$$P\left(D_{E,E}|x\right) = P\left\{\left[\left(X_{E,1}<L\right) \text{ AND } \left(X_{E,2}>L\right)\right] \text{OR} \left[\left(X_{E,1}>L\right) \text{ AND } \left(X_{E,2}<L\right)\right] \Big| x\right\}. \tag{3}$$