S-99.

META-ANALYSIS OF DIAGNOSTIC TESTS: TWO-STAGE HIERARCHIC MODEL FOR COMBINING LIKELIHOOD RATIOS FOR POSITIVE AND NEGATIVE TEST RESULTS

AUTHORS: **S. Mantha**[1], E. Mascha[2], J. F. Foss[2], J. E. Ellis[3], M. F. Roizen[2];
AFFILIATION: [1]Nizam's Institute of Medical Sciences, Hyderabad, India, [2]Cleveland Clinic Foundation, Cleveland, OH, [3]University of Chicago, Chicago, IL.

Introduction: Traditionally, diagnostic tests are employed to "rule-out" or "rule-in" the underlying disease states. Alternatively, in the context of perioperative care, diagnostic tests are also used to predict adverse postoperative outcomes resulting from the underlying disease states. For diagnostic tests summarized in 2 x 2 tables, sophisticated methods of meta-analysis that use relative diagnostic odds ratios, summary receiver operating characteristic curve and likehood ratio scatter plots have been employed (1-3) for better evaluation of the discriminative value of diagnostic tests. A combined point estimate and 95% confidence intervals (CIs) for the likehood ratio (LR) for a positive test (LR-pos) and LR for a negative test (LR-neg) would also be useful.

Methods: We propose a two-stage hierarchic model (4) to combine information about LR separately for positive and negative test. An empirical Bayes procedure with a normal-normal heirarchic model was used get a meta-analytic confidence interval for overall median for LR-pos and LR-neg. For a single study, 95%CIs for LR-pos and LR-neg were computed using standard formulas (5). The variance between the studies was estimated using method of moments approach described by DerSimonian and Laird (6). The methodology was applied to the data summarized in a recent meta-analysis that evaluated accuracy of six diagnostic tests for predicting perioperative cardiac risk in patients undergoing major vascular surgery (2)

Results: The combined summary measure point estimate and 95% CIs for each of LR-pos and LR-neg are given in table 1.

Table 1

Test	LR-pos (95% CI	LR-neg (95% CI)
Ambulatory ECG	1.81 (1.30 to 2.51)	0.99 (0.86 to 1.13)
Exercise ECG	2.72 (1.85 to 4.00)	0.51 (0.30 to 0.86)
Radionuclide ventriculography	6.52 (2.43 to 17.51)	0.72 (0.54 to 0.97)
Myocardial perfusion scintiggraphy	1.65 (1.43 to 1.89)	0.49 (0.34 to 0.71)
Dipyridamole stress echocardiography	4.96 (2.31 to 10.62)	0.40 (0.17 to 0.90)
Dobutamine stress echocardiography	3.03 (2.06 to 4.46)	0.35 (0.23 to 0.51)

Discussion: For periperative risk stratification, LR-pos and LR-neg should be greater than 10 or less than 0.2 respectively, because these values indicate a substantial change in risk from the pretest level (7). Using such interpretation, even the dobutamine stress echocardiography, which was concluded as the best among the 6 tests for prediction of adverse cardiac outcome after vascular surgery (2), cannot be regarded as the ideal test.

References:
1. Stengel D et al. J Med Screening. 2003;10 :47-51
2. Kertai MD et al. Heart 2003:89; 1327-34
3. Beattie WS et al. Anesth Analg 2006;102:8-16
4. Mantha S et al. Anesth Analg 1994;79 :422-33
5. Simel DL et al. J Clin Epidemiol 1991;44: 763-70
6. DerSimonian R and Laird N. Control Clin Trials 1986;7: 177-88.
7. Grayburn PA and Hills D. Ann Intern Med 2003;138: 506-511.

S-100.

THE CLINICAL EFFICIENCY OF NEW GENERATION CO2 ABSORBENTS: AMSORB PLUS AND SODASORB LF

AUTHORS: F. V. Cobos II, **R. Shaffer**, J. Tinker;
AFFILIATION: University of Nebraska Medical Center, Omaha, NE.

Introduction: Given the cost advantages of low flow anesthesia, and the potential risks of using older absorbents which contain strong bases, new generation CO_2 absorbents have been developed that do not significantly react with inhalational anesthetics. These new absorbents are more expensive and competing products need to be evaluated under clinical conditions to measure efficiency and confirm in vitro performance[1]. We designed a retrospective, single-blinded study to compare the duration of clinical usefulness of the CO_2 absorbents Sodasorb LF and Amsorb Plus.

Methods: Without the knowledge of any participating anesthesia providers, canisters of both absorbents were tested in anesthesia machines throughout our academic institution. Times and dates were recorded when these absorbents were changed. Changes were made when CO_2 rebreathing signaled the exhaustion of absorbent ($EtCO_2$ = 5 cm H2O). Then, the total time and fresh gas flow (FGF) rates used during general endotracheal anesthesia for each canister of absorbent were obtained through retrospective chart review. Because the rate at which CO_2 absorbent is exhausted is greater with lower FGFs, the total time spent at a given FGF rate was divided by the FGF rate to give a weighted time value. Thus in our calculations, the time the absorbent was utilized at 4 l/min. contributed only half as much to the time measured until absorbent exhaustion when compared to the time spent at 2 l/min. The average of these time weighted values for each brand was calculated, their coefficients of variance, the 95% confidence interval for those means, and p-value for the likelihood that the means of the brands differ using the student's 2-sample t-test.

Results: The mean flow corrected times for Sodasorb LF and Amsorb Plus were 1811.8 and 1088.7 respectively. The coefficients of variance for their means were 1624 and 209 respectively. Under clinical conditions, no statistically significant difference in performance was measured between Sodasorb LF and Amsorb Plus (p=0.132).

Discussion: While in vitro differences between various absorbents can be measured, the clinical performance is more important when making decisions regarding which absorbents to use. Only when clinical performances are known can sound decisions be made when evaluating which products are most efficient, and most economical. Difficulty in comparing two samples of absorbent arises from uncertainty in exactly how much CO_2 a sample of absorbent has been exposed to. Wide variations in CO_2 production occur between patients, and we have not attempted to correct for them in the present study. By blinding those using the absorbents, we were unable to perform the study under uniform FGFs, as has been done in previous in vitro studies[1]. These factors may explain the large coefficient of variance seen in this study.

References: 1. Woehlck, HJ et al. Anesthesiology 2005; 103: A1164

Meta-analysis of diagnostic tests: Two-stage hierarchic model for combining likelihood ratios for positive and negative test results

Srinivas Mantha, MD*, Joseph Foss, MD,†, John E. Ellis, MD,‡, Michael F. Roizen, †

Depts.of Anesthesiology, *Nizam's Institute of Medical Sciences, Hyderabad, India, † The Cleveland Clinic Foundation, Cleveland, OH, USA, ‡The University of Chicago, Chicago, IL

Abstract

Introduction: Traditionally, diagnostic tests are employed to "rule-out" or "rule-in" the underlying disease states. Alternatively, in the context of perioperative care, diagnostic tests are also used to predict adverse postoperative outcomes resulting from the underlying disease states. For diagnostic tests summarized in 2 x 2 tables, sophisticated methods of meta-analysis that use relative diagnostic odds ratios, summary receiver operating characteristic curve and likehood ratio scatter plots have been proposed (1) and employed (2,3) for better evaluation of the discriminative value of diagnostic tests. A combined point estimate and 95% confidence intervals (CIs) for the likehood ratio (LR) for a positive test (LR-pos) and LR for a negative test (LR-neg) would also be useful.

Methods: We propose a two-stage hierarchic model (4) to combine information about LR separately for positive and negative test results i.e for LR-pos and for LR-neg. An empirical Bayes procedure with a normal-normal hierarchic model was used get a meta-analytic confidence interval for overall median for LR-pos and LR-neg. For a single study, 95%CIs for LR-pos and LR-neg were computed using standard formulas (5). The variance between the studies was estimated using method of moments approach described by DerSimonian and Laird (6). The methodology was applied to the data summarized in a recent meta-analysis that evaluated accuracy of six diagnostic tests for predicting perioperative cardiac risk in patients undergoing major vascular surgery (2)

Results: The combined summary measure point estimate and 95% CIs for each of LR-pos and LR-neg are given in table 1.

Table 1

Test	LR-Pos (95% CI)	LR-Neg (95% CI)
Ambulatory ECG	1.81 (1.30 to 2.51)	0.99 (0.86 to 1.13)
Exercise ECG	2.72 (1.85 to 4.00)	0.51(0.30 to 0.86)
RNV	6.52 (2.43 to 17.51)	0.72 (0.54 to 0.97)
MPS	1.65 (1.43 to 1.89)	0.49 (0.34 to 0.71)
Dipyridamole stress echo	4.96 (2.31 to 10.62)	0.40 (0.17 to 0.90)
Dobutamine stress echo	3.03 (2.06 to 4.46)	0.35 (.23 to 0.51)

RNV- radionuclide ventriculography,
MPS- Myocardial perfusion scintigraphy

Discussion: Of the several variables used to evaluate discriminative ability of diagnostic tests, LR-pos and LR-neg provide more useful information as they precisely describe both directions of test performance. For peripberative risk stratification, LR-pos and LR-neg should be greater than 10 or less than 0.2 respectively, because these values indicate a substantial change in risk from the pretest level (7). Ideally, in meta-analysis, the combined summary ratio and CIs must be greater than 10 for LR-pos and less than 0.2 for LR-neg. Using such interpretation, even the dobutamine stress echocardiography, which was concluded as the best among the 6 tests for prediction of adverse cardiac outcome after vascular surgery (2), cannot be regarded as the ideal test.

References:

1. Stengel D et al. J Med Screening. 2003;10 :47-51
2. Kertai MD et al. Heart 2003:89; 1327-34
3. Beattie SW et al. Anesth Analg 2006;102:8-16
4. Mantha S et al. Anesth Analg 1994;79 :422-33
5. Simel DL et al. J Clin Epidemiol 1991;44: 763-70
6. DerSimonian R and Laird N. Control Clin Trials 1986;7: 177-88.
7. Grayburn PA and Hills D. Ann Intern Med 2003;138: 506-511.

Introduction and Methods

Introduction

Meta-analytic studies form an important tool in evidence-based medicine. Traditionally, diagnostic tests are employed to "rule-out" or "rule-in" the underlying disease states. Alternatively, in the context of perioperative care, diagnostic tests are also used to predict adverse postoperative outcomes resulting from the underlying disease states. For diagnostic tests summarized in 2 x 2 tables, sophisticated methods of meta-analysis that use relative diagnostic odds ratios, summary receiver operating characteristic curve and likehood ratio scatter plots have been proposed (1) and employed (2,3) for better evaluation of the discriminative value of diagnostic tests. A combined point estimate and 95% confidence intervals (CIs) for the likehood ratio (LR) for a positive test (LR-pos) and LR for a negative test (LR-neg) would also be useful.

Methods:

We propose a two-stage hierarchic model (4) to combine information about LR separately for positive and negative test results i.e for LR-pos and for LR-neg.

An empirical Bayes procedure with a normal-normal hierarchic model was used get a meta-analytic confidence interval for overall median for LR-pos and LR-neg. For a single study, 95%CIs for LR-pos and LR-neg were computed using standard formulas (Figure 1) (5). The variance between the studies was estimated using method of moments approach described by DerSimonian and Laird (6).

Figure 1

	Disease	Normal
Test postive	A	B
Test negative	C	D
	n1	n2

sensitivity (sn)= A/n1
specificity (sp)= D/n2

Formula to compute point estimate and 95% CIs for likelihood ratio in general for a single study is

$$LRx = \exp\left(\ln \frac{p1}{p2} \pm 1.96 \sqrt{\frac{1-p1}{p1n1} + \frac{1-p2}{p2n2}}\right)$$

For LR-pos, p1=sn and p2=1-sp
For LR-neg, p1=1-sn and p2=sp

Figure 1 (Contd)

<u>The formula for 95% confidence intervals for positive likehood ratios is</u>

$$LR+ = \exp\left(\ln \frac{\text{sensitivity}}{1-\text{specificity}} \pm 1.96 \times \sqrt{\frac{1-\text{sensitivity}}{A} + \frac{\text{specificity}}{B}}\right)$$

where A and B are values of the respective cells from the 2 x 2 table and exp is the base of the natural logarithm.

<u>The formula for 95% confidence interval for negative likelihood ratio is</u>

$$LR- = \exp\left(\ln \frac{1-\text{sensitivity}}{\text{specificity}} \pm 1.96 \times \sqrt{\frac{\text{sensitivity}}{C} + \frac{1-\text{specificity}}{D}}\right)$$

Where C and D are the values of the respective cells from the 2 x 2 table and exp is the base of the natural algorithm

<u>The two-stage Procedure</u>

•The two-stage hierarchic model was used to combine information about the LR-pos and LR-neg of test from various studies. In the following description LR implies separate computations for LR-pos and LR-neg.
•The quantity combined was not LR but rather log of LR. Taking the logarithm of LR makes the hierarchic model more appropriate, particularly for the first-stage model.
•The model assumes that the studies are exchangeable i.e., we are assuming a priori that the studies are equivalent in regard to their outcomes.

First-Stage Model

$$dj \mid Dj = dj \sim N(Dj, \sigma_j^2) j = 1, \cdots, J$$

Where dj is an estimator of dj. We treat dj as a realization of the random variable Dj; i.e., we treat log LR as a random variable, and the true log LR for a study as a realization of the random variable Dj. In the second stage of the hierarchy Dj is assumed to have a normal distribution, i.e., we assume that the study effects have a normal distribution(study effect is the log LR of a study)

Second-Stage Model

$$E(Dj \mid \mu, \Delta^2) \equiv \mu$$
$$V(Dj \mid \mu, \Delta^2) \equiv \Delta^2$$
$$Dj \mid \mu, \Delta^2 \sim N(\mu, \Delta^2)$$

•We then calculated conditional mean (i.e. combined LR from all the studies) and conditional CI with variance estimated by the method-of moments described by DerSimonian and Laird (5)
•Computational scheme is described in our earlier study (4)
•The methodology was applied to the data summarized in a recent meta-analysis that evaluated accuracy of six diagnostic tests for predicting perioperative cardiac risk in patients undergoing major vascular surgery (2)

Results and Discussion

Results: The combined summary measure point estimate and 95% CIs for each of LR-pos and LR-neg are given in table 1.

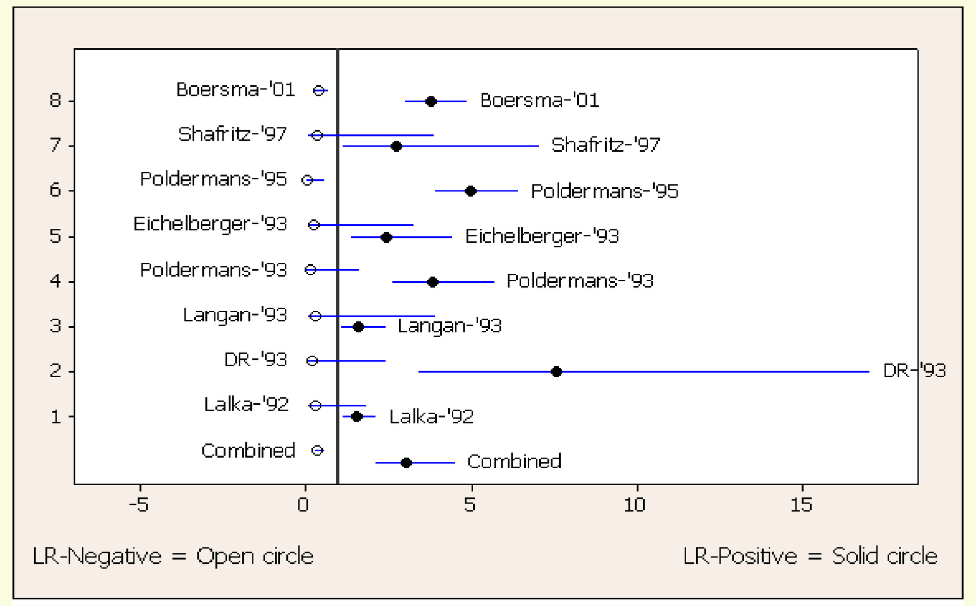
Table 1

Test	LR-Pos (95% CI)	LR-Neg (95% CI)
Ambulatory ECG	1.81 (1.30 to 2.51)	0.99 (0.86 to 1.13)
Exercise ECG	2.72 (1.85 to 4.00)	0.51(0.30 to 0.86)
RNV	6.52 (2.43 to 17.51)	0.72 (0.54 to 0.97)
MPS	1.65 (1.43 to 1.89)	0.49 (0.34 to 0.71)
Dipyridamole stress echo	4.96 (2.31 to 10.62)	0.40 (0.17 to 0.90)
Dobutamine stress echo	3.03 (2.06 to 4.46)	0.35 (.23 to 0.51)

RNV- radionuclide ventriculography,
MPS- Myocardial perfusion scintigraphy

Figure 2 depicts the graphical display data for dobutamine stress Echocardiography stuides

Figure 2



Discussion: Of the several variables used to evaluate discriminative ability of diagnostic tests, LR-pos and LR-neg provide more useful information as they precisely describe both directions of test performance. For peripberative risk stratification, LR-pos and LR-neg should be greater than 10 or less than 0.2 respectively, because these values indicate a substantial change in risk from the pretest level (7). Ideally, in meta-analysis, the combined summary ratio and CIs must be greater than 10 for LR-pos and less than 0.2 for LR-neg. Using such interpretation, even the dobutamine stress echocardiography, which was concluded as the best among the 6 tests for prediction of adverse cardiac outcome after vascular surgery (2), cannot be regarded as the ideal test.

References

1. Stengel D, Bauwens K, Sehouli J, Ekkernkamp A, Porzsolt F. A likelihood ratio approach to meta-analysis of diagnostic studies. J Med Screen 2003;10:47-51.
2. Kertai MD, Boersma E, Bax JJ, Heijenbrok-Kal MH, Hunink MG, L'Talien G J, Roelandt JR, van Urk H, Poldermans D. A meta-analysis comparing the prognostic accuracy of six diagnostic tests for predicting perioperative cardiac risk in patients undergoing major vascular surgery. Heart 2003;89:1327-34.
3. Beattie WS, Abdelnaem E, Wijeysundera DN, Buckley DN. A meta-analytic comparison of preoperative stress echocardiography and nuclear scintigraphy imaging. Anesth Analg 2006;102:8-16.
4. Mantha S, Roizen MF, Barnard J, Thisted RA, Ellis JE, Foss J. Relative effectiveness of four preoperative tests for predicting adverse cardiac outcomes after vascular surgery: a meta-analysis. Anesth Analg 1994;79:422-33
5. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. J Clin Epidemiol 1991;44:763-70.
6. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177-88.
7. Grayburn PA, Hillis LD. Cardiac events in patients undergoing noncardiac surgery: shifting the paradigm from noninvasive risk stratification to therapy. Ann Intern Med 2003;138:506-11.