Relationship Between Accurate Auscultation of a Clinically Useful Third Heart Sound and Level of Experience

Gregory Marcus, MD; Joshua Vessey, MD; Mark V. Jordan, MD; Michele Huddleston, RN; Barry McKeown, MD; Ivor L. Gerber, MD; Elyse Foster, MD; Kamu Chatterjee, MB; Charles E. McCulloch, PhD; Andrew D. Michaels, MD

Background: Poor performance by physicians-in-training and interobserver variability between physicians have diminished clinicians' confidence in the value of the third heart sound (S3).

Methods: To determine whether auscultation of a clinically useful S3 improves with advancing levels of experience, we performed a prospective, blinded, observational study of 100 patients undergoing left-sided heart catheterization. Patients underwent blinded auscultation by 4 physicians (each from 1 of 4 different levels of experience), phonocardiography, measurement of blood B-type natriuretic peptide levels, echocardiography for measurement of left ventricular ejection fraction, and cardiac catheterization for measurement of left ventricular end-diastolic pressure.

Results: Whereas residents’ and interns’ auscultatory findings demonstrated no significant agreement with phonocardiographic findings, an S3 auscultated by cardiology fellows (κ = 0.29; P = .003) and cardiology attendings (κ = 0.37; P < .001) agreed with phonocardiographic findings. Although the sensitivities of the S3 were low (13%-52%) for identifying patients with abnormal measures of left ventricular function, the specificities were high (85%-95%), with the best test characteristics exhibited by phonocardiography and more experienced physicians. The S3 detected by attendings and fellows was superior in distinguishing an elevated B-type natriuretic peptide level, a depressed left ventricular ejection fraction, or an elevated left ventricular end-diastolic pressure (P = .002-.47 for attendings and .02-.03 for fellows) compared with residents (P = .02-.47) or interns (P = .09-.64).

Conclusions: The S3 auscultated by more experienced physicians demonstrated fair agreement with phonocardiographic findings. Although correlations were superior for phonocardiography, the associations between the S3 and abnormal markers of left ventricular function improved with each level of auscultator experience.

Arch Intern Med. 2006;166:617-622

The third heart sound (S3) is a soft, low-frequency vibration that can be auscultated in early diastole. Its origins have been investigated for more than a century. Extensive studies in animal models and human studies involving older patients have demonstrated that the S3 reflects elevated ventricular diastolic pressure, rapid early diastolic filling, increased ventricular stiffness, volume overload, and an abrupt deceleration of early diastolic filling. In addition, the presence of an S3 seems to be clinically meaningful; it portends increased risk in patients undergoing noncardiac surgery, and it is associated with adverse cardiovascular outcomes in patients with heart failure and acute myocardial infarction.

Identification of this valuable physical finding requires relatively little time, and it is accessible to any physician with a stethoscope. However, studies have demonstrated very poor interobserver agreement among physicians and have suggested that trainees may not receive adequate instruction in auscultation. There is concern that the teaching and art of physical diagnosis may deteriorate as physicians increasingly rely on more sophisticated technology. To our knowledge, no previous study has examined agreement in auscultation of the S3 across groups of individuals representing different training levels or used both an objective correlate of the sound itself (eg, phonocardiography) and objective diagnostic testing of cardiac function.

With the hypothesis that auscultation of the S3 improves with level of experience, we sought to compare auscultative abilities of physicians of multiple training levels with the presence of an S3 detected by computerized heart sound analysis. To assess the clinical relevance of this physical examination finding, we also compared the auscultative and phonocardiographic findings with objective independent measures of left ventricular function.
METHODS

PATIENTS

Adults scheduled to undergo nonemergency left-sided heart catheterization for a clinical indication at the University of California, San Francisco, were eligible for enrollment. Exclusion criteria included younger than 18 years, hypotension (systolic blood pressure <90 mm Hg), vasopressor or inotropic pharmacotherapy, cardiac rhythm other than a sinus or paced atrial rhythm, severe mitral regurgitation or stenosis, constrictive pericarditis, a serum creatinine level of 4.0 mg/dL or greater (≥354 µmol/L), severe pulmonary hypertension (mean pulmonary artery pressure >50 mm Hg), and mechanical ventilation.

STUDY PROCEDURE AND MEASUREMENTS

Before angiography, measurement of left ventricular diastolic pressure was recorded. Blood samples were obtained from the arterial sheath for measurement of the B-type natriuretic peptide (BNP) level using a membrane immunofluorescence assay (Biosite Inc., San Diego, Calif). A BNP level greater than 100 pg/mL was prospectively specified as abnormal. Within 2 hours immediately before or after cardiac catheterization, auscultation, computerized heart sound analysis, and a transthoracic echocardiogram were performed. The auscultators were blinded to any other clinical findings at the time of auscultation.

From a review of the clinical record, the patients' primary cardiac diagnosis and significant comorbidities were recorded, including coronary artery disease (defined as ≥1 coronary artery with ≥70% diameter stenosis), systemic hypertension, cardiomyopathy (left ventricular ejection fraction [LVEF] ≤40%), moderate-to-severe aortic stenosis, moderate-to-severe mitral regurgitation, chronic renal insufficiency (creatinine clearance ≤30 mL/min [≤0.50 mL/s]), hypertrophic obstructive cardiomyopathy, and chronic obstructive pulmonary disease. The protocol was approved by the University of California, San Francisco, Committee on Human Research, and all the patients gave written informed consent before enrollment.

INVASIVE LEFT VENTRICULAR PRESSURE MEASUREMENTS

Patients underwent recording of left ventricular end-diastolic pressure (LVEDP) using a 6-French pigtail catheter and a properly zeroed fluid-filled pressure transducer. Pressure was recorded using a 50–100 mm Hg scale at 30–mm/s paper speed. One of 2 blinded physicians (G.M. and J.V.) measured the post-A wave pressure. A minimum of 5 consecutive cardiac cycles were used to determine the mean LVEDP. An LVEDP greater than 15 mm Hg was prospectively specified as abnormal.

ECHOCARDIOGRAPHY

Transthoracic echocardiographic data were obtained using an experienced echocardiographer (Acuson Sequoia; Siemens, Malvern, Pa, or SONOS 5500; Philips Medical Systems, Andover, Mass). Echocardiographic contrast (Optison; Amersham, Little Chalfont, England) (0.3-0.5 mL injected into a peripheral vein) was administered when required to improve endocardial border detection. Echocardiographic data were stored on magnetooptical disks and were analyzed off-line by a single experienced reader (B.M.) blinded to any clinical or study data. End-diastolic and end-systolic volumes were calculated using the biplane method of discs, and they were then indexed to body surface area. These volumes were used to calculate LVEF. An LVEF less than 50% was prospectively specified as abnormal.

CLINICAL AUSCULTATION

Each study participant was auscultated for a left ventricular S3 by 1 physician from each of 4 groups: board-certified cardiology attendings (n=26), cardiology fellows (n=18), internal medicine residents (n=34), and internal medicine interns (n=48). Auscultation was performed in a quiet room at and around the apex in the supine and left lateral decubitus positions. The auscultators were not permitted to elicit any patient history or to perform other components of the physical examination. An intermittent S3 was considered positive. All the auscultators were blinded to the patients’ conditions and to all study test results.

COMPUTERIZED HEART SOUND PHONOCARDIOGRAPHIC ANALYSIS

A 3-minute audioelectric phonocardiographic tracing (Audicor; Inovise Medical Inc, Portland, Ore) was obtained. Audioelectrocardiographic leads were attached to the V3 and V4 position connected to an electrocardiography machine (Marquette MAC 5000; General Electric Healthcare Technologies, Waukesha, Wis). The audioelectrocardiographic data were stored electronically on a compact disc. A 10-second segment free of artifact was selected offline by a blinded electrical and biomedical engineer for a computer-generated report determining the presence of an S3.

STATISTICAL ANALYSIS

Data are given as mean±SD for continuous variables. Comparisons between groups were assessed using Mann-Whitney and Fisher exact tests, where appropriate. κ Statistics were calculated for the degree of agreement among phonocardiography and the different auscultation groups. Generally, κ>0.7 identifies good agreement, κ=0 is equal to guessing, and κ<0 is worse than guessing. Clustering of ratings by individual physicians was checked using a x² test in a random-effects logistic regression of physicians’ rating on the phonocardiography rating. Two-tailed P<.05 was considered statistically significant. Statistical computations were performed using a software program (Stata version 9.1; StataCorp, College Station, Tex).

RESULTS

One hundred patients were enrolled. The mean age was 62±14 years (range, 24-91 years), and 65 were men. Twenty-nine of these patients had treated diabetes mellitus, 81 had systemic hypertension, 36 had a clinical diagnosis of heart failure, 68 had coronary artery disease, and 17 were hospitalized for an acute coronary syndrome. Seven patients had moderate-to-severe aortic stenosis, and 3 had severe hypertrophic obstructive cardiomyopathy. The mean creatinine level was 1.52±1.39 mg/dL (134±123 µmol/L) (>2.0 mg/dL [>177 µmol/L] in 12 patients).

One hundred patients had invasive measurement of central hemodynamic variables and LVEDP. The mean heart rate was 69±12 bpm. The mean central aortic pressure was 131±25 mm Hg systolic and 66±12 mm Hg diastolic. The mean LVEDP was 15.1±7.7 mm Hg (range, 1-31 mm Hg). Forty-six patients had an abnormal LVEDP (>15 mm Hg). Eighty-eight patients underwent adequate assessment of LVEF. The mean LVEF was 57%±19% (range, 7%-85%). Twenty-six patients had an abnormal LVEF (<50%). Ninety-eight patients underwent BNP measurement. The mean BNP level was 487±914 pg/mL (range, 5-4490 pg/mL). Fifty-eight patients had an abnormal BNP level (>100 pg/mL).
CLINICAL AUSCULTATION AND COMPUTERIZED HEART SOUND ANALYSIS

Ninety-eight patients were auscultated by 1 of 26 board-certified cardiologists. Ninety-nine patients were auscultated by 1 of 18 cardiology fellows. Ninety-four patients were auscultated by 1 of 48 internal medicine interns. The maximum number of patients auscultated by a single physician from each group was 19 for attendings and fellows and 6 for residents and interns. The percentage of study patients in whom an S3 was auscultated is given by level of experience in Table 1. Ninety patients had adequate phonocardiographic data: 2 were excluded because of atrial pacing and 8 owing to poor sound quality tracings. Phonocardiography detected an S3 in 21 (23%) of the 90 patients (Table 1).

INTEROBSERVER VARIABILITY

The fellows’ and attendings’ auscultation of an S3 had fair agreement with phonocardiographic findings ($\kappa = 0.37$; $P < .001$ and $\kappa = 0.29$; $P = .003$, respectively). Using phonocardiographic S3 as the reference standard, fellows tended to have fewer false-positive S3s (30%) compared with attendings (55%) but more false-negative S3s (61% for fellows vs 48% for attendings). The agreement was higher with residents than with interns, but there was no significant agreement between either interns or residents with phonocardiographic findings ($\kappa = 0.04$; $P = .36$ and $\kappa = 0.13$; $P = .11$, respectively). None of the 8 tests for clustering of ratings by individual physician (S3 for each category of physician) was significant ($P > .15$ by $\chi^2$ test for all).

S3 COMPARED WITH MEASUREMENTS OF CARDIAC FUNCTION

Patients with an S3 detected by phonocardiography had significantly higher BNP, lower LVEF, and higher LVEDP measurements than those without an S3 detected by phonocardiography (Figure 1 and Table 2). With decreasing levels of experience, each of the clinical auscultator groups performed progressively less well compared with phonocardiography for separating patients based on these functional and hemodynamic measurements (Table 2). Phonocardiography and attendings separated patients based on all 3 measures of ventricular function. Fellow detection of an S3 separated patients based on LVEF and LVEDP but not BNP level. Resident auscultation of an S3 separated patients based only on LVEDP and not BNP level or LVEF. Intern assessment of an S3 performed most poorly, with no significant separation based on any of these objective measures of cardiac function.

S3 TEST CHARACTERISTICS

Using prospectively defined cutoff points for abnormality in each hemodynamic measurement, the presence of an S3 as detected by phonocardiography had sensitivities of 33% to 52% and specificities of 88% to 92% (Figure 2). Compared with attendings, an S3 detected by fellow auscultation yielded consistently lower sensitivities (33%-50% for attendings compared with 13%-30% for fellows) but higher specificities (85%-89% for attendings compared with 93%-95% for fellows). Auscultation of an S3 by interns and residents demonstrated low sensitivities (30%-38%). The intern’s and residents’ S3 had somewhat high specificities (79%-88%) that generally remained lower than those obtained by the groups with more advanced levels of training. Phonocardiography, attendings, and fellows detected an S3 more frequently in those with elevated LVEDP and reduced LVEF compared with residents and interns (Table 3).

Although the clinical utility and prognostic power of the S3 have been demonstrated, previous literature focusing on the actual auscultation of the sound has not been promising. Physicians-in-training were shown to have inadequate cardiac auscultatory proficiency, and interobserver agreement among small numbers of fully trained physicians has been poor. Lok et al examined the accuracy and interobserver agreement of 2 cardiologists, 1 general internist, 3 senior residents, and 2 junior residents to auscultate an S3 in 40 patients. They found “slight agreement” among auscultators ($\kappa = 0.18$). There was no improvement in accuracy (comparing auscultative findings with those from phonocardiography) or interobserver agreement with more advanced levels of training. Unlike previous studies, we examined the ability of groups of physicians to represent each level of training and included a larger number of patients. A fair degree of agreement between an S3 detected by phonocardiography and one detected by cardiology fellows ($\kappa = 0.37$)

---

Table 1. Auscultator Agreement With Phonocardiography for the Presence of a Third Heart Sound (S3)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patients Assessed, No.</th>
<th>Patients With an S3, No. (%)</th>
<th>$\kappa$ vs Phonocardiography</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonocardiography</td>
<td>90</td>
<td>21 (23)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Attendings</td>
<td>98</td>
<td>23 (23)</td>
<td>0.29</td>
<td>.003</td>
</tr>
<tr>
<td>Fellows</td>
<td>99</td>
<td>12 (12)</td>
<td>0.37</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Residents</td>
<td>94</td>
<td>21 (22)</td>
<td>0.13</td>
<td>.11</td>
</tr>
<tr>
<td>Interns</td>
<td>86</td>
<td>22 (26)</td>
<td>0.04</td>
<td>.36</td>
</tr>
</tbody>
</table>

Abbreviation: NA, not applicable.
or cardiology attendings ($\kappa=0.29$) was observed. There was poor agreement between phonocardiography and medicine residents ($\kappa=-0.29$) and medicine interns ($\kappa=0.04$). In contrast, we demonstrated that the agreement with phonocardiography improved with a greater level of experience. We recently showed that the phonocardiographic S3 has fair sensitivities but high specificities for evidence of left ventricular dysfunction.22 Although several animal and human studies have elucidated the origins of the S3, few have compared the sound levels of B-type natriuretic peptide (ln[BNP]) levels in patients with (+S3) and without (−S3) a third heart sound by observer. The bottom line indicates the 25th percentile; the middle line, the median; and the top line, the 75th percentile. Asterisks represent outliers; error bars, ±1.5 times the interquartile range.

**Table 2. Functional and Hemodynamic Ventricular Measurements Based on the Third Heart Sound (S3) Determination by Observer**

<table>
<thead>
<tr>
<th>Measurement and Observer</th>
<th>Without S3</th>
<th>With S3</th>
<th>$P$ Value</th>
<th>AUROC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNP, median (IQR), pg/mL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonocardiography</td>
<td>101 (49-256)</td>
<td>568 (211-1280)</td>
<td>$&lt;.001$</td>
<td>0.75 (0.64-0.83)</td>
</tr>
<tr>
<td>Attendings</td>
<td>108 (46-317)</td>
<td>389 (147-1210)</td>
<td>.003</td>
<td>0.71 (0.61-0.80)</td>
</tr>
<tr>
<td>Fellows</td>
<td>115 (54-344)</td>
<td>904 (142-1869)</td>
<td>.03</td>
<td>0.71 (0.61-0.80)</td>
</tr>
<tr>
<td>Residents</td>
<td>115 (49-366)</td>
<td>249 (105-445)</td>
<td>.048</td>
<td>0.64 (0.53-0.74)</td>
</tr>
<tr>
<td>Interns</td>
<td>113 (41-394)</td>
<td>198 (101-445)</td>
<td>.09</td>
<td>0.62 (0.51-0.72)</td>
</tr>
<tr>
<td>LVEF, mean ± SD, %</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonocardiography</td>
<td>61 ± 17</td>
<td>45 ± 19</td>
<td>.002</td>
<td>0.74 (0.63-0.83)</td>
</tr>
<tr>
<td>Attendings</td>
<td>61 ± 18</td>
<td>45 ± 18</td>
<td>.002</td>
<td>0.73 (0.63-0.82)</td>
</tr>
<tr>
<td>Fellows</td>
<td>59 ± 18</td>
<td>42 ± 20</td>
<td>.02</td>
<td>0.72 (0.62-0.82)</td>
</tr>
<tr>
<td>Residents</td>
<td>58 ± 18</td>
<td>53 ± 21</td>
<td>.47</td>
<td>0.56 (0.45-0.67)</td>
</tr>
<tr>
<td>Interns</td>
<td>59 ± 17</td>
<td>54 ± 21</td>
<td>.64</td>
<td>0.54 (0.42-0.65)</td>
</tr>
<tr>
<td>LVEDP, mean ± SD, mm Hg</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonocardiography</td>
<td>13.8 ± 7.7</td>
<td>18.8 ± 7.0</td>
<td>.005</td>
<td>0.70 (0.59-0.79)</td>
</tr>
<tr>
<td>Attendings</td>
<td>14.3 ± 7.3</td>
<td>18.5 ± 7.8</td>
<td>.02</td>
<td>0.66 (0.56-0.76)</td>
</tr>
<tr>
<td>Fellows</td>
<td>14.4 ± 7.6</td>
<td>19.8 ± 7.2</td>
<td>.02</td>
<td>0.70 (0.61-0.79)</td>
</tr>
<tr>
<td>Residents</td>
<td>14.0 ± 7.4</td>
<td>18.6 ± 7.6</td>
<td>.02</td>
<td>0.67 (0.57-0.76)</td>
</tr>
<tr>
<td>Interns</td>
<td>14.0 ± 7.1</td>
<td>15.8 ± 7.8</td>
<td>.41</td>
<td>0.56 (0.45-0.66)</td>
</tr>
</tbody>
</table>

Abbreviations: AUROC, area under the receiver operating characteristic curve; BNP, B-type natriuretic peptide; CI, confidence interval; IQR, interquartile range; LVEF, left ventricular ejection fraction; LVEDP, left ventricular end-diastolic pressure.

*The $P$ values from the Mann-Whitney rank sum test and the c-statistic from the AUROC are shown.
with objective markers of hemodynamic aberration routinely used in clinical practice. In 1968, Shah et al demonstrated that the mean left atrial pressure was elevated in 10 patients with an S3 detected by phonocardiography; however, correlation with physician auscultation was not performed.

B-type natriuretic peptide is a neurohormone secreted from the myocardiun in response to myocyte stretch. With higher BNP levels indicating volume overload, this serologic marker has been shown to be useful in the clinical diagnosis of heart failure, and it is becoming a commonly used tool in determining the cause of dyspnea. It was previously demonstrated that an S3 auscultated by a senior cardiologist is highly specific for an elevated level of BNP.

Comparing findings from each group of auscultators and phonocardiography with BNP levels, LVEF, and LVEDP, we found a consistent association between the presence of an S3 and greater hemodynamic abnormality. The presence of an S3 correlated with markers of poor left ventricular function, volume overload, or both. Again, performance improved with more advanced levels of training. Whereas the mean levels of each of the 3 objective markers of cardiac function (BNP, LVEF, and LVEDP) were significantly more abnormal in patients with an S3 detected by phonocardiography and the attendings, 2 markers (LVEF and LVEDP) were significantly more abnormal in those with an S3 detected by fellows, 1 marker (LVEDP) was significantly more abnormal in those with an S3 detected by residents, and none were significantly more abnormal in those with an S3 detected by interns.

No formal assessment of the educational curriculum for physical diagnostic skills for each level of training was performed. The clear improvement in auscultatory accuracy by the fellows compared with the residents and interns may be due in part to the emphasis on the cardiac physical examination and regular bedside teaching by senior cardiologists provided to the cardiology fellows at the University of California, San Francisco. It is also possible that individuals with greater interest in or skill at clinical auscultation may pursue cardiology specialty training.

Finally, the test characteristics of the S3 demonstrated clinical utility and reflected improvement with higher levels of experience. In general, the sensitivity of the S3 to detect an abnormal hemodynamic marker was poor, regardless of the level of the auscultator. Therefore, the absence of an S3 is likely not helpful to exclude abnormal ventricular hemodynamics or function. However, the specificity of the S3 for an abnormal hemodynamic marker was consistently high: the specificities of an S3 detected by interns for an abnormal hemodynamic marker in the abnormal range were 79% to 84%, and the specificities of an S3 detected by residents were 82% to 88%. The presence of an S3 detected by fellows and attendings had specificities of 85% to 95% for detecting a marker of cardiac function in the abnormal range. Because auscultation of an S3 requires that the heart sound have sufficient amplitude, frequency, and coupling to the chest wall, this physical finding is an insensitive, but highly specific, sign of ventricular dysfunction. Therefore, auscultation of an S3 by individuals with advanced training can be useful to rule in an abnormally elevated BNP level, a depressed LVEF, or an elevated LVEDP.

Although correlation with phonocardiographic findings improved with each level of experience, the values for fellows (κ = 0.37) and attendings (κ = 0.29) were only fair. Rather than reflecting only fair auscultation of the S3 by these physicians, this level of agreement must

![Figure 2](http://archinte.jamanetwork.com/pdf-articles/166/1/B783-Fig2-res.jpg)

**Figure 2.** Sensitivity (A) and specificity (B) of the third heart sound for detecting abnormal measurements of cardiac function. BNP indicates B-type natriuretic peptide; LVEDP, left ventricular end-diastolic pressure; and LVEF, left ventricular ejection fraction.

<p>| Table 3. Prevalence of the Third Heart Sound Stratified by Left Ventricular End-Diastolic Pressure (LVEDP) and Left Ventricular Ejection Fraction (LVEF)* |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>LVEDP = 15 mm Hg</th>
<th>LVEDP = &gt;50% (n = 45)</th>
<th>LVEF = &gt;50% (n = 18)</th>
<th>LVEDP = &gt;15 mm Hg</th>
<th>LVEF = &gt;50% (n = 18)</th>
<th>LVEDP = &gt;15 mm Hg</th>
<th>LVEF = &gt;50% (n = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phonocardiography</td>
<td>11</td>
<td>20</td>
<td>22</td>
<td>61</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>Attendings</td>
<td>18</td>
<td>0</td>
<td>11</td>
<td>61</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>Fellows</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>39</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>Residents</td>
<td>14</td>
<td>0</td>
<td>29</td>
<td>50</td>
<td>.13</td>
<td></td>
</tr>
<tr>
<td>Interns</td>
<td>18</td>
<td>25</td>
<td>33</td>
<td>38</td>
<td>.53</td>
<td></td>
</tr>
</tbody>
</table>

*Data are given as percentage of observers.
†By Fisher exact test.
be interpreted with the understanding that the phonocardiographic S3 also had modest sensitivity for detecting patients with abnormal left ventricular function. Using the abnormal levels of the objective markers of ventricular function (BNP, LVEF, and LVEDP) as reference standards, the sensitivity of the phonocardiographic S3 was low (33%-52%), demonstrating that a false-negative phonocardiographic S3 might commonly occur. The specificity of the phonocardiographic S3 was high. Using the same markers as reference standards, the test specificities of the attendings’ and fellows’ auscultated S3 (85%-95%) were similar to those of the phonocardiographic S3 (88%-92%). The attendings’ and fellows’ S3 matched the phonocardiographic S3 in successfully distinguishing patients with abnormal levels of each of the 3 markers (Table 2). The increasing correlation with the phonocardiogram with each level of experience is encouraging; however, the lack of a robust χ value should not dissuade anyone from perceiving the physician-auscultated S3 as a clinically useful finding.

Insofar as these findings demonstrate the capacity for physicians to auscultate a clinically important S3, we believe they can be generalized to the practicing physician and physician-in-training. The full realization of that capacity requires both continuing interest on the part of the learner and mentorship and teaching by those with expertise.

Because we did not follow the auscultative abilities of any given individual across time and with further training, we cannot absolutely conclude that more training is sufficient to improve physical examination skills. However, lack of evidence for interdependence among an individual physician’s ratings reduced the likelihood that the superior performance of the cardiology fellows and attendings was due to a selection of particular physicians with an interest in or affinity for the cardiovascular examination (rather than a general improvement with more training). Although auscultators were blinded to patient histories and test results, they had the opportunity to “eyeball” the patient during auscultation. The appearance of the patient and visible physical examination findings may have biased the auscultator; however, in true clinical practice, the examining physician will always be guided by the general appearance and condition of the patient. Finally, although patients with pulmonary hypertension were excluded, a right ventricular S3 might have been detected.

In conclusion, compared with an objective measure of the sound itself (phonocardiography), auscultation of the S3 improves with more advanced levels of training. Compared with clinically useful objective markers of left ventricular function (BNP, LVEF, and LVEDP), the specificity of the S3 is high, with the best performance demonstrated by physicians representing more advanced levels of training. Phonocardiography performed better than any of the auscultator groups in correlating with measures of ventricular function.

Accepted for Publication: November 11, 2005.

Correspondence: Andrew D. Michaels, MD, Division of Cardiology, University of California, San Francisco, Medical Center, 505 Parnassus Ave, Box 0124, San Francisco, CA 94143-0124 (andrewm@medicine.ucsf.edu).

Financial Disclosure: None.

Funding/Support: This study was supported by an unrestricted educational grant from Inovice Medical Inc (Dr Michaels) and internal research funds from the Division of Cardiology, University of California, San Francisco.

Acknowledgment: We thank the patients who participated in the study; the staff of the Cardiac Catheterization Laboratory, University of California, San Francisco, for their technical assistance; Patti Arand, PhD, Nancy Forman, RN, BSN, and Robert Warner, MD, of Inovise Medical Inc for training and technical assistance; and Inovise Medical Inc for free use of the phonocardiographic equipment and interpretation of the tracings.

REFERENCES