Chapter 2
Screening Basics: Differentiating a Screen from a Diagnostic Tool

Objectives: To provide knowledge and understanding regarding differentiating a screen from a diagnostic tool and why a reliable and validated swallow screen is valuable for patient care.

Methods: Definition of statistical terms necessary for understanding the bases for swallow screening.

Results: Terms defined were sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio, negative likelihood ratio, false positive rates, and false negative rates.

Conclusions: Understanding the statistical underpinnings pertinent to swallow screening will provide the clinician with the knowledge necessary to make informed decisions regarding use of the best swallow screen for patient care.

Keywords: Deglutition, Deglutition disorders, Swallow screening, Statistics, Oral alimentation

Introduction

Is Screening for Aspiration Risk Unique?

The direct and accurate answer is No. Dysphagia specialists like to think that screening tests for aspiration risk are different from, for example, screening tests for pregnancy, high blood pressure,

S.B. Leder and D.M. Suiter, *The Yale Swallow Protocol: An Evidence-Based Approach to Decision Making*, DOI 10.1007/978-3-319-05113-0_2,

tuberculosis, or diabetes but screening for aspiration risk is no different from screening for any other potential symptom, condition, or disease that is screened for in medicine. What is critically important, however, is use of a validated, reliable, and generalizable screen that can be used for the vast majority of at-risk individuals. You must then trust the screen to do its job and abide by its results. If this is not the case, then the screen is useless and you must test all referred individuals with the more expensive, time-consuming, and personnel-laden instrumental evaluation which in the case of dysphagia is either an endoscopic or videofluoroscopic procedure.

In order to be beneficial, a screening test must be easy to administer, accurate, and require less time, expense, and staff resources than a diagnostic test. In medicine, screening tests are generally given to groups of asymptomatic people to detect potential disease indicators. Groups can range from the population as a whole to more individualized, case-finding approaches focusing on individuals at-risk for a specific target condition [1]. Conversely, the purpose of a diagnostic test is to confirm the presence or absence of the target condition. A screening test, therefore, will have high sensitivity but low specificity, thereby allowing detection of most patients with the target condition while having the acceptable disadvantage of a high false positive rate. Subsequent diagnostic testing will eliminate the over-referrals, i.e., false positives, from receiving treatment.

A screen has both different goals and different endpoints than a diagnostic evaluation. A screen is a simple, noninvasive, and inexpensive method of detecting the probability that an individual has a disease. For the purposes of this book just substitute aspiration risk for disease. A screen cannot and does not provide a definitive diagnosis. Typically, screening tests are given to large groups of asymptomatic people, whereas diagnostic tests are given either to symptomatic individuals to establish a diagnosis or to asymptomatic individuals who fail a screening.

In clinical practice, it is neither practical nor feasible to screen everyone. Therefore, it is customary for a screening test to adopt a more individualized or case-finding approach [1]. In case-finding, individuals who are considered to be at high risk for a particular

condition are tested. For instance, it is common knowledge that health-care workers are considered to be at higher risk than the general population for exposure to tuberculosis and, therefore, receive an annual screening test for that disease. Other examples of screening tests that use a case-finding approach include mammograms for women over age 40 years and colonoscopies for individuals over age 50 years. *It is important to note that when a screening test is passed no further treatment or assessment is indicated*. Only when a screening test is failed is the more extensive and expensive diagnostic examination performed.

The use of a screening test is based upon two general assumptions [1]. First, the course of the targeted disorder will, if undetected and untreated, result in serious and preventable health problems. For example, undetected prandial aspiration has the potential to cause a number of negative health outcomes including aspiration pneumonia, acute respiratory failure, dehydration, malnutrition, and sepsis with the possibility of leading to death. Therefore, early identification and management is essential in prevention and health maintenance. Second, treatment for the target condition being screened exists and is effective. Again, in the case of prandial aspiration, effective interventions such as postural changes, bolus volume adjustments, and viscosity alterations can be implemented with the goal of either eliminating or minimizing its consequences. A nil per os, i.e., no food or drink by mouth, order is generally effective when prandial aspiration is unresolved and before instrumental testing is performed.

A screening test yields a binary response. A positive result, i.e., pass, indicates that further testing is necessary to determine if the disease or condition being screened for is truly present. A negative result, i.e., fails, indicates that the individual most likely does not have the disease or condition being screened for and further testing is not necessary. Specific to swallow screening, the binary pass or fail determination will identify an individual who either is in no need of additional testing and is not an aspiration risk or requires further diagnostic assessment to confirm a swallowing disorder. A diagnostic swallowing test is not a pass or fail procedure. Rather, the purpose is to identify the underlying pathophysiology

of the swallowing disorder that results in signs or symptoms of dysphagia and then to determine an appropriate treatment to address this pathophysiology with the goal of promoting safe and successful swallowing.

Why Use a Screening Test to Determine Aspiration Risk?

Despite the fact that pulmonary aspiration remains a leading cause of nosocomial infection in the critically ill [2], it is neither medically necessary nor fiscally defensible to perform instrumental swallow testing on all patients referred for swallowing testing. It is, however, vitally important to have a plan with the goal of identifying those specific patients who may have a high aspiration risk prior to initiating oral alimentation and medications.

All patients deemed to have potential swallowing problems should be screened for aspiration risk. The goal is to administer a reliable swallowing screen to all appropriately identified patients, deemed appropriate to begin oral alimentation, before starting oral ingestion of foods, fluids, or medications. This is especially important as many patients have an a priori increased aspiration risk due to comorbidities and concomitant interventions including short- (<24 h) [3] or longer-term (>24 h) endotracheal intubation [4–6], traumatic brain injury [5], other cognitive issues [7], severe deconditioning and reduced functional reserve [8], medication side effects [8], and advanced age [9, 10].

The use of screening tools to identify patients with potential aspiration risk has garnered substantial interest from a number of health-care organizations, including the American Heart Association, the Veterans Health Administration, and the American Speech-Language-Hearing Association, who recognize the need for early identification and appropriate intervention for identified individuals. Use of screening tests to detect aspiration risk is supported in that hospitals which used a mandatory and formal swallow screening procedure reported lower rates of pneumonia, and concomitant cost savings than those without [11–13].

Accuracy of Screening Tests

In order to determine accuracy of a screening tool, a 2 × 2 contingency table is frequently used. Such a table allows for results of the screening tool to be compared to those of a reference standard test. Results are classified as either positive or negative and the criterion for each of these is set a priori. Table 2.1 shows a typical 2 × 2 table which includes:

1. The number of true positives or those individuals that have the disease and test positive.
2. The number of false positives or those individuals that do not have the disease and test positive.
3. The number of false negatives or those individuals that do have the disease and test negative.
4. The number of true negatives or those individuals that do not have the disease and test negative.

Both true and false positives and negatives are used to calculate the statistical measures of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). These measures provide information about the validity of a test. The accuracy of a screening test is generally expressed in terms of its sensitivity and specificity. A high sensitivity, therefore, is good.

TABLE 2.1. A typical 2 × 2 contingency table.

		Reference standard (FEES or VFSS)	
		Positive	Negative
Screening	Positive	a or True Positives (TP)	b or False Positives (FP)
	Negative	c or False Negatives (FN)	d or True Negatives (TN)

Sensitivity = a/(a + c)
Specificity = d/(b + d)
PPV = a/(a + b)
NPV = d/(c + d)
Positive Likelihood Ratio = Sensitivity/(1 − Specificity)
Negative Likelihood Ratio = (1 − Sensitivity)/Specificity

On the other hand, low specificity, i.e., a low false negative rate, is an inherent shortcoming for all screening tests so as not to miss people who actually have the target condition. Similarly, an acceptable disadvantage of a screening test is a high rate of false positives, which means that screening tests often over-refer individuals for full diagnostic testing. Ideally, a reliable and clinically useful screen for aspiration risk should have a high sensitivity, a high negative predictive value, and a low false negative rate.

In contrast, diagnostic or criterion standard tests should have high sensitivity in order to identify people who truly have the target condition as well as high specificity in order to eliminate people who truly do not have the target condition. It must be remembered that the goal of a diagnostic tests is to make a definitive diagnosis.

Sensitivity

Sensitivity of a test is the proportion of individuals with the target condition who have a positive test result. In other words, sensitivity measures a test's ability to identify an individual with the disease as positive. Tests that are found to be highly sensitive indicate that there are few false negative results and thus fewer cases of disease are missed [14]. A test with high sensitivity means that a negative result rules out the target condition. Tests with high sensitivity are designed to identify everyone with a particular condition even if some people are identified who do not actually have the condition, i.e., false positives. Over-referral of these false positives is considered an acceptable outcome of screening and subsequent diagnostic testing eliminates them from receiving treatment.

Specificity

Specificity of a test is the proportion of individuals without the target condition who have a negative test result. In other words, specificity measures a test's ability to identify an individual without the disease as negative. Tests that are found to be highly specific indicate fewer

false positive result and are able to help rule in the disorder of interest [14]. A test with high specificity means that a positive result rules in the patient as most likely having the target condition.

Positive and Negative Predictive Values

The accuracy of screening tests may also be reported in terms of predictive values. A positive predictive value is the probability that an individual with a positive, i.e., abnormal, test result actually has the target condition and represents the proportion of individuals who fail the screen and are identified as having the target condition based upon corroborating diagnostic testing. A negative predictive value is the probability that a person with a negative, i.e., normal, test result truly does not have the target condition and represents the proportion of individuals who pass the screen and are identified as not having the target condition based upon diagnostic testing.

False Positive and False Negative Rates

Accuracy of screening tests can also be described in terms of false positive and false negative rates. A false positive result occurs when the screen result is positive for an individual who does not have the condition being tested. The false positive rate is calculated as: FP/(TP + FP). The ideal value for a false positive rate is 0. However, it is nearly impossible to achieve this when using a screening test in a large population.

A false negative result occurs when the screen reports a negative result for an individual who actually has the condition being screened. False negative rates are calculated as: FN/(TN = FN).

- When evaluating the usefulness of a screening test the consequences of false positive and false negative rates need to be considered. Due to their inherent purpose screening tests have high sensitivities and low specificities. This allows for detection of most patients with aspiration risk while having the acceptable disadvantage of a high rate of false positives.

Likelihood Ratios

Another means of assessing accuracy of tests is the use of likelihood ratios. Likelihood ratios (LR) are defined as the likelihood that a given test result would be expected in an individual with the target disorder compared to the likelihood that the same result would be expected in an individual without the target disorder. One advantage likelihood ratios offer over sensitivity and specificity is that they are less likely to change due to the prevalence of a given disorder.

Likelihood ratios are expressed in terms of positive and negative likelihood ratios. A positive likelihood ratio indicates how the probability of a disease shifts when the finding is present. In other words, a positive likelihood ratio reflects the number of times it is more likely that a positive test comes from an individual with the disease rather than from an individual without the disease. A negative likelihood ratio indicates how the probability of disease shifts when it is absent. In other words, a negative likelihood ratio is equivalent to the number of times it is more likely that a negative test comes from an individual with the disease rather than from an individual without the disease.

Illustrative Example

To demonstrate how a 2×2 contingency table would be used to determine accuracy of a given screening test, let us suppose a hospital wants to implement a new screening procedure to determine if newly admitted patients are at risk for aspiration. The plan is to allow individuals who pass the screen to begin oral diets and oral medications without further instrumental evaluation. Patients who fail the screen will be referred for further testing. In this scenario, before the hospital can adopt the screen for use a determination must be made regarding the accuracy of the screen. To accomplish this, a pilot study is first performed with 100 individuals admitted with the diagnosis of suspected stroke. All participants in the pilot study complete both the screen followed

immediately by an instrumental assessment of swallowing. The results are as follows:

		Aspiration on instrumental examination	
		Positive	Negative
Screening	Positive	a or True Positives (TP) $n=50$	b or False Positives (FP) $n=25$
	Negative	c or False Negatives (FN) $n=2$	d or True Negatives (TN) $n=23$

Sensitivity = a/(a + c) = 50/(50 + 5) = 96.1 %
Specificity = d/(b + d) = 25/(25 + 23) = 52.1 %
PPV = a/(a + b) = 50/(50 + 25) = 66.7 %
NPV = d/(c + d) = 23/(2 + 23) = 92.0 %
Positive Likelihood Ratio = Sensitivity/(1 − Specificity) = 0.961/1–0.521 = 2.006
Negative Likelihood Ratio = (1 − Sensitivity)/Specificity = (1–0.961)/.521 = 0.075
False Positive Rate = FP/(TP + FP) = 25/(50 + 25) = 0.333
False Negative Rate = FN/(TN + FN) = 2/(23 + 2) = 0.080

In this scenario, the screening test results yield a high sensitivity (96.1 %) and a high negative predictive value (92 %), meaning that most individuals who passed the screen also did not aspirate on instrumental examination. This is a good result. Clinicians, therefore, can be confident that when patients pass the screen they are at very low aspiration risk and safe to begin oral alimentation. Specificity (52.1 %) and positive predictive value (66.7 %) for this screening test are considerably lower, indicating that many individuals who fail the screen (have a positive result) are not in actuality aspirating and will be referred for instrumental assessment unnecessarily. However, as previously discussed, over-referral is inherent in all screening instruments and is considered to be an acceptable limitation.

References

1. Streiner DL. Diagnosing tests: using and misusing diagnostic and screening tests. J Pers Assess. 2003;81:209–29.

2. Hafner G, Neuhuber A, Hirtenfelder S, Schmedler B, Eckel ME. Fiberoptic endoscopic evaluation of swallowing in intensive care unit patients. Eur Arch Otorhinolaryngol. 2008;265:441–6.
3. Heffner JE. Swallowing complications after endotracheal extubation. Chest. 2010;137:509–10.
4. de Larminat V, Montravers P, Dureuil B, Desmonts JM. Alteration in swallowing reflex extubation in intensive care unit patients. Crit Care Med. 1995;3:486–90.
5. Leder SB, Cohn SM, Moller BA. Fiberoptic endoscopic documentation of the high incidence of aspiration following extubation in critically ill trauma patients. Dysphagia. 1998;13:208–12.
6. Ajemian MS, Nirmul GB, Anderson MT, Zirlen DM, Kwasnik EM. Routine fiberoptic endoscopic evaluation of swallowing following prolonged intubation: implications for management. Arch Surg. 2001; 136:434–7.
7. Leder SB. Serial fiberoptic endoscopic swallowing examinations in the management of patients with dysphagia. Arch Phys Med Rehabil. 1998;79:1264–9.
8. Leder SB, Suiter DM, Lisitano HL. Answering orientation questions and following single step verbal commands: effect on aspiration status. Dysphagia. 2009;24:290–5.
9. Solh A, Okada M, Bhat A, Pietrantoni C. Swallowing disorders post orotracheal intubation in the elderly. Intensive Care Med. 2003; 29:1451–5.
10. Leder SB, Suiter DM. An epidemiologic study on aging and dysphagia in the acute care hospitalized population: 2000–2007. Gerontology. 2009;55:714–8.
11. Odderson IR, Keaton JC, McKenna BS. Swallow management in patients on an acute stroke pathway: quality is cost effective. Arch Phys Med Rehabil. 1995;76:1130–3.
12. Doggett DL, Tappe KA, Mitchell MD, Chapell R, Coates V, Turkelson CM. Prevention of pneumonia in elderly stroke patients by systematic diagnosis and treatment of dysphagia: an evidence-based comprehensive analysis of the literature. Dysphagia. 2001; 16:279–95.
13. Hinchey JA, Shephard T, Furie K, Smith D, Wang D, Tonn S. Formal dysphagia screening protocols prevent pneumonia. Stroke. 2005; 36:1972–6.
14. Rosenbek JC, McCullough GH, Wertz RT. Is the information about a test important? Applying the methods of evidence-based medicine to the clinical examination of swallowing. J Commun Disord. 2004;37:437–50.

http://www.springer.com/978-3-319-05112-3

The Yale Swallow Protocol
An Evidence-Based Approach to Decision Making
Leder, S.B.; Suiter, D.M.
2014, XIV, 157 p. 4 illus., Softcover
ISBN: 978-3-319-05112-3