Impact of clinical experience on quantification of clinical signs at physical examination

Luciano F Drager
Paulo A Lotufo, Universidade de São Paulo
Isabela M Bensenor
Impact of clinical experience on quantification of clinical signs at physical examination

From the Department of Medicine, School of Medicine, University of São Paulo, São Paulo, Brazil

Abstract. Drager LF, Abe JM, Martins MA, Lotufo PA, Bensenõr IJM (School of Medicine, University of São Paulo, São Paulo, Brazil). Impact of clinical experience on quantification of clinical signs at physical examination. J Intern Med 2003; 254: 257–263.

Purpose. Although physical examination is a fundamental component of medical decision making, relatively few studies have evaluated how physicians quantify clinical signs and whether different methods of assessment have different effects on clinical practice.

Objectives. To evaluate a possible impact of clinical experience when attending physicians, medical residents and medical students quantify qualitative signs of physical examination in a teaching hospital.

Setting. Hospital das Clínicas, University of São Paulo, Brazil.

Subjects. A total of 244 randomly selected physicians and medical students completed a reliable and consistent eight-item questionnaire.

Main outcome measures. To compare how they quantified clinical signs of cyanosis, anaemia, jaundice, oedema and dehydration, why they used the method(s) they described, and whether the method used could affect diagnosis, further testing or patient management. A chi-square test was used to calculate differences between the groups.

Results. Whilst the majority of those surveyed tended to use a four-level evaluation for these clinical signs, attending staff physicians were more likely to employ two-level evaluations than were residents or medical students. For all five signs, attending physicians’ use of dichotomous evaluations was significantly higher than that of residents or medical students: anaemia (P = 0.004), cyanosis (P < 0.001), oedema (P = 0.005), dehydration (P < 0.001) and jaundice (P = 0.002).

Conclusion. Although medical students and residents are routinely taught to use a four-level evaluation for these clinical signs, many of those surveyed tend to abandon this experience for a dichotomous approach. Given that the clinicians in this survey tended not to change their initial approach to a patient based on the intensity of this semi-quantitative method, increased emphasis on teaching dichotomous approach evaluations in medical school should be encouraged.

Keywords: clinical signs, diagnosis, investigation approach, medical decision, physical examination, quantification of clinical signs.

Introduction

Despite the importance of physical examination in patient care, relatively few studies have critically evaluated this key step of medical diagnosis and decision making. Such assessments are sorely needed, given that many methods used in physical examination were developed generations ago, few have been subjected to standardization or careful evaluation, and the reproducibility of the evaluation of clinical signs between different observers is problematic. Although extreme clinical findings have great interobserver reliability, clinicians often disagree when it comes to more subtle or intermediate signs [1, 2]. The quantification of clinical signs such as cyanosis, anaemia, jaundice, oedema and dehydration are examples of this problem. Different evaluations of these clinical signs are common, and this divergence affects the global assessment of patients [1–5].

Statements such as these often appear in medical journals: ‘A new grade 2 systolic murmur was
heard. The lungs were clear. There was ++ peripheral oedema’ [6]; or ‘There was a moderate amount of ascites’ [7] or ‘There was slight oedema in both legs’ [8]. Whilst such statements suggest a level of confidence and accuracy in the results of a physical examination, the reality is quite different.

In a study evaluating the diagnosis of anaemia, Wurapa et al. found that conjunctival pallor, a commonly used sign for anaemia, had a sensitivity of 18.6% and a specificity of 95.8% when compared with actual haemoglobin levels [9]. Gjørup et al. studied interobserver agreements between three different physicians evaluating anaemia based on inspection of skin, nail beds and conjunctiva. Using haemoglobin levels as a standard, the sensitivity of these evaluations ranged from 27 to 44% and the specificity ranged from 88 to 95% [3]. Interestingly, compared with the haemoglobin-based diagnosis, the physicians tended not to diagnose patients as anaemic; even amongst those with the lowest haemoglobin concentrations, the expected probability of detecting anaemia did not exceed 75%. A double-blind study involving 62 medical observers at various levels of training showed that only 58% detected scleral icterus in six patients at a total serum bilirubin concentration of 2.5 mg dL\(^{-1}\), and 68% detected it at a concentration of 3.1 mg dL\(^{-1}\). The level of training did not appear to influence sensitivity [10].

The availability of laboratory or technical analyses is no guarantee of accuracy. The results from transcutaneous devices used to assess jaundice by estimating bilirubin levels, for example, have been shown to be quite variable when compared with a combination of clinical evaluation and serum bilirubin levels [11, 12]. Moreover, the agreement between the cut-off points of some laboratory tests and clinical signs may be inappropriate, as is the case for central cyanosis. In 1923, Lundsgaard and Van Slyke defined clinical cyanosis as haemoglobin levels <5 g dL\(^{-1}\) of capillary blood [13]. These definitions have been disputed [14–17], although some of the criticisms emerged from analyses of reduced haemoglobin in the arterial, not capillary, blood. A more recent revision supporting the original values described by Lundsgaard and Van Slyke also discussed the problems of defining clinical cut-off points [18]. For clinical signs that cannot be further evaluated with a laboratory test, such as oedema, quantification must be based on physical examination alone.

In this study, we assessed how a sample of medical students, medical residents and attending physicians in a teaching hospital quantify qualitative signs of the physical examination such as cyanosis, anaemia, jaundice, oedema and dehydration.

### Methods

A random sample of attending physicians, medical residents and fifth-year medical students working in the Hospital das Clínicas of the University of São Paulo and the Faculty of Medicine of the University of São Paulo were asked to complete an eight-item questionnaire regarding the quantification of cyanosis, anaemia, jaundice, oedema and dehydration during physical examination. The reliability of the questionnaire was evaluated using Cronbach alpha.

This study was approved by the Institutional Review Board and informed written consent of all participants was obtained prior to the study.

The questionnaire asked which of the following methods were used to quantify clinical signs: a four-level evaluation (crosses method –+/++/++++), a three-level evaluation (mild, moderate or severe); a dichotomous evaluation (present or absent) or other (Appendix 1). It also asked participants to explain why they chose to use a particular method, whether they felt they could precisely discriminate between the different levels of clinical signs at physical examination, and whether their initial approach to patients with cyanosis, anaemia, jaundice, oedema, and dehydration changed according to the method used (four-level, three-level and dichotomous evaluation) to quantify clinical signs at physical examination.

We examined whether length of training, number of years since graduating from medical school (five or less and more than 5 years since graduation), or medical specialty influenced the methods used to quantify clinical signs. We also analysed whether the methods used changed with time since graduation to clinical practice.

For three- and four-level evaluations, we analysed whether the intensity of the clinical sign (mild, moderate or severe, or 1–4 crosses) changed the physician’s approach in specific cases.

We used a chi-square approach to evaluate interobserver variability by medical specialty and...
by years of clinical practice (student, resident and assistant physician).

**Results**

A total of 81 attending physicians, 81 medical residents and 82 fifth-year medical students completed the questionnaire twice, with a 1-month interval. Amongst the physicians and residents, medical specialties included Internal Medicine, General Surgery, Paediatrics, and Obstetrics/Gynaecology (Table 1). Cronbach alpha was 0.78, which is considered an acceptable value for reliability of the scale.

The majority of attending physicians, residents and medical students used a four-level evaluation to quantify clinical signs of anaemia, cyanosis, oedema and jaundice at physical examination (Table 2). When evaluating dehydration, however, attending physicians tended to use a three-level observation (mild, moderate or severe), whilst residents and medical students used a four-level evaluation.

The use of a four-level evaluation for all five signs was inversely associated with clinical experience, with attending physicians using it less often than residents and residents using it less often than medical students. Attending physicians were more likely to use a three-level or dichotomous evaluation compared with residents and medical students for cyanosis ($P < 0.001$), anaemia ($P = 0.004$), jaundice ($P = 0.002$), oedema ($P = 0.005$) and dehydration ($P < 0.001$). About 25% of paediatricians classified jaundice using Krammer’s zones, a specific method of evaluation used primarily by paediatricians devoted to neonatology.

---

**Table 1** Distribution of survey participants by clinical experience and medical specialty

<table>
<thead>
<tr>
<th>Medical specialty</th>
<th>Number (%)</th>
<th>Attending physicians</th>
<th>Residents</th>
<th>Medical students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal medicine</td>
<td>42 (17.2)</td>
<td>20</td>
<td>22</td>
<td>–</td>
</tr>
<tr>
<td>General surgery</td>
<td>40 (16.4)</td>
<td>21</td>
<td>19</td>
<td>–</td>
</tr>
<tr>
<td>Paediatrics</td>
<td>40 (16.4)</td>
<td>20</td>
<td>20</td>
<td>–</td>
</tr>
<tr>
<td>Obstetrics/gynaecology</td>
<td>40 (16.4)</td>
<td>20</td>
<td>20</td>
<td>–</td>
</tr>
<tr>
<td>Medical students</td>
<td>82 (33.6)</td>
<td>–</td>
<td>–</td>
<td>82</td>
</tr>
<tr>
<td>Total (%)</td>
<td>244 (100)</td>
<td>81 (33.2)*</td>
<td>81 (33.2)*</td>
<td>82 (33.6)*</td>
</tr>
</tbody>
</table>

*P*-value not significant.

**Table 2** Method chosen for evaluating clinical signs by attending physicians, residents and medical students

<table>
<thead>
<tr>
<th>Methods</th>
<th>Attending physicians (%)</th>
<th>Residents (%)</th>
<th>Medical students (%)</th>
<th>$\chi^2$ ($P$)</th>
<th>$P$ for trend</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cyanosis (n = 244)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-level</td>
<td>38 (46.9)</td>
<td>60 (74.1)</td>
<td>63 (76.8)</td>
<td>0.00061</td>
<td></td>
</tr>
<tr>
<td>3-level</td>
<td>20 (24.7)</td>
<td>6 (7.4)</td>
<td>7 (8.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dichotomous</td>
<td>23 (28.4)</td>
<td>15 (18.5)</td>
<td>12 (14.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Anaemia (n = 242)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-level</td>
<td>60 (74.1)</td>
<td>67 (82.7)</td>
<td>76 (95)</td>
<td>0.00413</td>
<td></td>
</tr>
<tr>
<td>3-level</td>
<td>17 (21)</td>
<td>8 (9.9)</td>
<td>2 (2.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dichotomous</td>
<td>4 (4.9)</td>
<td>6 (7.4)</td>
<td>2 (2.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Jaundice (n = 244)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-level</td>
<td>57 (70.4)</td>
<td>74 (91.4)</td>
<td>78 (95.1)</td>
<td>0.00199</td>
<td></td>
</tr>
<tr>
<td>3-level</td>
<td>13 (16.1)</td>
<td>3 (3.7)</td>
<td>0 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dichotomous</td>
<td>4 (4.9)</td>
<td>1 (1.2)</td>
<td>0 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krammer’s zones</td>
<td>7 (8.6)</td>
<td>3 (3.7)</td>
<td>4 (4.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Oedema (n = 243)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-level</td>
<td>58 (72.5)</td>
<td>75 (92.6)</td>
<td>76 (92.7)</td>
<td>0.00493</td>
<td></td>
</tr>
<tr>
<td>3-level</td>
<td>16 (20)</td>
<td>3 (3.7)</td>
<td>2 (2.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dichotomous</td>
<td>6 (7.5)</td>
<td>3 (3.7)</td>
<td>4 (4.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dehydration (n = 241)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-level</td>
<td>21 (26.6)</td>
<td>39 (48.1)</td>
<td>51 (63)</td>
<td>0.00008</td>
<td></td>
</tr>
<tr>
<td>3-level</td>
<td>40 (50.6)</td>
<td>32 (39.5)</td>
<td>19 (23.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dichotomous</td>
<td>18 (22.8)</td>
<td>10 (12.3)</td>
<td>11 (13.6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2003 Blackwell Publishing Ltd *Journal of Internal Medicine* **254**: 257–263
The seven attending physicians with 5 years or less of clinical experience tended to use four-level evaluations to quantify clinical signs, as did residents and medical students. This was particularly true for anaemia (six of seven assistants with <5 years since graduation), jaundice (six of seven assistants), and oedema (seven of seven assistants).

In contrast, attending physicians with more than 5 years of clinical experience were more likely to use dichotomous or three-level evaluations, especially for cyanosis ($P = 0.007$), jaundice ($P = 0.009$), oedema ($P < 0.001$) and dehydration ($P = 0.006$).

In the evaluation of dehydration, the three-level evaluation is used more than the four-level, a trend observed even amongst residents and medical students.

A small percentage of those interviewed (2.5%) used other methods to quantify clinical signs. Training in a particular medical specialty did not appear to influence the method used to quantify clinical signs. Similar methods were employed by different specialists for all analysed signs ($P$ not significant), with the exception of paediatricians’ high use of Krammer’s zones for evaluating jaundice.

When we asked why clinicians used a specific method for quantifying a clinical sign, the most common reason given for use of a four-level method was having learned it in medical school – 66.6% of attending physicians, 91.4% of residents and 90.2% of medical students gave this response.

In response to a question asking whether the quantification of cyanosis, anaemia, jaundice, oedema and dehydration would change a physician’s initial patient-care plan, about two-thirds of the physicians and medical students using the four-level method said no as well as a similar proportion of those using a three-level evaluation. Data are not significant for any of the signs examined in this study. Dichotomous evaluation was not included in this analysis because the presence or absence of a sign clearly changed the initial approach to management.

Discussion

In this assessment of the quantification of clinical signs amongst a sample of 244 attending physicians, residents and medical students, young clinicians or those in training tended to rely on four-level evaluations whilst more experienced clinicians tended to use a three-level or dichotomous evaluation, largely to simplify the possibilities.

The fact that most of the attending physicians, residents and medical students use a four-level method and justify this approach by having learned it in medical school comes as no surprise, as medical education has an enormous and long-lasting impact on clinical practice. What is surprising is that the method used by a substantial proportion of clinicians for these five signs, the dichotomous evaluation, is not one taught in most medical schools.

Whilst medical students and clinicians know that a detailed medical history and a careful physical examination form the foundation for evaluation of a patient, studies show that the contribution of physical examination to clinical diagnosis is poor. In a study of 630 patients followed for 2 years, Sandler et al. found that physical examination alone contributed to only one quarter of the diagnoses, whilst medical history alone contributed to two-thirds of the patients’ diagnoses. Together, medical history and physical examination were the basis for the correct diagnosis in 91% of cases [19]. Physical examination was most helpful for the diagnosis of respiratory diseases such as chronic bronchitis and emphysema, less helpful for cardiovascular diagnoses, and least helpful for alimentary diagnoses. For the 160 patients with chest pain, history alone was sufficient for diagnosis in 90% of cases whilst physical examination alone did not account for any correct diagnoses; routine tests such as chest X-rays and electrocardiograph helped with only 3% of diagnoses. In a study of 80 new outpatients, Hampton et al. observed that physical examination alone provided enough information for making an initial diagnosis in 9% of patients compared with 83% for medical history [20]. In evaluating the usefulness of physical examination for the detection of anaemia, Nardone et al. reported that it had a sensitivity of 65% and a specificity of 95%. Even so, the absence of pallor did not rule out a diagnosis of anaemia [21].

At the same time, physical findings can provide important prognostic information. Recently, Drazner et al. performed a retrospective analysis of the Studies of Left Ventricular Dysfunction Treatment Trial (SOLVD), showing in multivariate analyses that elevated jugular venous pressure and a third heart sound were independently associated with adverse outcomes in heart failure, such as an increased risk of hospitalization and death [22].
It is difficult to determine at that time some methods of quantification for qualitative signs were created. Textbooks do not discuss the origin of these techniques in much detail [23–26]. Another problem is that signs are quantified by different scales. Physicians commonly use a six-point scale for evaluating the intensity of a heart murmur [27] or quantifying muscle strength [28], a five-point scale for evaluating reflexes [29], and a four-point scale for evaluating the pulse [23]. Compared with a simple dichotomous evaluation, evaluations with three or more levels are often confusing and less accurate, leading to difficulty with standardization, increasing subjectivity of clinical judgement, and possibly more errors in diagnosis [1]. Moreover, the reliability of clinicians’ observations or judgements is not to be confused with accuracy [29] – two physicians may be in complete agreement regarding their findings on physical examination yet may be wrong when compared with an independent gold standard. Such conflicts led Fletcher to urge physicians to abandon unreliable methods and stop teaching unreliable signs to their students [30].

Other than the studies by Hampton et al. [20] and Sandler et al. [19] regarding the impact of physical examination on diagnosis, relatively little attention has been paid to whether quantification is useful in changing a clinician’s initial approach or decision making, or its impact on further testing and thus the economics of health care.

In this study, we evaluated whether physicians and physicians-in-training change their approach to a patient based on the quantification of clinical signs, justifying the usefulness of quantifying signs at clinical examination. Surprisingly, the majority of those quantifying cyanosis, anaemia, jaundice, oedema and dehydration using the three- and four-level techniques said they would not change their initial approach independent of quantification. In other words, the initial management of two patients with different quantifications for anaemia (for example, + and ++++) would be the same. This brings into question the usefulness of quantifications involving several variables, as in the case of three- or four-level evaluations.

We did not undertake this study to establish rules for quantifying the five clinical signs examined here. Instead, we wanted to determine how experienced physicians quantify these common clinical signs, to see whether physicians-in-training did it the same way, and to determine how physicians are trained to quantify these signs. We found that the quantification methods commonly taught in medical schools appear to be gradually abandoned with experience in favour of the simpler and more reliable dichotomous evaluation. The development and improvement of simple techniques that can be incorporated in the physical examination that are more accurate and less subject to personal interpretations, such as replacing four-level evaluations when appropriate with dichotomous evaluations, and increased emphasis on these evidence-based techniques in medical school and in continued medical education for attending physicians, may contribute to improved and less costly health care. However, new studies need to be carried out comparing the various methods (dichotomous versus three- to four-level evaluation) used on quantification of clinical signs at physical examinations with gold standard tests as serum levels of haemoglobin or bilirubin.

Conclusions

Although medical students and residents are routinely taught to use a four-level evaluation for clinical signs such as cyanosis, anaemia, jaundice, oedema and dehydration, the majority of those participating in this survey said they would not change their initial approach to treatment based on the intensity assigned to a particular sign. Indeed, most tend to forget this approach as they gain experience in favour of a dichotomous approach. Altogether this may be due to the difficulty of distinguishing between mild alterations, and that concluding that a sign is present or absent is what is most important for determining or changing the initial therapeutic approach.

Conflict of interest statement

No conflict of interest was declared.

Acknowledgements

The authors would like to acknowledge P. J. Skerrett for his help in preparing the manuscript.

Drs Martins, Lotufo and Benseñor are recipients of fellowship from Conselho Nacional de Pesquisa (CNPq), Brasília, Brazil.

© 2003 Blackwell Publishing Ltd Journal of Internal Medicine 254: 257–263
References

21 Nardone DA, Roth KM, Mazur DJ, McAfee JH. Usefulness of physical examination in detecting the presence or absence of anemia. Arch Intern Med 1990; 150: 201–4.
27 Levine AS. Notes on graduation of the intensity of cardiac murmurs. JAMA 1961; 177: 261.

Appendix 1: Questionnaire about the medical evaluation of cyanosis, jaundice, oedema, dehydration and anaemia

(1) Level of performance:
  - Attending physician
  - Medical resident
  - Medical student
  - Other (specify)

(2) Year of graduation from medical school:

(3) Medical specialty:
  - Internal medicine
  - General surgery
  - Paediatrics
  - Gynaecology/obstetrics
  - Other (specify)

(4) What method do you use to quantify the following clinical signs at physical examination?

Anæmia:
  - Crosses method [four-level evaluation (+/4+, 2+/4+, 3+/4+ or 4+/4+)]
  - Light, moderate and serious (three-level evaluation)
  - Presence or absence (dichotomous evaluation)
  - Other (specify)

Cyanosis:
  - Crosses method [four-level evaluation (+/4+, 2+/4+, 3+/4+ or 4+/4+)]
  - Light, moderate or severe (three-level evaluation)
Presence or absence (dichotomous evaluation)
Other (specify)

Dehydration:
☐ Crosses method [four-level evaluation (+/4+, 2+/4+, 3+/4+ or 4+/4+)]
☐ Light, moderate and serious (three-level evaluation)
☐ Presence or absence (dichotomous evaluation)
☐ Other (specify)

Oedema:
☐ Crosses method [four-level evaluation (+/4+, 2+/4+, 3+/4+ or 4+/4+)]
☐ Light, moderate or severe (three-level evaluation)
☐ Presence or absence (dichotomous evaluation)
☐ Other (specify)

Jaundice:
☐ Crosses method [four-level evaluation (+/4+, 2+/4+, 3+/4+ or 4+/4+)]
☐ Light, moderate or severe (three-level evaluation)
☐ Presence or absence (dichotomous evaluation)
☐ Other (specify)

(5) Why?
☐ I learned this method in medical school
☐ I read it in a textbook
☐ It is the easiest for quantifying this clinical sign with high interobserver reproducibility
☐ Other (specify)

(6) Do you think that doctors can precisely discriminate the levels of clinical signs at physical examination?
☐ Yes
☐ No

(7) Do you believe that the initial approach of patients with anaemia, cyanosis, oedema, dehydration or jaundice will change according to the method used (four-level, three-level, dichotomous evaluation) to quantify clinical signs at physical examination?

Cyanosis ☐ Yes ☐ No
Jaundice ☐ Yes ☐ No
Oedema ☐ Yes ☐ No
Dehydration ☐ Yes ☐ No
Anaemia ☐ Yes ☐ No ☐ No
☐ I don’t use three- or four-level evaluation

(8) Would you be interested in participating in further research regarding the quantification of these signs?
☐ Yes ☐ No

Received 5 September 2002; revision received 4 April 2003; accepted 8 April 2003.

Correspondence: Dr Isabela M. Bensenor, Hospital das Clínicas da FMUSP-PAMB, Av Enéas Carvalho de Aguiar, 155, 8o andar bloco 3-Clinica Médica, São Paulo, SP, Brazil CEP: 05403-000 (fax: 55-11-30830827; e-mail: isabensenor@hcnet.usp.br).