ACTIVITY: Test Validity Rests in Evidence, Part 2

Instructions:

The **goal** of this activity is to give you experience with an in-depth analysis and comparison of two diagnostic accuracy studies. You will complete it in two parts. In Part 1, you will identify two diagnostic accuracy studies and complete a comparison table. In Part 2, you will compare your findings. Follow the steps for each part below.

Relevant articles for this assignment:

- Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Better decisions through science. *Scientific American*, 283(4), 82-87.
- Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., ... & STARD Group*. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Annals of internal medicine*, *138*(1), 40-44.
- Vermiglio, A. J. (2014). On the clinical entity in audiology: (Central) auditory processing and speech recognition in noise disorders. *Journal of the American Academy of Audiology*, 25(09), 904-917.
- Vermiglio, A. J. (2016). On diagnostic accuracy in audiology: Central site of lesion and central auditory processing disorder studies. *Journal of the American Academy of Audiology, 27*(02), 141-156. (With errata)

Vermiglio, A. J., & Fang, X. (2022). The World Health Organization (WHO) hearing impairment guidelines and a speech recognition in noise (SRN) disorder. *International Journal of Audiology*, *61*(10), 818-825. [Note: This article is basically a tutorial on diagnostic accuracy studies.]
[Note: Do not confuse reliability with validity. Reliability refers to the repeatability of a test. <u>Validity</u> refers to the ability of a test to measure what it is supposed to measure.]

<u>Part 1</u>

 Find a diagnostic accuracy study on a test used in audiology or speech-language pathology for the detection of a specific target disorder. Recall from the video lecture that the terms diagnostic accuracy, index test, target disorder, and "gold" or reference standard test may not appear in these types of articles.

Suggested search strategies include using Boolean cues in Google Scholar.

For example: "sensitivity" AND "specificity" AND "name of index test"

"sensitivity" AND "specificity" AND "name of the target disorder"

At minimum the article should report the sensitivity and specificity of the index test or tests, it should contain a control group and at least one disordered group and it should report the index test results including the means, standard deviations, and the number of subjects.

- 2. Find a second article with the same index test and/or target disorder as the first article for comparison.
- 3. Use the information from the article(s) to complete the data fields in the Excel Table template. *Data Field Descriptions*

Author(s) (year) – article citation

Reference – complete article reference

Diagnostic system - index test, target disorder, "gold" or reference standard test

Index test – the test under evaluation. This may appear in the form of an index test battery with multiple tests.

Target disorder – This may also be called a clinical entity (see Vermiglio, 2014) or diagnostic target.

Index test cut-point – This value delineates between the positive (disorder is presence) and negative (disorder is absent) index test results.

"Gold" or reference test – This is the considered the best method for the identification of the presence and absence of the target disorder. In a diagnostic accuracy study, this is the test (or test battery) used to sort research participants into the control and disordered groups.

"Gold" or reference test cut-point – This value delineates between participants with and without the target disorder.

Control group – The group without the target disorder as identified by the "gold" or reference standard test.

Disordered group – The group with the target disorder as identified by the "gold" or reference standard test.

Index test group difference – This is the most elemental calculation in a diagnostic accuracy study. A statistically significant index test result group difference indicates that, to some degree, the index test can delineate between participants with and without the disorder.

Sensitivity – This is the same as the percentage of true positive index test results.

Specificity – This is the same as the percentage of true negative index test results.

Convergent validity – This describes relationship between index test vs "gold" or reference standard test results. A strong and statistically significant relationship indicates that the results of the reference standard test may be predicted from the index test results.

Area under the curve (AUC) – This describes the strength of the index test as a predictor of the target disorder. An AUC of 1 indicates perfect separation of the participants into groups with and without the disorder. An AUC of 0.5 indicates that the index test has a 50/50 chance of making a correct diagnosis of the target disorder (no better than flipping a coin). See Swets et al. (2000) for an explanation.

<u>Part 2</u>

Write a report that compares and analyzes the findings from the diagnostic studies.

1. Introduction

- a. Introduce the topic and explain why it is relevant for audiologists and/or speechlanguage pathologists. Include the appropriate article citations.
- b. Give some background information on the purpose of diagnostic accuracy studies.
- c. Describe the components of a diagnostic system.
- d. Describe how the diagnostic system is used to determine test validity.
- e. Explain why it is important for a clinical audiologist or speech-language pathologist to know the diagnostic accuracy or validity of their diagnostic tests.

2. Study Overview: Article 1

Describe the goal of first research article. Give some information on the target disorder. Explain the importance of identifying the presence of this disorder. Include the relevant citations.

3. Components of the Diagnostic System(s): Article 1

- a. Index Test(s) or Index Test Battery:
 - i. Give the details of the index test or test battery.
 - ii. Was the index test or index test battery described clearly enough to allow for replication of the study?
 - iii. Include a table showing the diagnostic system(s) used in the study.

Table 1

Diagnostic Systems Used in Article 1

| Author, year | Diagnostic | Index Test | Target | "Gold" or Reference |
|--------------|------------|------------|----------|---------------------|
| | System | | Disorder | Standard Test |
| | 1 | | | |
| | 2 | | | |
| | 3 | | | |
| | 4 | | | |

b. Target Disorder:

Is the target disorder a clinical entity according to the Sydenham-Guttentag criteria? Briefly describe each criterion and how it relates to the target disorder (see Vermiglio, 2014 for two examples).

- c. "Gold" or Reference Standard Test:
 - i. Did the author(s) give any evidence showing that the "gold" or reference standard test is valid? In other words, did they indicate that the "gold" or reference standard test is the best way or one of the best ways to identify the presence and absence of the target disorder? For example, was the reference standard test used as an index test in a prior study showing good diagnostic accuracy (sensitivity, specificity) of this test?
 - ii. Was the "gold" or reference standard test described clearly enough to allow for replication?
- d. Participant Characteristics:
 - i. What was the inclusion criteria for the study participants?
 - ii. How many participants were included in each group?
 - iii. What was the mean age of the participants?
 - iv. Was the index test and reference standard test administered to all subjects in the control (non-disordered) and disordered group?
 - v. Did the control group include those without the disorder?
 - vi. Did the disordered group contain those with the disorder?
 - vii. Did the control or disordered subjects have other disorders or conditions that may have affected the index and reference standard test results?

4. Results: Article 1

- a. Enter the appropriate results in Tables 2 and 3 below.
- b. Describe the descriptive statistics (from Table 2) for the index test and reference standard test results (mean and SD) for each group (required).
- c. Describe the minimum maximum, range, (if available). Note: if not provided, this may be estimated from a scatter plot.
- d. Describe the sensitivity and specificity of the index tests (from Table 3).
- e. Describe the group differences in index test performances and *p*-values if available.
- f. Describe the correlation coefficients between the reference standard test vs. the index test results and the *p*-values (if available).
- g. Do the correlation coefficients (if available) indicate that the reference standard test results can be predicted from the index test result?
- h. Describe the receiver operating characteristics (ROC) curve(s) and the area under the curve (AUC; if available).

Table 2

| Article | 1 | Descriptive | Statistics |
|---------|---|-------------|------------|
|---------|---|-------------|------------|

| Author, year | Group | Test | Mean (unit of measure) | Standard Deviation | Maximum Result (unit of measure) | Minimum Results (unit of measure) | Range (e.g., dB or percentage points) |
|-----------------|------------|-------|------------------------------|-----------------------|---|--|--|
| | Control | Indov | | | | | |
| | Disordered | muex | | | | | |

| Control | Gold | Not | | |
|------------|----------|------------|--|--|
| Disordered | Standard | applicable | | |

Table 3

Article 1 Results of the Diagnostic Accuracy Study

| Author, year | Diagnostic System | Index Test | Sensitivity | Specificity | Index Test vs. Reference Standard Test Results, Correlation Coefficients (p-values in parentheses) | Area Under the Curve (p- value) |
|--------------|----------------------|---------------|-------------|-------------|---|--|
| | 1 | | | | | |
| | 2 | | | | | |
| | 3 | | | | | |
| | 4 | | | | | |

5. Discussion: Article 1

Discuss the implications of the results for your work as a clinician and/or researcher. Include the appropriate citations throughout.

6. Critique: Article 1

- a. Describe any limitations or shortcomings of the article.
- b. Describe anything that was unclear in the article.
- c. Describe the generalizability of the study results. Are there limitations in using the information from this study to your clinical patients?

7. Study Overview: Article 2

Describe the goal of second research study. Give some information on the target disorder. Explain the importance of identifying the presence of this disorder. Include the relevant citations.

8. Components of the Diagnostic System(s): Article 2

- a. Index Test(s) or Index Test Battery:
 - i. Give the details of the index test or test battery.
 - ii. Was the index test or index test battery described clearly enough to allow for replication of the study?
 - iii. Include a table showing the diagnostic system(s) used in the study.

Table 4

Diagnostic Systems Used in Article 2

| Author, year | Diagnostic System | Index Test | Target Disorder | "Gold" or Reference Standard Test |
|--------------|----------------------|------------|--------------------|--------------------------------------|
| | 1 | | | |
| | 2 | | | |
| | 3 | | | |

| 4 | | |
|---|--|--|
| | | |

b. Target Disorder:

Is the target disorder a clinical entity according to the Sydenham-Guttentag criteria? Briefly describe each criterion and how it relates to the target disorder (see Vermiglio, 2014 for two examples).

- a. "Gold" or Reference Standard Test:
 - i. Did the author(s) give any evidence showing that the "gold" or reference standard test is valid? In other words, did they indicate that the "gold" or reference standard test is the best way or one of the best ways to identify the presence and absence of the target disorder? For example, was the reference standard test used as an index test in a prior study showing good diagnostic accuracy (sensitivity, specificity) of this test?
 - ii. Was the "gold" or reference standard test described clearly enough to allow for replication?
- b. Participant Characteristics:
 - iii. What was the inclusion criteria for the study participants?
 - iv. How many participants were included in each group?
 - v. What was the mean age of the participants?
 - vi. Was the index test and reference standard test administered to all subjects in the control (non-disordered) and disordered group?
 - vii. Did the control group include those without the disorder?
 - viii. Did the disordered group contain those with the disorder?
 - ix. Did the control or disordered subjects have other disorders or conditions that may have affected the index and reference standard test results?

9. Results: Article 2

- a. Enter the appropriate results in Tables 5 and 6 below.
- b. Describe the descriptive statistics (from Table 5) for the index test and reference standard test results (mean and SD) for each group (required).
- c. Describe the minimum maximum, range, (if available). Note: if not provided, this may be estimated from a scatter plot.
- d. Describe the sensitivity and specificity of the index tests (from Table 6).
- e. Describe the group differences in index test performances and *p*-values if available.
- f. Describe the correlation coefficients between the reference standard test vs. the index test results and the *p*-values (if available).
- g. Do the correlation coefficients (if available) indicate that the reference standard test results can be predicted from the index test result?
- h. Describe the receiver operating characteristics (ROC) curve(s) and the area under the curve (AUC; if available).

Table 5

Article 2 Descriptive Statistics

| Author, year | Group | Test | Mean (unit of measure) | Standard Deviation | Maximum Result (unit of measure) | Minimum Results (unit of measure) | Range (e.g., dB or percentage points) |
|-----------------|------------|----------|------------------------------|-----------------------|---|--|--|
| | Control | Indox | | | | | |
| | Disordered | muex | | | | | |
| | Control | Gold | Not | | | | |
| | Disordered | Standard | applicable | | | | |

Table 6

Article 2 Results of the Diagnostic Accuracy Study

| Author, year | Diagnostic | Index | Sensitivity | Specificity | Index Test vs. Reference | Area |
|--------------|------------|-------|-------------|-------------|---------------------------------|-----------|
| | System | Test | | | Standard Test Results, | Under the |
| | | | | | Correlation Coefficients | Curve (p- |
| | | | | | (p-values in parentheses) | value) |
| | 1 | | | | | |
| | 2 | | | | | |
| | 3 | | | | | |
| | 4 | | | | | |

10. Discussion: Article 2

Discuss the implications of the results for your work as a clinician and/or researcher. Include the appropriate citations throughout.

11. Critique: Article 2

- a. Describe any limitations or shortcomings of the article.
- b. Describe anything that was unclear in the article.
- c. Describe the generalizability of the study results. Are there limitations in using the information from this study to your clinical patients?

12. Comparison of Two Diagnostic Accuracy Studies

Complete Table 7 to compare the results of each diagnostic accuracy study.

Table 7

Comparison of the Diagnostic Accuracy Results Found in Two Different Diagnostic Accuracy Studies

| Author, | Diagnostic | Index | Target | Reference | Sensitivity | Specificity | Index Test. | Area |
|---------|------------|-------|----------|-----------|-------------|-------------|-------------|-------|
| year | System | Test | Disorder | Standard | | | Vs. | Under |
| | | | | Test | | | Reference | the |
| | | | | | | | Standard | Curve |
| | | | | | | | Correlation | |

| | | | Coefficients (p-values) | (p- value) |
|---|--|--|----------------------------|---------------|
| 1 | | | | |
| 2 | | | | |