# The Hypothesis Driven Physical Examination

Brian T. Garibaldi, MD[a,*], Andrew P.J. Olson, MD[b]

**KEYWORDS**

• Evidence-based physical diagnosis • Hypothesis-driven physical examination
• Likelihood ratios • Accuracy • Reliability

**KEY POINTS**

- The physical examination remains a vital part of the clinical encounter.
- Many physical examination maneuvers are just as reliable as diagnostic gold standard tests.
- A hypothesis-driven approach to the physical examination emphasizes the performance of specific physical examination maneuvers that are able to alter the likelihood of disease in a given patient.
- The physical examination should be taught to trainees in a context-specific manner as opposed to the traditional head-to-toe approach.
- Likelihood ratios are diagnostic weights that facilitate interpretation of physical examination findings at the bedside.

## INTRODUCTION

For centuries, clinicians have used bedside observation to make diagnostic decisions. Over time, additional modalities have been added to aid in the diagnostic process. Perhaps the greatest example is the introduction of the stethoscope by Laennec in the early nineteenth century.[1] As technology has advanced beyond the stethoscope, diagnosis has moved further away from the bedside in the form of laboratory testing and diagnostic imaging.[2] However, the key to the accurate diagnosis of many conditions still lies in the bedside observations of an astute clinician. In some patients, these observations are the only way to determine the presence or absence of disease (eg, herpes zoster, Parkinson disease, cellulitis, and so forth).[3] In other conditions, additional tests are needed for a definitive diagnosis (eg, myocardial infarction); but the physical examination plays a key role in substantially revising the probability of

Disclosure Statement: The authors have no conflicts of interest to disclose.
[a] Division of Pulmonary and Critical Care, Johns Hopkins University School of Medicine, 1830 East Monument Street, 5th Floor, Baltimore, MD 21205, USA; [b] Division of General Internal Medicine, Department of Medicine, Office of Medical Education, University of Minnesota, 420 Delaware Street Southeast, MMC 284, Minneapolis, MN 55455, USA
* Corresponding author.
*E-mail address:* bgariba1@jhmi.edu

Med Clin N Am ■ (2017) ■–■
https://doi.org/10.1016/j.mcna.2017.12.005 **medical.theclinics.com**

disease in order to effectively guide further evaluation. Even after a diagnosis is made, the physical examination is important in following the disease's trajectory and severity. For example, the presence of an S3 gallop in patients with heart failure predicts mortality and might prompt more aggressive intervention beyond the information found on an echocardiogram.[4]

### *Several Factors Have Led to a Decline in Physical Examination Skills*

Despite its enduring importance, several factors have led to a decline in physical examination skills in recent years.[5–7] In the modern hospital, graduate medical trainees spend as little as 12% of their time in direct contact with patients and their families.[8] This lack of time at the bedside has decreased opportunities for deliberate practice and reduced the number of practitioners who are confident in their ability to teach examination skills.[6,9]

There are also many practitioners and learners who question the value and relevance of the physical examination in the age of technology.[10] Some fail to recognize that many physical examination maneuvers are just as reliable as gold standard technology-based tests. Reliability is commonly measured either through simple agreement or by calculating a kappa score. A kappa score of 0 means that agreement between two observers happens by chance alone. A kappa score of 1 indicates perfect agreement. In general, a kappa score greater than 0.4 is considered reasonable for a diagnostic test.[3,11] Many physical examination maneuvers have kappa scores between 0.4 and 0.75, which indicates intermediate to good reliability (**Table 1**). Many diagnostic gold standards have kappa scores that are in that same range. Technology-based tests are not inherently more reliable than the physical examination.

Another reason that some clinicians, particularly those who trained more recently, hold a nihilistic view about the utility of the physical examination is that the examination is often taught as a list of maneuvers to be performed, regardless of the clinical context. Students learn this head-to-toe approach instead of tailoring their examination to each individual patient.[12] They are then assessed on their ability to perform

**Table 1**
**Interobserver reliability (kappa score) for common physical examination findings and common diagnostic standard tests**

| Physical Finding | Kappa | Diagnostic Standard | Kappa |
|---|---|---|---|
| Liver span >9 cm by percussion | 0.11 | Classification of coronary artery lesions (by catheterization) | 0.33 |
| Delayed carotid upstroke | 0.26 | Pulmonary infiltrate (by chest radiograph) | 0.38 |
| Diminished cardiac dullness | 0.49 | Cardiomegaly (by chest radiograph) | 0.48 |
| Facial palsy (present or absent) | 0.57 | Severity of valvular regurgitation (by echo) | 0.32–0.55 |
| Clubbing (Schamroth sign) | 0.64 | Cirrhosis (by liver biopsy) | 0.59 |
| Systolic hypertension (SBP >160 mm Hg) | 0.75 | Calf DVT (by ultrasound) | 0.69 |
| Tachycardia (pulse >100 bpm) | 0.85 | Diagnosis of narrow complex tachycardia (by ECG) | 0.70 |
| Abdominal jugular test | 0.92 | Interstitial edema (by chest radiograph) | 0.83 |

*Abbreviations:* bpm, beats per minute; DVT, deep venous thrombosis; ECG, electrocardiogram; echo, echocardiogram; SBP, systolic blood pressure.

*Adapted from* McGee S. Reliability of physical findings. In: McGee S, editor. Evidence-based physical diagnosis. 4th edition. Philadelphia: Elsevier; 2018. p. 13; with permission.