

RESEARCH ARTICLE

Physical examination skills training: Faculty staff vs. patient instructor feedback—A controlled trial

Markus Krautter¹, Katja Diefenbacher², Jobst-Hendrik Schultz², Imad Maatouk², Anne Herrmann-Werner³, Nadja Koehl-Hackert⁴, Wolfgang Herzog², Christoph Nikendei^{2*}

1 Department of Nephrology, University of Heidelberg, Heidelberg, Germany, **2** Department of General Internal Medicine and Psychosomatics, University of Heidelberg, Medical Centre, Heidelberg, Germany, **3** Department of Psychosomatic Medicine and Psychotherapy, University of Tübingen, Tübingen, Germany, **4** Department of General Practice and Health Services Research, University of Heidelberg, Heidelberg, Germany

* christoph.nikendei@med.uni-heidelberg.de



Abstract

Background

Standardized patients are widely used in training of medical students, both in teaching and assessment. They also frequently lead complete training sessions delivering physical examination skills without the aid of faculty teaching staff—acting as “patient instructors” (PIs). An important part of this training is their ability to provide detailed structured feedback to students which has a strong impact on their learning success. Yet, to date no study has assessed the quality of physical examination related feedback by PIs. Therefore, we conducted a randomized controlled study comparing feedback of PIs and faculty staff following a physical examination assessed by students and video assessors.

Methods

14 PIs and 14 different faculty staff physicians both delivered feedback to 40 medical students that had performed a physical examination on the respective PI while the physicians observed the performance. The physical examination was rated by two independent video assessors to provide an objective performance standard (gold standard). Feedback of PI and physicians was content analyzed by two different independent video assessors based on a provided checklist and compared to the performance standard. Feedback of PIs and physicians was also rated by medical students and video assessors using a questionnaire consisting of 12 items.

Results

There was no statistical significant difference concerning overall matching of physician or PI feedback with gold standard ratings by video assessment ($p = .219$). There was also no statistical difference when focusing only on items that were classified as major key steps ($p = .802$), mistakes or parts that were left out during physical examination

OPEN ACCESS

Citation: Krautter M, Diefenbacher K, Schultz J-H, Maatouk I, Herrmann-Werner A, Koehl-Hackert N, et al. (2017) Physical examination skills training: Faculty staff vs. patient instructor feedback—A controlled trial. PLoS ONE 12(7): e0180308. <https://doi.org/10.1371/journal.pone.0180308>

Editor: Joerg Latus, Robert Bosch Krankenhaus, GERMANY

Received: February 21, 2017

Accepted: June 13, 2017

Published: July 10, 2017

Copyright: © 2017 Krautter et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

($p = .219$) or mistakes in communication items ($p = .517$). The feedback of physicians was significantly better rated than PI feedback both by students ($p = .043$) as well as by video assessors ($p = .034$).

Conclusions

In summary, our study demonstrates that trained PIs are able to provide feedback of equal quantitative value to that of faculty staff physicians with regard to a physical examination performed on them. However, both the students and the video raters judged the quality of the feedback given by the physicians to be significantly better than that of the PIs.

Introduction

The physical examination of a patient is an essential clinical competence of physicians, and along with comprehensive history taking, flags the beginning of primary patient-doctor relationships. Moreover, history taking and physical examination form the basis for establishing a diagnosis, planning further diagnostic steps, and developing a therapeutic scheme for the patient's care. Accordingly, the acquisition of physical examination skills constitutes a centerpiece of medical education [1]. However, several studies indicate severe shortcomings in students' physical examination competencies [2,3]. In a recent study, we revealed significant deficits in the ability of final-year medical students to perform a detailed physical examination of standardized patients [4], with only 63% of correctly performed procedural steps. However, the manner in which physical examination skills should be delivered is still subject to discussion [5–7].

Standardized Patients (SPs) are specially trained laypersons who present learned symptoms or diseases in a standardized, non-varying manner for didactic purposes. The assignment of SPs, both in teaching and assessment, has a long tradition in medical education [8–14]. SPs have also been used successfully for the teaching of physical examinations [15–19] – mostly by assisting a faculty staff trainer – and for the assessment of examination skills via objective structured clinical examinations (OSCE) [15,20]. In some studies, SPs have even led a complete training session, sometimes being termed “patient instructors” (PIs), who deliver physical examination skills without the aid of faculty teaching staff [15,21].

A major didactic element in medical education in general, and in the assignment of SPs in particular, is seen in the delivery of professional, structured feedback, which has been shown to exert an enduring effect when training medical students and physicians [22–27]. SPs undergo extensive feedback training prior to their deployment in student teaching [28–31]. In terms of the delivery and feedback related to physical examination skills, PIs have to take into account not only the quality of communication, but also the correctness of medical procedures, often without having a medical professional background. To the best of our knowledge, no study to date has assessed the quality of physical examination-related feedback by PIs. Specifically, randomized controlled studies comparing PI feedback to faculty staff feedback following a physical examination are completely lacking.

The present study therefore aimed to evaluate the following hypotheses: Compared to physician feedback, PI feedback is not inferior in terms of a) quantitative measures, in terms of the number of named feedback items observed by objective video-assessor ratings and b) qualitative measures, reflected in questionnaire ratings both by students and by video assessors.

Material and methods

Trial design

The study prospectively investigated the quality of feedback given by PIs vs. faculty staff physicians—in terms of congruence with gold-standard ratings and questionnaire ratings. For this purpose, we created physical examination small-group teaching settings, each comprising one PI, one faculty staff member and one medical student. In total, fourteen different PIs and 14 different faculty staff physicians delivered feedback to 40 medical students (two or three medical students per PI and faculty staff member). First, the medical student performed a physical examination of the respective PI while the physician observed his/her performance. Second, the student received feedback from both the PI and the physician, independently from each other. The feedback of the PI and the physician was then content-analyzed by two different independent video assessors based on a provided checklist. Finally, the results were compared to an objective performance standard as previously described elsewhere [4,32,33].

Participants

Patient instructors sample. Patient instructors ($n = 14$; 4 female, 9 male; mean age 41.1 ± 15.3 years) all had considerable experience in communication skills training and feedback prior to the physical examination training (mean time of serving as an SP 3.4 ± 2.7 years; mean number of roles 10.3 ± 13.3). PIs were informed that the purpose of the study was to evaluate the final-year students' physical examination skills and were otherwise blinded to the study design.

Physicians sample. Physicians ($n = 14$; 4 female, 9 male; mean age 33.6 ± 6.0 years) were experienced internal medicine residents, who had substantial experience in teaching and supervising physical examination (mean work experience in internal medicine 2.9 ± 2.2 years, mean teaching experience 4 ± 4.0 years). In line with the instruction of the PIs, physicians were also told that the purpose of the study was to evaluate the final-year students' physical examination skills and were otherwise blinded to the study design.

Student sample. A total of 40 (25 female, 15 male; mean age 24.8 ± 1.4 years) final-year students agreed to participate in the study. The students were not told that the study aimed to directly compare feedback skills, and were instead informed that the purpose of the study was to investigate their physical examination skills and that they would receive feedback from both the SP and the physician.

Acquisition of data

The trial was conducted over a three-week period at the beginning of the winter semester at the University of Heidelberg, Germany. Data were collected on the premises of the Department of Internal and Psychosomatic Medicine at Heidelberg University Hospital.

Ethics

Ethical approval was granted by the ethics committee of the University of Heidelberg (Nr. S-009/2015). Written consent was obtained from all participants. Study participation was voluntary and all candidates were assured of anonymity and confidentiality.

Patient instructor physical examination training

All PIs underwent two four-hour-long physical examination training sessions. The training began with instruction on basic anatomical and physiological features of the cardiovascular system, the lungs, the abdomen and the thyroid gland, followed by instruction on physical

examination skills. By the end of the training, the PIs were therefore themselves able to perform a physical examination of the respective organ systems. Additionally, the training also focused on the correct recognition of physical examination skills (e.g. the correct placement of the stethoscope during auscultation or the correct depth of palpation during abdominal examination). Finally, PIs were trained in giving specific feedback on physical examination skills, including accompanying communication skills.

Physical examination feedback session. All 14 PIs were randomly assigned to one of the 14 physicians, and the final-year medical students were randomly assigned to these 14 PI-physician dyads. Twelve dyads performed three sessions, while two dyads performed two sessions, resulting in 40 sessions with 40 student participants overall. The students were given role-playing instructions asking them to perform a pre-employment medical check-up of a PI, including a detailed physical examination of the cardiovascular system, the lungs, the abdomen and the thyroid gland [4]. In line with normal teaching situation procedure, during the examination of the respective PI, the physician was able to make notes on mistakes made by the student on a provided checklist. To avoid interrupting the procedure, the PI had to remember any noticeable problems but was able to make notes afterwards. Following the examination, both the PI and the physician gave feedback to the student separately while the other one left the room (Fig 1). The PI gave feedback first in half of the cases, and the physician gave feedback first in the other half. To enable a subsequent comparison of the feedback, PIs and physicians were asked to give complete feedback, including all mistakes or omitted parts of the physical examination as well as things that were performed well.

Comparative assessment of standardized patient and physician feedback

First, medical students' physical examination performances were video-recorded and subsequently rated by two independent video assessors using binary checklists [34] ("item performed correctly", "item not performed/not performed correctly") in line with the University-wide physical examination standards. This enabled the development of a gold-standard rating against which PI and medical staff feedback could be compared [4,32]. To obtain a more

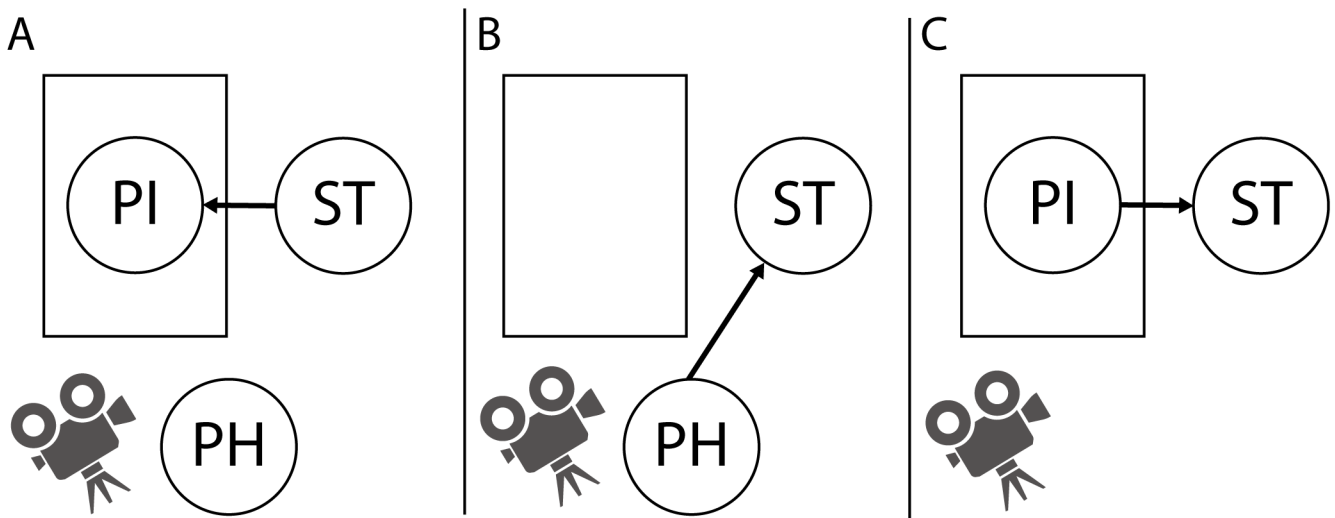


Fig 1. A) Student (ST) performs a physical examination of a patient instructor (PI), while the physician (PH) watches. Student then receives feedback from either physician (B) or PI first (C) while the other one leaves the room.

<https://doi.org/10.1371/journal.pone.0180308.g001>

differentiated pattern, items were categorized into “major” procedural key steps, which are indispensable for a high standard of physical examination, and “minor” procedural steps, which contribute to a more detailed physical examination. Simultaneously, feedback of PIs and physicians was content-analyzed by two different independent video assessors based on the same checklist (“item mentioned was performed correctly”, “item mentioned was not performed or performed incorrectly”, “item was not mentioned”). The resulting ratings were then compared to the gold standard.

Feedback of PIs and physicians was also rated by medical students using a questionnaire consisting of 12 items referring to content (7 items), communication (2 items), and quality (3 items) of the provided feedback. The same questionnaire was used by independent video raters to assess quality of feedback.

Video raters. The video raters (one female, one male, 35.0 ±1.0 years) were experienced internal medicine residents who had experience in training and rating students’ physical examination skills.

Statistical analysis

Data are presented as means ± standard deviation (SD). Non-parametric Mann-Whitney U tests were used for ordinal data. Distribution of group characteristics was compared by Chi-square tests. All data are presented as absolute numbers and percentages or mean and standard deviation. To calculate differences between objective video ratings, the rating scores for PI and physician feedback were aggregated for all of the students assessed by a single teacher or PI (nested design). Ratings from the two video raters were combined into one single rating. A p-value <0.05 was considered to be statistically significant. Standardized inter-rater reliability for the two video assessors was calculated based on intraclass correlation coefficient (ICC) type C. Raw data were processed using Microsoft EXCEL. The software package STATISTICA (Statsoft, Inc, Tulsa, OK, USA) was used for statistical analysis.

Results

Objective quantitative comparison of PI and physician feedback by video raters

There was no statistically significant difference concerning overall matching of physician or PI feedback with gold-standard ratings by video assessment (p = .219). There was also no statistically significant difference when focusing only on items that were classified as major key steps (p = .802), mistakes or parts that were left out during physical examination (p = .219) or mistakes in communication items (p = .517; see [Table 1](#)).

Table 1. Objective comparison of physician and PI feedback with gold standard. Values are percentage of matching with gold-standard checklist ratings.

	Physician (%)	% items	ICC ⁵	PI (%)	% items	ICC	p ⁶
Overall matching of feedback with gold standard	86.79	26.33 ¹	0.697	89.83	22.19 ¹	0.846	.219
Matching with gold standard concerning key steps by students	75.95	34.20 ²	0.684	76.73	26.79 ²	0.908	.802
Matching with gold standard concerning mistakes by students	86.79	27.00 ³	0.697	89.83	25.31 ³	0.846	.219
Matching with gold standard concerning communication by students	71.48	15.45 ⁴	0.493	73.64	11.48 ³	0.058	.517

¹Percentage of items that received feedback based on all 147 items.

²Percentage of items that received feedback based on 53 key step items.

³Percentage of items that received feedback based on students’ mistakes (individual number of items for every student).

⁴Percentage of items that received feedback based on 61 communication items.

⁵Intraclass correlation coefficient (ICC) type C.

⁶t-test for independent samples

Table 2. Qualitative comparison of PI (n = 14) and physician (n = 14) feedback by video assessors.

Items	PI (n = 14)		Physician (n = 14)		p-value
	Mean	(SD)	Mean	SD	
...gave detailed feedback on the execution of the physical examination (PE).	3.71	(1.25)	2.57	(1.34)	<0.001
...gave detailed feedback on the compliance with framework conditions and on the sequence of the performed PE.	3.66	(1.17)	2.82	(1.27)	0.001
...gave constructive feedback regarding the performance of the physical examination of the thyroid gland	3.63	(1.06)	2.75	(1.16)	<0.001
...gave constructive feedback regarding the performance of the physical examination of the heart.	3.64	(1.20)	2.67	(1.11)	<0.001
...gave constructive feedback regarding the performance of the physical examination of the lungs.	3.79	(1.21)	2.79	(1.25)	<0.001
...gave constructive feedback regarding the performance of the physical examination of the abdomen.	3.78	(1.19)	2.78	(1.23)	<0.001
...gave constructive feedback regarding the doctor-patient communication during the PE.	3.05	(1.20)	3.21	(1.32)	0.500
...was attentive and focused while giving feedback.	2.19	(0.56)	2.05	(0.85)	0.456
...was friendly while giving feedback.	1.96	(0.37)	1.91	(0.51)	0.504
...gave specific examples while giving feedback.	2.41	(0.70)	2.13	(0.74)	0.038
...seemed competent in the field of PE while giving feedback.	3.19	(1.28)	1.91	(0.65)	<0.001
...seemed well prepared with regard to giving feedback (sandwich technique; feedback rules).	3.11	(1.45)	2.57	(1.22)	0.033
Overall	3.18	(1.05)	2.51	(1.05)	0.043

Values are shown as mean and standard deviation. Likert-scale ratings ranging from 1 (I fully agree) to 6 (I completely disagree)

<https://doi.org/10.1371/journal.pone.0180308.t002>

Questionnaire assessment of PI and physician feedback by video assessors

Concerning overall ratings by video assessors, physician feedback was significantly better rated than PI feedback ($p = .043$). There was no statistically significant difference concerning four items (see Table 2).

Questionnaire assessment of PI and physician feedback by students

Concerning overall ratings by students, physician feedback was significantly better rated than PI feedback ($p = .034$). There was no statistically significant difference concerning five items (see Table 3).

Table 3. Qualitative comparison of PI (n = 14) and physician (n = 14) feedback by students.

Items	PI (n = 14)		Physician (n = 14)		p-value
	Mean	(SD)	Mean	SD	
...gave detailed feedback on the execution of the physical examination (PE).	1.83	(1.03)	1,28	(0.55)	0.006
...gave detailed feedback on the compliance with framework conditions and on the sequence of the performed PE.	1.75	(0.84)	1,50	(0.78)	0.126
...gave constructive feedback regarding the performance of the physical examination of the thyroid gland	1.83	(1.08)	1,28	(0.68)	0.004
...gave constructive feedback regarding the performance of the physical examination of the heart.	1.80	(0.91)	1,25	(0.54)	0.001
...gave constructive feedback regarding the performance of the physical examination of the lungs.	1.98	(1.10)	1,40	(0.84)	0.005
...gave constructive feedback regarding the performance of the physical examination of the abdomen.	1.95	(1.13)	1,38	(0.59)	0.014
...gave constructive feedback regarding the doctor-patient communication during the PE.	1.33	(0.57)	1,30	(0.56)	0.812
...was attentive and focused while giving feedback.	1.03	(0.16)	1,20	(0.56)	0.133
...was friendly while giving feedback.	1.10	(0.30)	1,10	(0.30)	1.00
...gave specific examples while giving feedback.	1.36	(0.63)	1,30	(0.52)	0.836
...seemed competent in the field of PE while giving feedback.	2.25	(1.50)	1,21	(0.47)	0.001
...seemed well prepared with regard to giving feedback (sandwich technique; feedback rules).	1.73	(1.09)	1,35	(0.74)	0.092
Overall	1.66	(0.86)	1.30	(0.59)	0.034

Values are shown as mean and standard deviation. Likert-scale ratings ranging from 1 (I fully agree) to 6 (I completely disagree)

<https://doi.org/10.1371/journal.pone.0180308.t003>

Interrater reliability of feedback ratings

Interrater reliability of PI and physician feedback ratings (Table 2) was satisfactory to good (0.52 and 0.58, respectively).

Discussion

The present study prospectively examined whether feedback provided by specially trained PIs regarding a physical examination carried out on them is comparable with feedback provided by faculty staff physicians observing the examination. To this aim, both the examination procedure conducted by the student and the feedback from the PI and the physician were video-recorded. In a first step, the examination was evaluated by video raters using binary checklists, which subsequently served as gold standard. The feedback was then compared with this checklist in order to quantify the points mentioned in the feedback. Overall, both the physician feedback and the PI feedback contained only a small percentage of the possible items (between 11 and 34%). In terms of matching with the gold standard, no significant difference in performance was found between PIs and physicians. As the 147-item checklist is very comprehensive, it is understandably not possible to list all of these points individually in one feedback session. Therefore, an examination of the pure mistakes, the key steps and the communicative aspects was conducted. Again, no significant differences emerged between the PI and physician feedback on any of these points.

This equally good result of faculty staff physicians and PIs is particularly striking as the PIs were unable to make any notes during the examination, and had to remember the individual points, while the physicians—as is customary in normal teaching—already made notes for later feedback during the examination. Moreover, the finding shows that with corresponding training [35], it is possible for PIs to become very well qualified in a very short time.

However, the video raters evaluated the quality of the physician feedback as significantly better overall than that of the PI feedback. When observing the individual items, it becomes apparent that this difference is primarily—though not exclusively—based on the better feedback regarding the technical implementation of the physical examination. This indicates that despite intensive training, the PIs have not yet been able to acquire sufficient routine in this field, and possibly did not make as confident an impression on the students and video raters, even though no objective difference in the feedback on the items was apparent. As these were the first assignments of the trained PIs, this aspect might improve in the future with greater experience and more assignments. The subjective evaluation by the students also showed higher ratings for the physicians than for the PIs, although it should be noted that both groups received high ratings.

To our knowledge, the present study is the first to provide insight into the quality of PI feedback in direct comparison with physician feedback in relation to a complex and extensive task like the physical examination of four organ systems. The use of SPs as PIs in the area of physical examination is nothing new, having already been described some 20 years ago [36,37], although in most cases, no feedback on the examination steps is given. An exception to this is the use of PIs within OSCEs: In a survey of all German-speaking universities, it was found that in 31 of 39 participating universities, SPs were used for the provision of feedback, and in five universities they were also used as raters within the framework of OSCEs [38]. In this context, SPs frequently do not assess the examination per se, but rather assess, for example, whether the student deals with the SP in a professional manner [39]. Moreover, in the case of OSCEs, the extent of the task is also smaller than in our study, and the checklists accordingly shorter. Another finding that has not been previously described is the overall low feedback rate of both the physicians and the PIs. Presumably, this arises from the feedback provider's need to

primarily give feedback on the most salient positive and negative points, so that the student is not overwhelmed by a large amount of items. Nevertheless, it should be noted that only around 30% of key steps received feedback.

Future studies should employ different examinations to investigate how feedback providers select the items on which to give feedback, and the extent to which this proportion of feedback increases in situations encompassing fewer items to be rated. On the whole, the area of feedback provision by SPs remains relatively unexplored, with the AMEE Guide 2009 already concluding that “There is also a clear lack of studies with regard to the training for and the effect of giving feedback by SPs” [29]. However, the studies published since then have been devoted to other situations: For instance, in their study, Bowman et al studied the feedback of 8 licensed physical therapists serving as standardized patients for practical examinations in comparison to a course instructor [40]. Based on the intra-class correlation coefficient there was a significant difference between scores so the authors concluded, that standardized patients might not be an adequate replacement for an instructor. However, the results of the standardized patients and the instructor were not compared to a gold standard. May et al. describe the use of PIs within a Clinical Performance Examination (CPX) Test, in which both history-taking and physical examination as well as patient consultation are assessed [41]. Through intensive training, a checklist congruence of >85% was reached, which was also achieved in our study, at least in terms of matching with the gold standard. However, May et al. do not describe the total extent of the checklist and the proportion of physical examination. Nevertheless, it can be assumed that the number of assessed items is a great deal higher in the present study.

Limitations

Several limitations of this study should be mentioned. First, the percentage of points that received feedback is relatively low compared to the gold standard. This does, however, reflect everyday life, in which feedback must be limited to the points that appear to be most important, and it is almost never possible to provide comprehensive feedback. Although our PIs had completed extensive training, these were their first “real” assignments as a PI who is examined and provides feedback directly afterwards, while the physicians already had several years of experience both in physical examination and in teaching itself. Therefore, it may be the case that with the experience that the PIs will gather over time, the results will improve further. Although the sample size of the present study and the number of involved PIs and physicians was rather small, we were able to collect two to three cases per physician and PI in order to minimize the between-case variance. Furthermore, the existence of a pool of 12 different PIs trained in physical examination feedback skills is rather unique.

Conclusions

In summary, our study demonstrates that trained PIs are able to provide feedback of equal quantitative value to that of faculty staff physicians with regard to a physical examination performed on them. However, both the students and the video raters judged the quality of the feedback given by the physicians to be significantly better than that of the PIs. This is interesting insofar as PIs are already trained in providing feedback prior to the specialization in physical examination. Thus, under considerations of personnel and financial resources, it is reasonable to deploy PIs to assess students’ physical examination skills. Further studies should investigate whether these results could be improved further by PIs who have completed longer training or who possess greater experience in the area of physical examination.

Author Contributions

Conceptualization: Markus Krautter, Jobst-Hendrik Schultz, Nadja Koehl-Hackert, Wolfgang Herzog, Christoph Nikendei.

Data curation: Katja Diefenbacher, Imad Maatouk, Anne Herrmann-Werner, Nadja Koehl-Hackert.

Formal analysis: Markus Krautter, Katja Diefenbacher, Imad Maatouk, Anne Herrmann-Werner, Nadja Koehl-Hackert, Christoph Nikendei.

Investigation: Markus Krautter, Katja Diefenbacher, Imad Maatouk, Anne Herrmann-Werner, Nadja Koehl-Hackert, Christoph Nikendei.

Methodology: Markus Krautter.

Project administration: Jobst-Hendrik Schultz, Wolfgang Herzog, Christoph Nikendei.

Resources: Jobst-Hendrik Schultz, Wolfgang Herzog, Christoph Nikendei.

Supervision: Markus Krautter, Jobst-Hendrik Schultz, Wolfgang Herzog, Christoph Nikendei.

Validation: Markus Krautter, Katja Diefenbacher, Imad Maatouk, Anne Herrmann-Werner, Nadja Koehl-Hackert, Christoph Nikendei.

Visualization: Markus Krautter.

Writing – original draft: Markus Krautter, Christoph Nikendei.

Writing – review & editing: Markus Krautter, Jobst-Hendrik Schultz, Imad Maatouk, Anne Herrmann-Werner, Nadja Koehl-Hackert, Wolfgang Herzog, Christoph Nikendei.

References

1. Reilly BM (2003) Physical examination in the care of medical inpatients: an observational study. *Lancet* 362: 1100–1105. [https://doi.org/10.1016/S0140-6736\(03\)14464-9](https://doi.org/10.1016/S0140-6736(03)14464-9) PMID: 14550696
2. Nikendei C, Kraus B, Schrauth M, Briem S, Junger J (2008) Ward rounds: how prepared are future doctors? *Med Teach* 30: 88–91. <https://doi.org/10.1080/01421590701753468> PMID: 18278658
3. Ortiz-Neu C, Walters CA, Tenenbaum J, Colliver JA, Schmidt HJ (2001) Error patterns of 3rd-year medical students on the cardiovascular physical examination. *Teach Learn Med* 13: 161–166. https://doi.org/10.1207/S15328015TLM1303_5 PMID: 11475659
4. Krautter M, Diefenbacher K, Koehl-Hackert N, Buss B, Nagelmann L, Herzog W, et al. (2015) Short communication: final year students' deficits in physical examination skills performance in Germany. *Z Evid Fortbild Qual Gesundhwes* 109: 59–61. <https://doi.org/10.1016/j.zefq.2015.01.003> PMID: 25839370
5. Barley GE, Fisher J, Dwinell B, White K (2006) Teaching foundational physical examination skills: study results comparing lay teaching associates and physician instructors. *Acad Med* 81: S95–97. PMID: 17001147
6. Dull P, Haines DJ (2003) Methods for teaching physical examination skills to medical students. *Fam Med* 35: 343–348. PMID: 12772936
7. Martens MJ, Duvivier RJ, van Dalen J, Verwijnen GM, Scherpbier AJ, van der Vleuten CP (2009) Student views on the effective teaching of physical examination skills: a qualitative study. *Med Educ* 43: 184–191. <https://doi.org/10.1111/j.1365-2923.2008.03283.x> PMID: 19161490
8. Barrows HS (1993) An overview of the uses of standardized patients for teaching and evaluating clinical skills. *AAMC. Acad Med* 68: 443–451; discussion 451–443. PMID: 8507309
9. Holmboe ES, Hawkins RE (1998) Methods for evaluating the clinical competence of residents in internal medicine: a review. *Ann Intern Med* 129: 42–48. PMID: 9652999
10. May W, Park JH, Lee JP (2009) A ten-year review of the literature on the use of standardized patients in teaching and learning: 1996–2005. *Med Teach* 31: 487–492. PMID: 19811163

11. Tuzer H, Dinc L, Elcin M (2016) The effects of using high-fidelity simulators and standardized patients on the thorax, lung, and cardiac examination skills of undergraduate nursing students. *Nurse Educ Today* 45: 120–125. <https://doi.org/10.1016/j.nedt.2016.07.002> PMID: 27449150
12. Lopreiato JO, Sawyer T (2015) Simulation-based medical education in pediatrics. *Acad Pediatr* 15: 134–142. <https://doi.org/10.1016/j.acap.2014.10.010> PMID: 25748973
13. Wright B, McKendree J, Morgan L, Allgar VL, Brown A (2014) Examiner and simulated patient ratings of empathy in medical student final year clinical examination: are they useful? *BMC Med Educ* 14: 199. <https://doi.org/10.1186/1472-6920-14-199> PMID: 25245476
14. Swanson DB, van der Vleuten CP (2013) Assessment of clinical skills with standardized patients: state of the art revisited. *Teach Learn Med* 25 Suppl 1: S17–25.
15. Davidson R, Duerson M, Rathe R, Pauly R, Watson RT (2001) Using Standardized Patients as Teachers: A Concurrent Controlled Trial. *Academic Medicine* 76: 840–843. PMID: 11500289
16. Hatala R, Issenberg SB, Kassen B, Cole G, Bacchus CM, Scalese RJ (2008) Assessing cardiac physical examination skills using simulation technology and real patients: a comparison study. *Med Educ* 42: 628–636. <https://doi.org/10.1111/j.1365-2923.2007.02953.x> PMID: 18221269
17. Hatala R, Issenberg SB, Kassen BO, Cole G, Bacchus CM, Scalese RJ (2007) Assessing the relationship between cardiac physical examination technique and accurate bedside diagnosis during an objective structured clinical examination (OSCE). *Acad Med* 82: S26–29. <https://doi.org/10.1097/ACM.0b013e31814002f1> PMID: 17895683
18. Aamodt CB, Virtue DW, Dobbie AE (2006) Trained standardized patients can train their peers to provide well-rated, cost-effective physical exam skills training to first-year medical students. *Fam Med* 38: 326–329. PMID: 16673193
19. Chalabian J, Dunnington G (1997) Standardized patients: a new method to assess the clinical skills of physicians. *Best Pract Benchmarking Healthc* 2: 174–177. PMID: 9362616
20. Adamo G (2003) Simulated and standardized patients in OSCEs: achievements and challenges 1992–2003. *Med Teach* 25: 262–270. <https://doi.org/10.1080/0142159031000100300> PMID: 12881047
21. Fletcher KE, Stern DT, White C, Gruppen LD, Oh MS, Cimmino VM (2004) The physical examination of patients with abdominal pain: the long-term effect of adding standardized patients and small-group feedback to a lecture presentation. *Teach Learn Med* 16: 171–174. https://doi.org/10.1207/s15328015tlm1602_9 PMID: 15276894
22. Hattie J, Timperley H (2007) The Power of Feedback. *Review of Educational Research* 77: 81–112.
23. Veloski J, Boex JR, Grasberger MJ, Evans A, Wolfson DB (2006) Systematic review of the literature on assessment, feedback and physicians' clinical performance*: BEME Guide No. 7. *Medical Teacher* 28: 117–128. <https://doi.org/10.1080/01421590600622665> PMID: 16707292
24. Bokken L, Linssen T, Scherpbier A, van der Vleuten C, Rethans JJ (2009) Feedback by simulated patients in undergraduate medical education: a systematic review of the literature. *Med Educ* 43: 202–210. <https://doi.org/10.1111/j.1365-2923.2008.03268.x> PMID: 19250346
25. Doyle Howley L, Martindale J (2009) The Efficacy of Standardized Patient Feedback in Clinical Teaching: A Mixed Methods Analysis.
26. Kluger AN, Denisi A (1996) The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin* 119.
27. Van De Ridder JMM, Stokking KM, McGaghie WC, Ten Cate OTJ (2008) What is feedback in clinical education? *Medical Education* 42: 189–197. <https://doi.org/10.1111/j.1365-2923.2007.02973.x> PMID: 18230092
28. Bosek MS, Li S, Hicks FD (2007) Working with standardized patients: a primer. *Int J Nurs Educ Scholarsh* 4: Article 16.
29. Cleland JA, Abe K, Rethans J-J (2009) The use of simulated patients in medical education: AMEE Guide No 421. *Medical Teacher* 31: 477–486. PMID: 19811162
30. Stillman PL, Regan MB, Philbin M, Haley HL (1990) Results of a survey on the use of standardized patients to teach and evaluate clinical skills. *Acad Med* 65: 288–292. PMID: 2337429
31. Wykurz G, Kelly D (2002) Developing the role of patients as teachers: literature review. *BMJ* 325: 818–821. PMID: 12376445
32. Pjontek R, Scheibe F, Tabatabai J (2012) Heidelberger Standarduntersuchung—Interdisziplinäre Handlungsanweisungen zur Durchführung der körperlichen Untersuchung; Heidelberg MF, editor: Bezug über: standard.untersuchung@med.uni-heidelberg.de.
33. Nikendei C, Ganschow P, Groener JB, Huwendiek S, Kochel A, Kohl-Hackert N, et al. (2016) "Heidelberg standard examination" and "Heidelberg standard procedures"—Development of faculty-wide

- standards for physical examination techniques and clinical procedures in undergraduate medical education. *GMS J Med Educ* 33: Doc54. <https://doi.org/10.3205/zma001053> PMID: 27579354
34. Regehr G, MacRae H, Reznick RK, Szalay D (1998) Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Academic Medicine* 73: 993–997. PMID: 9759104
 35. Nikendei C, Diefenbacher K, Kohl-Hackert N, Lauber H, Huber J, Herrmann-Werner A, et al. (2015) Digital rectal examination skills: first training experiences, the motives and attitudes of standardized patients. *BMC Med Educ* 15: 7. <https://doi.org/10.1186/s12909-015-0292-7> PMID: 25638247
 36. De Champlain AF, Margolis MJ, King A, Klass DJ (1997) Standardized patients' accuracy in recording examinees' behaviors using checklists. *Acad Med* 72: S85–87. PMID: 9347749
 37. Epstein RM, Hundert EM (2002) Defining and assessing professional competence. *JAMA* 287: 226–235. PMID: 11779266
 38. Hartl A, Bachmann C, Blum K, Hofer S, Peters T, Preusche I, et al. (2015) Desire and reality—teaching and assessing communicative competencies in undergraduate medical education in German-speaking Europe—a survey. *GMS Z Med Ausbild* 32: Doc56. <https://doi.org/10.3205/zma000998> PMID: 26604998
 39. Homer M, Pell G (2009) The impact of the inclusion of simulated patient ratings on the reliability of OSCE assessments under the borderline regression method. *Med Teach* 31: 420–425. <https://doi.org/10.1080/01421590802520949> PMID: 19142798
 40. Bowman DH, Ferber KL, Sima AP (2016) Inter-rater Agreement on Final Competency Testing Utilizing Standardized Patients. *J Allied Health* 45: 3–7. PMID: 26937875
 41. May W (2008) Training standardized patients for a high-stakes Clinical Performance Examination in the California Consortium for the Assessment of Clinical Competence. *Kaohsiung J Med Sci* 24: 640–645. [https://doi.org/10.1016/S1607-551X\(09\)70029-4](https://doi.org/10.1016/S1607-551X(09)70029-4) PMID: 19251559