Accuracy of glucose breath testing for small intestine bacterial overgrowth (SIBO) using endoscopy aspirate cultures as a reference standard: A meta-analysis

Jack Stylli BSc¹, Shaoying Nikki Lee PhD¹, Collen Kelly PhD², Sharat Singh PhD¹, Chris Wahl MD¹, Emil Chuang MD¹, and Mitchell Jones MD PhD¹ ¹Progenity, Inc. Ann Arbor, Michigan, United States; ²Kelly Statistical Consulting, San Diego, California, United States

Introduction

Small Intestinal Bacterial Overgrowth (SIBO) is associated with abnormally high bacterial counts in the small intestine. SIBO is under-diagnosed and there are significant limitations with currently available testing methods.

Methodology

A literature search was performed to find relevant studies using the following keywords: glucose breath test accuracy, endoscopy aspirate culture, and small intestine bacterial overgrowth. Eleven studies with data on patients tested with both a glucose breath test and endoscopy aspirate culture were found. Fixed-effects and random-effects models were fit to the eleven studies' data using the Metafor meta-analysis package for R. The mean positive and negative percent agreement across studies is estimated as the weighted average of each study's percent agreement, where the weights are the inverse of the standard errors, squared, plus the variance of the random study effects.

Study	Ν	Glucose administration	Length of breath test	Positive percent agreement	Negative percent agreement
Donald et al. 1992	52	50g	120 min.	5/26=20%	20/26=77%
Bauer et al. 2000	40	75g	120 min.	12/29=41%	5/11=45%
Berthold et al. 2009	21	50g	120 min.	5/12=42%	4/9=44%
Corazza et al. 1990	20	75g	180 min.	5/8=62%	10/12=83%
Erdogan et al. 2015	139	75g	120 min.	17/25=68%	87/114=76%
Ghoshal et al. 2006	73	Not stated	Not stated	12/29=41%	36/44=82%
Ghoshal et al. 2014	80	100g	180 min.	4/15=27%	65/65=100%
Kerlin et al. 1988	45	50g 50g	120 min. 240 min.	25/27=93% 25/27=93%	14/18=78% 13/18=72%
MacMahon et al. 1996	30	50g	120 min.	15/20=75%	3/10=30%
Metz et al. 1976	17	50g	120 min.	8/12=67%	5/5=100%
Stotzer et al. 2000	46	50g	120 min.	14/24=58%	19/22=86%

Table 1. Summary of the eleven studies included in the meta-analysis.

Threshold of a change of 12ppm were established on the basis of control data; since the classifier is trained and tested on the same set of data, agreement is likely biased upward.

Results

The homogeneity of effects across studies was tested with a Chi-square test and homogeneity of effects was rejected for both positive and negative percent agreement (c2 = 91.5 and 107.5, respectively, df = 10, p-value < 0.0001) (**Table 1**). The mean percent agreement across studies was estimated with both the fixed-effects and the random-effects models (Table 2). Figures 1A and 1B present each study's estimate of the positive and negative percent agreement, respectively, and their associated 95% confidence intervals, as well as the overall estimate and its 95% confidence interval.

"SIBO is under-diagnosed and there are significant limitations with currently available testing methods."

A. Positive percent agreement



Figure 1. Forest plots showing the percentage agreement of 11 studies comparing the result of glucose breath test and endoscopy aspirate culture.

This figure shows the estimated mean (A.) positive percent agreement and (B.) negative percent agreement with corresponding 95% confidence intervals for the individual studies and overall based on a random-effects (RE) model. The polygon at the bottom shows the summary estimate based on the model (with the outer edges of the polygon indicating the confidence interval limits.) The square sizes are drawn proportional to the precision of the estimates

Table 2. Fixed-effects and random-effects model estimates of the mean positive and negative percent agreement combining eleven studies.

Model	Mean positive percent agreement (95% CI)	Mean negative percent agreement (95% CI)
Fixed-effects model	61% (55%, 66%)	99% (98%, 99.9%)
Random-effects model	54% (40%, 68%)	76% (65%, 88%)

"Novel tools are needed for evaluating patients with suspected SIBO."

progenity®

B. Negative percent agreement



Conclusions

- Heterogeneity was found in the study designs and the implementation of breath tests, such as the differences in dose of the substrate (50g or 75-100g) and the duration of the breath test (120-240min).
- In addition, the endoscopy aspirate culture is not standardized, as the amount of liquid collected, the site of collection and the technical details of the microbiological tests may differ.
- This heterogeneity in practice of the two tests certainly contributes to the heterogeneity in positive and negative percentage agreement observed between the two tests.
- ► The random effects models' estimates of 54% positive agreement and 76% negative agreement indicate moderate to poor agreement between the breath test and endoscopy aspirate culture.
- Given these limitations, there is a consensus to standardize the process, and novel tools are needed for evaluating patients with suspected SIBO.