

## An overview of and approach to selecting appropriate patient representations in teaching and summative assessment in medical education

Bauer Daniel<sup>a</sup>, Lahner Felicitas-Maria<sup>b</sup>, Huwendiek Sören<sup>a</sup>, Schmitz Felix M.<sup>a</sup>, Guttormsen Sissel<sup>a</sup>

<sup>a</sup> Institute for Medical Education, University of Bern, Switzerland

<sup>b</sup> Department of Health Professions, Bern University of Applied Sciences, Switzerland

### Summary

Medical education has a long tradition of using various patient representations in teaching and assessment. With this literature review we aim, first, to provide an overview of the most important patient representations used to teach and assess clinical skills, considering in particular “summative exams” that have a pass or fail outcome; second, to provide arguments for choosing certain patient representations; and third, to show the advantages and limitations of different patient representations, especially simulated patients (SPs) and real patients (RPs).

Typical patient representations include case narratives, anatomical models, simulators and mannequins, as well as SPs and RPs. The literature indicates that there are multiple ways of using various patient representations in teaching and that the intended didactical purpose informs the choice of representation. Early in the educational programme, even low-fidelity patient representations can be a good fit for assessment purposes if chosen to match the educational level. The use of RPs in summative, high-stakes assessments (exams with particularly important consequences for the examinee) is limited for methodological and ethical reasons. The methodological implementation of summative exams also entails specific challenges, such as ensuring measurement reliability and fairness towards the examinees. Carefully prepared, SPs can perform their roles with a sufficient degree of authenticity, making summative exams more manageable, and imposing no strain or risk on RPs. The ongoing debate concerning the use of SPs and RPs in summative assessment highlights perceived limitations of SPs in relation to RPs that are often not supported by research. Evidence shows that SPs, in combination with additional simulation modalities as needed, represent the first choice for summative clinical assessment. We also consider the strengths and limitations of this review and reflect on the applicability of our findings.

We conclude that in order to select the right patient representations in clinical teaching and/or assessment, a number of perspectives must be considered: (i) the learning goals, aligned with the stage of study, (ii) the correspond-

ing requirements of the clinical task itself (e.g., performing a phlebotomy or a communication task), (iii) the level of authenticity required and (iv) the resources needed, taking patient safety and feasibility into consideration.

**Keywords:** *patient representations, teaching, summative assessment, medical education*

### Introduction

Effective medical teaching relies on educators selecting, based on specific educational goals and the curricular setting, the most appropriate learning aids from a multitude of options. Because the central purpose of teaching, learning and assessment in medical education is to optimise patient care, educators must ensure that students have established the connections between what is taught, what is assessed, and how these relate to patient care. Medical teaching must also consider how to present a patient while accommodating the demands of the particular learning and assessment task, the students’ progress within the medical programme, and the educators’ preferences. A “patient” may be presented in several ways: as a real patient (RP) (e.g., live during a lecture or through bedside teaching) or by making use of alternative formats such as case narratives, cadavers, anatomical models, mannequins, simulators or simulated patients (SPs) (see [figure 1](#) for an overview).

Clinical summative assessment aims to determine a learner’s level of competence, adding another set of factors to be considered when choosing the best patient representation. Summative assessments are exams that can eventually determine whether a student continues the programme or not. To put this in context, it is necessary to first look at why different patient representations are used in medical education. In order to constructively align assessments with learning and teaching, they must be viewed within the integrated continuum of medical education [15, 16]. Questions regarding when students should interact with RPs, which tasks they should practice on such patients and the associated benefits of such interactions have been topics of discussion in medical schools across the world, as captured by a systematic review [17]. The main goal of employing

#### Correspondence:

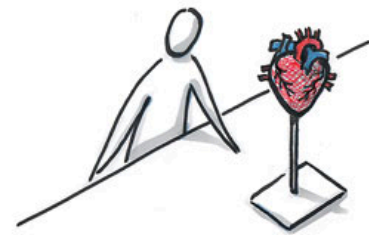
Prof. Sissel Guttormsen, MD, Institute for Medical Education, University of Bern, Mittelstrasse 43, CH-3012 Bern, [sissel.guttormsen\[at\]iml.unibe.ch](mailto:sissel.guttormsen[at]iml.unibe.ch)

**Figure 1:** Various forms of patient representation. (A) *Case narratives:* a narrative approach describes a patient and their medical history [1], not unlike a doctor's letter. Narratives are often based on a real patient's story. Delivered in text or digitally and sometimes media-enriched, these patient representations are often encountered in problem-based learning settings, in blended-learning scenarios or as virtual patients [2]. (B) *Anatomical models,* including cadavers [3] and some e-learning/virtual reality tools [4], are patient representations that reduce physiological complexity to emphasise anatomy. They are used in undergraduate education and at the postgraduate level, for example, for practicing surgical skills [5]. (C) *Simulators and mannequins* reduce the complexity of a patient to a set of defined functions, and range from part-task trainers to highly authentic simulators [6]. They can be combined with simulated patients in so-called hybrid simulations (combining simulation modalities) [7]. (D) *Simulated patients (SPs),* also known as standardised patients, are professional or lay actors who have been carefully trained to assume a patient's role, following an elaborate script on how to react to certain questions and how to present specific findings [8]. The range of portrayable pathologies can be expanded, for example, by using simulation technology [9] or moulage techniques [10], in "hybrid simulations". SPs and hybrid simulations are employed in dedicated courses [11] or in assessment [12]. (E) *Real patients* are hospitalised or in ambulant care due to a medical problem. They are an important resource in education and formative assessment, as they have the highest ecological validity of any form of patient representation [13]. In order to protect them from potential harm, especially in undergraduate settings, they should be carefully chosen for the specific teaching or

formative assessment scenario [14].



A. Case narrative



B. Model



C. Simulator / mannequin



D. Simulated patient



E. Real patient

various patient representations is to enable students to master essential clinical skills and knowledge prior to entering real-life clinical settings [18]. Educational patient representations simulate a variety of situations. It has been shown that students benefit from the use of SPs more than from interactions with RPs during their education [19].

The three research questions addressed in this paper are Q1: what should educators and assessors know about various patient representations? Q2: which considerations should influence the choice of patient representations in medical teaching and assessment? Q3: what are the specific advantages and limitations of SPs and RPs for assessment purposes?

## Method

We performed a literature review to clarify the research questions according to Grant and Booth [20]. This was a recursive process with the following steps:

(a) *Discussing the focus and structure of a paper in relation to the research questions:* The use of SPs representing RPs in summative assessment continues to be a topic of debate. We suggest setting this debate within a broader context including medical teaching and the challenges of summative assessment. Therefore, we present an overview of various patient representations and arguments for their use and relevance in teaching and assessment. Only against this background can a sound discussion of the benefits and disadvantages of SPs and RPs in medical assessment follow. Our chosen approach thus considers patient representations in medical teaching and in summative assessment and includes a discussion on SPs and RPs in summative assessment, which will appeal to the journal's broad readership, as well as to medical education experts. Therefore, psychometric information (i.e., methodological aspects related to the trustworthiness of psychological data and the development of corresponding measurement instruments) and other methods have been kept to a minimum. Our fo-

cus is on summative assessment (assessment of learning), rather than formative assessment (assessment for learning), which is a separate research area and beyond the scope of the paper.

*(b) Searching for and selecting materials for inclusion:* Searches were performed in Scopus, PubMed and Google Scholar, as well as the authors' and their colleagues' own literature collections. The search terms varied according to the specific issues addressed. When multiple publications were found covering an issue, the most recent were selected, with a preference for renowned, peer-reviewed journals. In order to balance aspects related to teaching and to assessment, particular literature was searched. Textbooks were also included, particularly for well-established concepts (e.g., OSCE exams, i.e., practical clinical skills exams). "Grey literature" was considered in a few cases where better sources could not be found and an argument had to be further elaborated.

*(c) Synthesising the material for presentation in the paper:* We identified various patient representations based on our own experience and verified these in the literature. We identified factors related to the selection of patient representations for teaching and assessment based on our own experience and continued to search for literature documenting all relevant and further aspects.

*(d) Analysing the outcome and deriving conclusions:* The authors wrote various sub-sections of the paper according to their specialist knowledge. The team met regularly to discuss and review drafts. The analysis was led by critical reflections and alignment to the overall aims. An illustrator was asked to visualise the patient representations. The drawings were revised until they fitted the aims. The table representing a Swiss medical curriculum was reviewed by a curriculum expert. The conclusions emerged from the authors' critical reflections and additional expert feedback. Finally, in order to check for readability and comprehensibility, we asked two independent clinicians with no background or formal training in medical education to highlight areas that were difficult to understand and to critically review our manuscript. The issues they highlighted were revised accordingly.

## Patient representations in medical teaching

### Benefits of employing RPs in teaching

Patient contact is integral to medical education [13]. For example, both the Association of American Medical Colleges and the UK General Medical Council recommend early patient contact [21]. It has been shown that integrating patients into undergraduate medical education has several advantages. It provides real-life learning opportunities in which students can link their theoretical knowledge with practice. Patient contact motivates students by giving relevance and context to the theoretical content of their educational course, which otherwise might remain abstract. In this way, real patient contact supports learning in both biomedical and social sciences. Additionally, interacting with patients during medical education fosters students' self-awareness, confidence, self-reflection and self-appraisal skills [21–25]. Regular patient contact is known to reduce stress, as it enables students to become familiar with their future role [21–23]. It also helps students to de-

velop a professional identity and to hone their clinical reasoning skills [21, 22]. There is evidence that it encourages students to develop a more patient-centred approach [24]. From the patient perspective, many patients enjoy participating in teaching. They find it a reassuring experience that helps them to better understand their illness [13]. However, the main advantage of including RPs in teaching is the presentation of actual pathological findings and allowing students to gain unique insights from exploring the patient's perspective.

### Limiting aspects of employing RPs in teaching

Patients participating in teaching and assessment are, unfortunately, also at risk of suffering negative consequences from their commendable willingness to be involved. First, using RPs as a teaching resource can put a strain on the patient-physician relationship: a good rapport between patient and student cannot be established when the encounters are too short, happen too often or are with too many different students. Such occurrences can impair the one-to-one relationship between a patient and their assigned attending physician. Patients' trust in physicians can be violated if they were initially expecting to be treated by a trained physician but instead find themselves being attended and examined by students [26]. Second, using RPs in teaching may violate ethical principles if patients' participation in teaching is imposed as a moral obligation. Waterbury [26] addresses the issue that a patient's informed consent to participate in teaching is often not addressed. This might be due to patients' unawareness of their right to refuse participation [27]. Third, participating in teaching may result in harm and discomfort to patients despite supervision. Students have, after all, not yet mastered their skills and lack some precision when performing routine procedures [26]. Last but not least, reduced inpatient bed capacities and shorter hospital stays have resulted in fewer suitable patients being available, while a simultaneous increase in the number of students requiring patient contact has put an additional strain on the remaining patients. Students usually spend much of their education in university hospitals, where the sample of patients is not necessarily representative of the wider medical reality. Both these aspects can be alleviated somewhat when community hospitals and private practices are involved in teaching, as is increasingly the case at many teaching institutes [28]. Therefore, the use of RPs in medical education must always balance patient safety and well-being with the need for RPs when teaching students. This represents a fundamental ethical tension in medical education [29].

### Didactical aspects of teaching and their impact on patient representations

As well as patient safety and ethical aspects, there are didactic reasons for using patient representations other than RPs in teaching activities. Students learn to become skilled physicians by building knowledge from elementary facts into larger units of knowledge and partial skills. Individual knowledge emerges from a constant process of teaching and learning with deconstructed units of complex objectives on the one hand and reassembling the full complexity during progressing clinical teaching [30] on the other. Cognitive load theory [31] describes typical strains on learners and suggests how cognitive constraints can be overcome

when the learning content is too complex. Both in medical education and in other educational fields, it has been shown that students' learning is easily hindered when they are tasked with combining various elements of a complex issue in order to gain consolidated and applicable knowledge [32, 33]. Hence, during clinical training, a stepwise approach to acquiring complex clinical skills is a fundamental aspect of teaching [30]. Typically, students approach interactions with RPs and the performance of clinical tasks by first going through a process of learning specific skills in a context of reduced authenticity and complexity. Examples include students learning to auscultate and identify pathological murmurs by practicing on each other or by using auscultation simulators; they might also use e-learning resources with a catalogue of auscultatory findings to practice differentiating the various sounds and their pathologies (for an example, see <https://elearning.medizin.unibe.ch/>). Also, students can make use of so-called "virtual patients", which can prepare learners for what questions to ask when taking a history or which physical exams are to be performed in a specific patient case, as well as enable them to practice their clinical reasoning skills, all in preparation for encounters with real patients [34, 35]. Many clinicians are familiar with the guiding principle of "see one, do one, teach one" [36]. Obviously, this principle does not mean that a single demonstration suffices, but still nicely illustrates that performing and mastering (teach one) clinical procedures is preceded by safe, consecutive learning steps [36]. While this approach is not wrong, it no longer fulfils the needs of modern medical education. Today, there is a need for flexible and scalable teaching approaches which address prevailing and challenging patient cases, as well as safe ways of performing/practicing such (patient) interactions [37]. This means that the choice of the most suitable patient representation in a given teaching scenario depends on the demands of the educational task, the students' progress within the programme, the resources available and ethical considerations.

#### Various forms of patient representations in a typical Swiss curricular context

The objectives of medical education in Switzerland are described in the Federal Act of Medical Professions (Med-

BG) [38] and in the national catalogue of learning objectives [39]. In this framework, the knowledge and skills expected by the end of the medical programme are grouped into three levels, i.e., "roles of a doctor" [40], "entrustable professional activities" and "situations as a starting point". According to the MedBG [38], undergraduate medical education at Swiss universities aims to prepare students to enter further professional development. Only after professional specialisation do physicians act independently. Until this level of professional practice is reached, education proceeds through many stages of fundamental learning and training in which patients are presented in different ways. Typical forms of patient representation, with complexity increasing over the course of education, are shown in figure 1. Each of these patient representations has advantages and limitations. If implemented as intended, in an appropriate didactical context and at an appropriate level of training, they are known to serve their purpose well [6, 13, 41, 42].

The composition of a typical Swiss medical curriculum is illustrated in table 1 (adapted from the curriculum at the Medical Faculty in Bern). It illustrates possible implementations of various patient representations in clinical teaching over the course of medical education and training.

#### Using SPs and RPs in medical teaching

Before addressing the implications of their use in *assessment*, we need to examine the use of SPs and RPs in *medical education* in further detail. There is an unrelenting debate on the usefulness of SPs in training and assessment [43]. Due to the shift to integrated curricula that introduce early clinical practice to medical programmes, patient participation in teaching and learning has increased [44, 45]. However, early patient contact can be counterproductive for trainees lacking a minimum of basic clinical knowledge and skills, such as knowing how to take a patient history [46]. Therefore, it is essential that students get early and safe training to prepare them for tasks with real patients.

Despite existing reservations [43], SPs are widely employed in medical teaching to prepare students for clinical practice. There are many reasons for this, but patient safety and the need for predictable patients in training and as-

**Table 1:** Examples of patient representations during a Swiss medical education curriculum.

	Pre-medical school	Bachelor year 1	Bachelor year 2	Bachelor year 3	Master year 1	Master year 2	Master year 3	Federal licensing exam	Residency	Post-graduate practice
<b>Educational Framework</b>		<i>From anatomy and physiology in healthy humans to pathological anatomy functions</i>			<i>From patients and pathology to diagnostics and therapy</i>				<i>Decreasing supervision</i>	
<b>Real patients</b>	Nursing clerkship	Ambulatory care, lectures, clinical skills training			Internships, ambulatory care, lectures		Elective internship		Real patients in various settings	
<b>Simulated patients</b>		History taking and feedback training		Exam: OSCE	Communication skills, geriatrics training	Communications skills training Exam: OSCE	Communication skills training	Clinical skills exam		
<b>Simulators, mannequins</b>	First aid course		Clinical skills training	Exam: OSCE		Exam: OSCE			Various training courses: surgical skills, anaesthesia, emergency, resuscitation	
<b>Models</b>		Gross anatomy, cadaver course								
<b>Case narratives</b>		Problem-based learning		Written / electronic exam	Written / electronic exam	Written / electronic exam		Written / electronic exam		



assessment situations are among the most prominent. An SP presents the patient: not just the history, but also the body language, aspects of the physical exam, and emotional and personality-related characteristics [47]. Therefore, SPs can be seen as “expert patients” [48]. The advantages of SPs in teaching are that they are available when required, can be trained in a broad range of clinical cases, and can be used for repeated practicing of new tasks. Thus, they can repeat scenarios many times, their behaviour is predictable, and they can portray patient roles where the use of a real patient would be inappropriate (e.g., in an emergency situation) [44]. Another self-evident example where SPs are useful is in training students to handle challenging communication scenarios [49, 50], such as breaking bad news (e.g., a terminal diagnosis) to a patient, before they are allowed to do so in real life. It is a quasi-imperative of medical professionalism and ethos that an RP in such a grave situation is cared for by someone proficient in the matter. Such skills must therefore be taught with the next best patient representation – an SP. Ideally, when to employ SPs, RPs or other representations in teaching activities and the way this is embedded in the curriculum as a whole should be carefully considered.

## Patient representations in medical summative assessment

### Aims of medical assessment

Whether to use SPs or RPs in assessment is mostly a continuation of the debate related to education: assessment must be aligned with the particular aims and level of the education. Ideally, it should use the same didactical instruments that were used in the teaching. Hence, the patient representations shown in figure 1 are also relevant and utilised for summative assessment at different educational levels.

When choosing the optimal patient representation in the context of summative assessment, additional factors beyond those related to teaching must be considered. The assessment of clinical skills has formed a key aspect of medical education for centuries [51]. One of the main goals of these assessments is to decide whether a student is competent enough to continue his or her education, or in the case of a licensing examination, whether he or she can become a physician [52]. In fact, patient safety is the main reason for elaborating the assessment process in medical education [53]. It is to protect patients' well-being that students are regularly tested to ascertain if they can be entrusted to interact with patients, to perform ever more challenging tasks and to meet patient needs [54, 55]. This determination, however, should optimally be based on programmatic assessment through a series of tasks [56].

Also, the recognised “failure to fail effect”, where teachers and examiners struggle to communicate to their students that the observed performance was substandard, emphasises the need for objective exams where underperforming students are evaluated on objective grounds. Assessors should be relieved from conflicting responsibilities, i.e., supporting students' learning and at the same time assessing them [57, 58]. Hence, in summative clinical exams there are many arguments that advocate methodological rigor at the cost of the ultimate authenticity of a real and

complex patient [52, 59]. In the following, we address key factors that should be optimised in exams: reliability, validity and authenticity.

### Reliability – why exams should be trustworthy

Just as we want airplane pilots to be assessed in a way that only permits those able to perform their job to qualify, we also require medical exams to differentiate between competent and incompetent doctors in order to safeguard patient safety. A summative exam is an instrument that must deliver reliable measures; it is not a teaching situation [60]. When discussing assessments from a psychometric perspective, the main argument is reproducibility, or reliability. Reliability describes the assumption that a student who has passed (or failed) an exam would also pass (or fail) a similar exam [52]. To achieve reliability, it is necessary to minimise measurement errors, i.e. all factors that impact an examinee's result that arise from anything *other* than his or her ability [61]. Both SPs and RPs can influence the measurement error in an exam. A contribution from the 1960s [62] illustrates this in context: a professor at the Department of Neurology of the University of South California in Los Angeles noticed that when employing RPs in assessments, their reaction to the examinees was difficult to predict, and they did not always present the same history or physical findings to one examinee as they did to a previous examinee. Barrows and colleagues continued to publish on this issue over several years [63]. While these inconsistencies might be due to memory limitations in the patients [64] or the absence of training for the task, they could also be influenced by the fact that the patients know some of the examinees from previous encounters during their internships [44]. Thus, comparing the performances of several students can be difficult even when they interact with the same RP, as the examinees' results are not only influenced by their ability, but also by the RP. In response to this, SPs were introduced to medical assessment. As SPs are trained to portray a role, they perform similarly in repeated encounters and their responses to students are predictable and similar from one encounter to another [44]. Additionally, several SPs can be trained for the same role, ensuring similar conditions in concurrent exams in various locations, such as the Swiss federal licensing examination [65]. In this case, candidates encounter the same exam situations even if they are geographically distributed, and their performance can be attributed to their ability rather than other variables. Thus, in a reliable exam we anticipate that a candidate will perform similarly regardless of the exam site and/or examiners. From a psychometric point of view, we conclude that employing SPs rather than RPs increases reliability in summative assessment.

### Authenticity, validity and feasibility of assessments with SPs and RPs

It is generally accepted that RPs are a more valid patient representation in terms of an authentic presentation of clinical reality. However, assessment validity also requires the test to accurately assess what it is intended to measure [66, 67]. In this context, valid exams are subject to the same considerations as described previously for teaching, i.e., they need to match the educational level, align with previous teaching and avoid aspects outside the scope of what is being tested. Consequently, full authenticity in summative

exams is unfeasible (with respect to cost and effort), not didactically sound, not ethically defensible, and therefore also not valid [52, 66]. We look at the challenging balance between authenticity, validity and feasibility below.

In the context of summative medical assessment, it is not always the aim to maximise authenticity even when feasible, but rather the targeted assessment aim is the pivotal issue: maximising authenticity by employing RPs in pre-clinical assessments (e.g., demonstrating how to perform a phlebotomy according to a prescribed procedure) may be feasible in terms of organising the RPs for the assessment, but it will fail to justify the RPs' time, risk or the assessment objective (demonstrating mastery of the procedure and basic technical skills). In addition, the validity of a test's assessment of the key issues would be reduced if RPs were employed because they would introduce task-irrelevant aspects into the situation (e.g., each RP's individual medical history or communication style). In the above example of a phlebotomy, an SP combined with a model arm, or simply a model arm alone, is a valid representation with which to assess the task. Towards the end of medical education, however, demonstrating clinical skills in an authentic assessment setting becomes more relevant, for example when demonstrating how to take a patient history. Performing such a task with an RP is likely to give relevant performance cues, given that the RPs are specifically selected for the summative exam. Still, the risks of reducing a test's reliability and causing harm to the patients remain. Therefore, sufficient authenticity can also be provided by SPs when the exam aims to, for example, assess procedural and physical exam skills, empathic communication and complete history taking.

The reliability and consistency of SPs' role play is a critical issue in summative exams and is addressed through research [68]. An exam which is not reliable also does not fulfil the criteria for validity (i.e., it does not assess what it intends to assess). Consistent role play depends on the quality of the training provided, and any statement on role play quality depends on the method used to determine its quality. Therefore, many studies reporting on role play consistency also report on efforts to develop instruments to assess SPs and factors that influence good role play [63, 69–71]. Training the assessors of a practical exam is an equally important factor, one which also influences the exam's validity. Therefore, standards for their instruction, as well as approaches to training such assessors, are important considerations when preparing a summative exam. Guidelines for this process are regularly published [72, 73].

#### Experiences with RPs in summative assessment

There are a few studies reporting the use of RPs instead of SPs in summative assessments such as an OSCE [74]. An early Association for Medical Education in Europe (AMEE) guide [43] shares positive experiences of RP use in OSCEs. Sayer et al. [75] report on validity benefits (authenticity) when using RPs in summative assessment in the UK. However, the necessary effort of recruiting and involving RPs is described as laborious. Notably, Sayer's publication also reports adverse effects occurring in all 19 medical schools that employed RPs in standardised assessments. The adverse effects ranged from deterioration of an

existing medical condition during the exam to the onset of labour. Three schools even reported having experienced a patient dying during an exam. In response to Sayer's article, Newble [76] argues that the use of outpatients would minimise the risks of the reported adverse events and that managing outpatients would have many similarities with using SPs. This emphasises that even with trained RPs, the issues of ethics and feasibility remain.

We therefore need to take a closer look at the debate on using SPs and RPs in summative assessment. Most of the arguments related to the use of SPs in assessment have been conveniently summarised [77]. A review on using SPs in speech pathology education programmes highlights many of the same challenges and limitations with respect to assessment [78]. We have compiled the known limitations and advantages of using SPs and RPs in assessment from the literature and discuss these further below.

### Advantages and limitations of SPs and RPs in summative assessment

#### The quality of SPs and RPs with regard to displaying pathological findings

It can be argued that RPs inherently present the appropriate clinical findings, as they actually exhibit the illness, and that the medical history they report is authentic, whereas SPs must simulate these findings. For symptoms that cannot be authentically simulated, another representation must be employed [79]. It is argued that this hinders valid assessment when using an SP [44, 77].

Although RPs are physically authentic, there are many arguments against employing RPs in summative exams. These are best encapsulated by the professional credo "above all do no harm – *primum nihil nocere*". This rules out using RPs in exams that involve medical emergencies or psychiatric states such as suicidal tendencies. It also excludes many acute infections, challenging communication tasks (e.g. breaking bad news, domestic abuse), paediatric problems and any task too painful or intimate. Such situations are challenging in a way that threatens both the validity of the test and medical ethos. Furthermore, scenarios that put examinees at risk of making critical errors or unnecessarily coming into contact with infectious diseases must be avoided. Lastly, it is impossible to guarantee the availability of enough RPs who present the state of pathology required by a predefined exam blueprint.

By being able to cover the whole range of assessment objectives, however, SPs can facilitate most of the above-mentioned situations in an exam [44]. SPs, themselves healthy, can endure much more than RPs, including intimate examinations, if selected and prepared accordingly. In the case of invasive procedures during an exam, either a simulator could be used or it could be combined with an SP in a hybrid solution. Hybrid simulations combining an SP with additional simulation modalities creates new possibilities [80], such as use of modern medical moulage to simulate almost any skin condition [81] or technology that allows for auscultation of pathological murmurs on a healthy SP [9]. Skilled SPs can very accurately portray important clinical scenarios, e.g., acute respiratory infection combined with the monitoring of computer-simulated vital functions in an emergency room setting, psychiatric cases

or demonstration of pain, as mentioned above. An SP can learn a medical history from the case script and they can even be instructed on how to improvise if necessary.

The validity of SPs in assessment has been evidenced by research. For example, in a systematic review it was demonstrated that trained SPs are able to act realistically as so-called unannounced standardised patients [82] – in the 40 studies reviewed, all but one reported data on unannounced SP visits, and the SP detection rate was mostly between 0% and 20%. It has also been shown that there is a correlation between low trainee communication scores from assessments using SPs and subsequent patient complaints in later professional work [83]. Even the use of healthy SPs under the age of 18 has become, with the careful monitoring of their needs and well-being given highest priority, well-established in assessment [84].

Overall, we see the value of RPs in *formative* clinical assessment throughout medical education, such as in workplace-based assessments. There is a multitude of methods for formative assessment which can enrich clinical training and provide structured feedback in ways that respect the needs of both patients and students [85]. A new implementation guide for the Swiss national learning objectives, PROFILES [39], stresses the importance of implementing multiple approaches for assessment and employing various patient representations in order to test students' competencies [86].

#### Assessing higher order skills with SPs and RPs

It is often stated that SPs do not support an overall approach to assessing the care provided to the patient, that they only allow the assessment of a limited sample of competencies rather than higher order skills, such as clinical reasoning or differential diagnoses; and that when assessments are performed with SPs, students' deep knowledge and creativity are not rewarded [44, 77].

Assessment during an exam is indeed a "single case consultation" regardless of whether SPs or RPs are employed, so any structured exam has limitations compared to dealing with real-time, long-term treatment. It is, however, feasible and valid to address higher order skills around differential diagnoses and clinical reasoning in structured medical assessments [74]. This is achieved, for example, in the US medical licensing exam, which provides unique information about students' achievements [87]. The implemented assessment task depends more on the assessment aim than on whether SPs or RPs are involved. In both cases, the patient representation must be deliberately selected for the particular competences to be assessed.

#### Immersion in the assessment situation when using SPs and RPs

It has been argued that students' expectations surrounding the test situation are an omnipresent issue. When students know that the exam is a simulation, they may not engage or immerse themselves fully, and may not behave as they would in a real situation, but instead simply do what is expected of them in the test [77]. There is work showing that immersion is reduced when students are aware of the simulation [88].

However, other studies reveal that simulated settings can elicit a high degree of immersion, and also stress [89, 90].

Consequently, debriefing is a standard educational measure after simulation training activities [91, 92]. Thus, it appears that a high degree of immersion is not solely a result of authenticity, but can also be induced by the task and the setting in which an SP, or even a mannequin, is involved.

#### Resource intensity of using SPs and RPs

The use of SPs is often criticised for being resource intensive. Indeed, resources are needed to recruit, train and deploy SPs for OSCEs. In Switzerland, all the main medical faculties maintain considerable numbers of SPs, between 100 and >200 per faculty, who are engaged on hourly rates when participating in teaching and assessment scenarios. In addition, clinical experts must be engaged to rate the students' performances during an exam. These circumstances make structured practical exams costly [44]. However, the recruitment of RPs would be similarly resource intensive for clinicians, who would be responsible for the recruitment of RPs according to a blueprint, as well as for the well-being of these patients during the exam.

Brown et al. [93] make the point that cost overviews of assessments are important but difficult to achieve. A full cost overview must also include the consequences of any medical errors [94]. In order to optimise authenticity and to reduce the resources needed to train SPs, the idea of using RPs instead in summative assessment has been explored [95]. In this study, student performance with RPs and SPs in 'long cases' was compared. It was calculated that 3.5 hours of testing time with 10 RPs would be needed in order to have a reliable test. The authors concluded: "long cases are therefore, in terms of reliability, no worse or better than OSCEs (with SPs) in assessing clinical competence" (p. 321). In a separate review on the use of SPs portraying psychiatric cases, the conclusion was that SPs are expensive and therefore should only be used when needed to produce a reliable test [96], but that they offered the opportunity to "significantly upgrade the validity, reliability, and fairness of high-stakes examinations of psychiatry..." (p. 191). The challenge of balancing the costs of OSCEs using SPs and assessments using RPs remains an issue, one in which the cost balance cannot include only the costs of running the exam.

#### Discussion

In this paper, we have compiled an overview of the most relevant patient representations and their use in both teaching and assessment of clinical skills. The three aspects covered in this review aim to present the state of the art of research on the use (Q1), selection (Q2) and merits or otherwise (Q3) of patient representations.

The strengths of this review are first, that it presents what is to our knowledge the only overview of patient representation use in medical education that has been conducted in this way. We found no existing reviews or papers directly reflecting the research questions we have addressed herein. Second, as a team, we have the necessary combination of both practical expertise and knowledge to perform a critical review of the existing literature in this area. We searched for and analysed selected papers that specifically relate to our questions. Third, we sought to find and detail applicable answers to our questions, information that

educators and assessors who are not experts on medical education would benefit from (Q1, what do they need to know?). The feedback from clinicians helped us to improve the paper in this respect. We sought to share information (Q2) about how to choose a suitable patient representation in medical teaching and assessment. This is a multifaceted issue, where didactical considerations for teaching and special methodological requirements for summative assessment must be met. Furthermore, our analysis suggests that there are no absolute rules for the optimal selection of a particular patient representation, and that the levels of authenticity, validity and feasibility needed must be evaluated for each learning and assessment setting. We also sought to summarise the debate on the advantages and limitations of SPs and RPs (Q3). This debate, as well as the findings throughout the paper, indicate that RPs are often perceived as the ideal for teaching and assessment. The many limitations of RPs, however, make the need for alternative forms of representations obvious. An SP has the highest authenticity in scenarios where an RP is not an option, and our analysis shows that SPs are the first choice for summative exams and a good alternative to RPs, despite their known limitations.

It was challenging to find evidence on whether particular patient representations improved a physician's performance later in life. This is undoubtedly due to the many factors influencing the causality between educational interventions, later clinical practice and patient care. We found one study [83] reporting the prognostic validity of communication scores in SP-based assessments for patient complaints to regulatory bodies. There is further evidence that OSCEs (commonly SP-based) have predictive value for later clinical work [97] and, separately, that reliable results [98] and predictive value [99] can be obtained by employing RPs in workplace-based settings (normally formative), too. Research directly comparing the use of SPs and RPs in summative assessment would be of interest, but we only found studies comparing them in *teaching*, favouring SPs [100] and reporting no difference [101].

There are many potential limitations of a literature review such as this. The three most obvious are: (i) it is possible that the authors were biased when selecting the papers; (ii) the authors may have missed relevant publications due to an unsystematic search; and (iii) the selected papers might have methodological flaws. Regarding the first issue, we sought to minimise author bias by searching specifically for papers reporting both advantages and disadvantages of various patient representations, as well as differentiated views of their use, especially with regard to RPs and SPs. We performed general searches on the topic, which recovered papers published in a variety of journals and written by authors with a variety of professional experiences (42% of the cited journals have a clinical profile). The second limitation is inherent to "unsystematic" reviews, and we felt it was overcome by adhering to the principles described above. Furthermore, due to the experience of the author team, we could readily establish when saturation was reached on the available papers relating to a particular issue. To address the final point, papers found to be methodologically weak on closer inspection were excluded. Our experience in both qualitative and quantitative empirical research was useful for this validation. However,

given the long history of research on this topic, some papers reporting early experiences and opinions were included to ensure that a balanced picture was presented.

## Conclusions

In this paper, we highlight the arguments that can influence the choice of patient representation for medical teaching and summative assessment. Learning is best supported when students can build knowledge and competency in small steps: from easy to complex, from low to high fidelity. Such support requires the structured use of various patient representations. We conclude that there is a long tradition of using various patient representations in the context of medical teaching, but that opinions relating to this in the context of summative assessment tend to be polarised between the pros and cons of SPs and RPs. Considering all arguments together, we conclude that SPs are the first choice for summative assessment.

In our opinion, the most relevant issue is not the choosing of either RPs or SPs. Rather, it is ensuring the selection of the most suitable patient representation. For more than 50 years, SPs have been established as a reliable, valid and feasible patient representation in summative exams. We propose that a broad range of possible patient representations can be utilised for summative exams: RPs are as much a patient representation as a case narrative, to highlight the two extremes. We make the point that the debate on finding the optimal patient representation in summative assessment should follow the principles of medical teaching. The selected patient representation must (i) be aligned with the learning goals and the competencies of the learner, (ii) correspond to the requirements of the clinical task itself, (iii) provide the necessary authenticity in order to realise (i) and (ii) while also respecting ethical considerations, and (iv) taking the above aspects into consideration, be feasible with the available resources. In addition, summative assessment must comply with methodical requirements, i.e. the exams must be reliable (reproducible), valid (fulfilling the test intention), feasible (practically and economically) and authentic to the extent demanded by the task.

Cases where RPs can be used in summative assessment will remain limited because many relevant clinical situations cannot, for ethical and methodological reasons, be incorporated into summative assessment, for example in an OSCE. Critical situations (e.g., emergencies, communicating fatal diagnoses) must instead be addressed by exams using SPs in combination with other patient representations (e.g., hybrid simulations), which have been proven a reliable and valid alternative to RPs. For high-stakes summative assessments, such as an OSCE at the end of study or even the Swiss federal licensing examination, we conclude that to address the criteria for good assessment including reliability, validity, authenticity and feasibility, an exam with only SPs is a good fit for the assessment purpose.

Overall, students should be assessed according to the degree of competence they are supposed to have attained and are allowed to exercise at a particular level. Therefore, compromises must be made with regard to absolute authenticity in summative exams at particular educational levels: they must be both as authentic as possible and as



straightforward as necessary. Nevertheless, RPs should be integrated for teaching and formative means as often (and as early) as possible, given that students can be prepared for such interactions by means of other patient representations. In situations where the obstacles to using RPs in summative exams can be overcome, we suggest focusing particularly on the added value of RPs, and pursuing a process of documenting and investigating feasible, ethical, reliable and valid assessment scenarios.

Decisions as to whether RPs would be suitable in a potential separate summative assessment track must be guided by additional research in this area. Also, studies on how patient representations in medical training and assessment affect future clinical work seem merited. However, the feasibility of such studies is limited due to challenges surrounding the control of chains of cause and effect.

#### Disclosure statement

No financial support and no other potential conflict of interest relevant to this article were reported.

#### References

- Bizzocchi J, Schell R. Rich-narrative case study for online PBL in medical education. *Acad Med.* 2009;84(10):1412–8. doi: <http://dx.doi.org/10.1097/ACM.0b013e3181b6ead0>. PubMed.
- Huwendiek S, De leng BA, Zary N, Fischer MR, Ruiz JG, Ellaway R. Towards a typology of virtual patients. *Med Teach.* 2009;31(8):743–8. doi: <http://dx.doi.org/10.1080/01421590903124708>. PubMed.
- Lim KHA, Loo ZY, Goldie SJ, Adams JW, McMenamin PG. Use of 3D printed models in medical education: A randomized control trial comparing 3D prints versus cadaveric materials for learning external cardiac anatomy. *Anat Sci Educ.* 2016;9(3):213–21. doi: <http://dx.doi.org/10.1002/ase.1573>. PubMed.
- Nicholson DT, Chalk C, Funnell WRJ, Daniel SJ. Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. *Med Educ.* 2006;40(11):1081–7. doi: <http://dx.doi.org/10.1111/j.1365-2929.2006.02611.x>. PubMed.
- Groscurth P, Eggli P, Kapfhammer J, Rager G, Hornung JP, Fasel JD. Gross anatomy in the surgical curriculum in Switzerland: improved cadaver preservation, anatomical models, and course development. *Anat Rec.* 2001;265(6):254–6. doi: <http://dx.doi.org/10.1002/ar.10030>. PubMed.
- Epps C, White ML, Tofil N. Mannequin Based Simulators. In: Levine AI, DeMaria S, Schwartz AD, Sim AJ, editors. *The Comprehensive Textbook of Healthcare Simulation*. 7th ed. New York, NY: Springer New York; 2013. pp 209–32.
- Lindsay Miller J, Avery MD, Larson K, Woll A, VonAchen A, Mortenson A. Emergency birth hybrid simulation with standardized patients in midwifery education: implementation and evaluation. *J Midwifery Womens Health.* 2015;60(3):298–303. doi: <http://dx.doi.org/10.1111/jmwh.12276>. PubMed.
- hopkinsmedicine.org. Maryland: The Johns Hopkins University, The Johns Hopkins Hospital, and Johns Hopkins Health System; 2020. Available from: <https://www.hopkinsmedicine.org>.
- Friederichs H, Weissenstein A, Ligges S, Möller D, Becker JC, Marschall B. Combining simulated patients and simulators: pilot study of hybrid simulation in teaching cardiac auscultation. *Adv Physiol Educ.* 2014;38(4):343–7. doi: <http://dx.doi.org/10.1152/advan.00039.2013>. PubMed.
- Goulart JM, Dusza S, Pillsbury A, Soriano RP, Halpern AC, Marghoob AA. Recognition of melanoma: a dermatologic clinical competency in medical student education. *J Am Acad Dermatol.* 2012;67(4):606–11. doi: <http://dx.doi.org/10.1016/j.jaad.2011.12.007>. PubMed.
- Daepfen J-B, Fortini C, Bertholet N, Bonvin R, Berney A, Michaud P-A, et al. Training medical students to conduct motivational interviewing: a randomized controlled trial. *Patient Educ Couns.* 2012;87(3):313–8. doi: <http://dx.doi.org/10.1016/j.pec.2011.12.005>. PubMed.
- Vu N, Baroffio A, Huber P, Layat C, Gerbase M, Nendaz M. Assessing clinical competence: a pilot project to evaluate the feasibility of a standardized patient -- based practical examination as a component of the Swiss certification process. *Swiss Med Wkly.* 2006;136(25-26):392–9. doi: <http://dx.doi.org/10.4414/smw.2006.11353>. PubMed.
- Bokken L, Rethans J-J, Scherpbier AJJA, van der Vleuten CPM. Strengths and weaknesses of simulated and real patients in the teaching of skills to medical students: a review. *Simul Healthc.* 2008;3(3):161–9. doi: <http://dx.doi.org/10.1097/SIH.0b013e318182fc56>. PubMed.
- Gierk B, Harendza S. Patient selection for bedside teaching: inclusion and exclusion criteria used by teachers. *Med Educ.* 2012;46(2):228–33. doi: <http://dx.doi.org/10.1111/j.1365-2923.2011.04054.x>. PubMed.
- Biggs JB, Tang CS-K. *Teaching for quality learning at university*. 4th ed. Maidenhead, England: McGraw-Hill/Society for Research into Higher Education & Open University Press; 2011.
- Holt EA, Young C, Keetch J, Larsen S, Mollner B. The greatest learning return on your pedagogical investment: alignment, assessment or in-class instruction? *PLoS One.* 2015;10(9):e0137446. doi: <http://dx.doi.org/10.1371/journal.pone.0137446>. PubMed.
- Jha V, Quinton ND, Bekker HL, Roberts TE. Strategies and interventions for the involvement of real patients in medical education: a systematic review. *Med Educ.* 2009;43(1):10–20. doi: <http://dx.doi.org/10.1111/j.1365-2923.2008.03244.x>. PubMed.
- Beigzadeh A, Bahmanbijari B, Sharifpoor E, Rahimi M. Standardized patients versus simulated patients in medical education: are they the same or different. *J Emerg Pract Trauma.* 2016;2(1):25–8. doi: <http://dx.doi.org/10.15171/jept.2015.05>.
- Bokken L, Rethans J-J, van Heurn L, Duvivier R, Scherpbier A, van der Vleuten C. Students' views on the use of real patients and simulated patients in undergraduate medical education. *Acad Med.* 2009;84(7):958–63. doi: <http://dx.doi.org/10.1097/ACM.0b013e3181a814a3>. PubMed.
- Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J.* 2009;26(2):91–108. doi: <http://dx.doi.org/10.1111/j.1471-1842.2009.00848.x>. PubMed.
- Spencer J, Blackmore D, Heard S, McCrorie P, McHaffie D, Scherpbier A, et al. Patient-oriented learning: a review of the role of the patient in the education of medical students. *Med Educ.* 2000;34(10):851–7. doi: <http://dx.doi.org/10.1046/j.1365-2923.2000.00779.x>. PubMed.
- Littlewood S, Ypinazar V, Margolis SA, Scherpbier A, Spencer J, Dorman T. Early practical experience and the social responsiveness of clinical education: systematic review. *BMJ.* 2005;331(7513):387–91. doi: <http://dx.doi.org/10.1136/bmj.331.7513.387>. PubMed.
- Dorman T, Littlewood S, Margolis SA, Scherpbier A, Spencer J, Ypinazar V. How can experience in clinical and community settings contribute to early medical education? A BEME systematic review. *Med Teach.* 2006;28(1):3–18. doi: <http://dx.doi.org/10.1080/01421590500410971>. PubMed.
- Howe A, Dagley V, Hopayian K, Lillicrap M. Patient contact in the first year of basic medical training—feasible, educational, acceptable? *Med Teach.* 2007;29(2-3):237–45. doi: <http://dx.doi.org/10.1080/01421590701294356>. PubMed.
- Dammers J, Spencer J, Thomas M. Using real patients in problem-based learning: students' comments on the value of using real, as opposed to paper cases, in a problem-based learning module in general practice. *Med Educ.* 2001;35(1):27–34. doi: <http://dx.doi.org/10.1046/j.1365-2923.2001.00841.x>. PubMed.
- Waterbury JT. Refuting patients' obligations to clinical training: a critical analysis of the arguments for an obligation of patients to participate in the clinical education of medical students. *Med Educ.* 2001;35(3):286–94. doi: <http://dx.doi.org/10.1046/j.1365-2923.2001.00865.x>. PubMed.
- O'Flynn N, Spencer J, Jones R. Consent and confidentiality in teaching in general practice: survey of patients' views on presence of students. *BMJ.* 1997;315(7116):1142. doi: <http://dx.doi.org/10.1136/bmj.315.7116.1142>. PubMed.
- Dent JA. AMEE Guide No 26: clinical teaching in ambulatory care settings: making the most of learning opportunities with outpatients. *Med Teach.* 2005;27(4):302–15. doi: <http://dx.doi.org/10.1080/01421590500150999>. PubMed.
- Ziv A, Wolpe PR, Small SD, Glick S. Simulation-based medical education: an ethical imperative. *Acad Med.* 2003;78(8):783–8. doi: <http://dx.doi.org/10.1097/00001888-200308000-00006>. PubMed.
- Peyton JWR. *Teaching and learning in medical practice*. Heronsgate Rickmansworth, Herts: Manticore Europe Ltd.; 1998.
- Young JQ, Van Merriënboer J, Durning S, Ten Cate O. Cognitive Load Theory: implications for medical education: AMEE Guide No. 86. *Med Teach.* 2014;36(5):371–84. doi: <http://dx.doi.org/10.3109/0142159X.2014.889290>. PubMed.
- Fraser K, Ma I, Teteris E, Baxter H, Wright B, McLaughlin K. Emotion, cognitive load and learning outcomes during simulation training. *Med Educ.* 2012;46(11):1055–62. doi: <http://dx.doi.org/10.1111/j.1365-2923.2012.04355.x>. PubMed.

- 33 Mancinetti M, Guttormsen S, Berendonk C. Cognitive load in internal medicine: What every clinical teacher should know about cognitive load theory. *Eur J Intern Med.* 2019;60:4–8. doi: <http://dx.doi.org/10.1016/j.ejim.2018.08.013>. PubMed.
- 34 Huwendiek S, Reichert F, Bosse HM, de Leng BA, van der Vleuten CPM, Haag M, et al. Design principles for virtual patients: a focus group study among students. *Med Educ.* 2009;43(6):580–8. doi: <http://dx.doi.org/10.1111/j.1365-2923.2009.03369.x>. PubMed.
- 35 Huwendiek S, Duncker C, Reichert F, De Leng BA, Dolmans D, van der Vleuten CPM, et al. Learner preferences regarding integrating, sequencing and aligning virtual patients with other activities in the undergraduate medical curriculum: A focus group study. *Med Teach.* 2013;35(11):920–9. doi: <http://dx.doi.org/10.3109/0142159X.2013.826790>. PubMed.
- 36 Kotsis SV, Chung KC. Application of the “see one, do one, teach one” concept in surgical training. *Plast Reconstr Surg.* 2013;131(5):1194–201. doi: <http://dx.doi.org/10.1097/PRS.0b013e318287a0b3>. PubMed.
- 37 Speirs C, Brazil V. See one, do one, teach one: Is it enough? *No. Emerg Med Australas.* 2018;30(1):109–10. doi: <http://dx.doi.org/10.1111/1742-6723.12927>. PubMed.
- 38 Bundesgesetz über die universitären Medizinalberufe: MedBG. 2007. <https://www.admin.ch/opc/de/classified-compilation/20040265/index.html>.
- 39 Michaud PA, Jucker-Kupper P, and members of the Profiles working group. PROFILES; Principal Objectives and Framework for Integrated Learning and Education in Switzerland. Bern: Joint Commission of the Swiss Medical Schools; 2017. <https://www.profilesmed.ch>.
- 40 Frank JR, Snell L, Sherbino J, eds. CanMEDS 2015 Physician Competency Framework. Ottawa: Royal College of Physicians and Surgeons of Canada; 2015.
- 41 Thistlethwaite JE, Davies D, Ekeocha S, Kidd JM, MacDougall C, Matthews P, et al. The effectiveness of case-based learning in health professional education: A BEME systematic review: BEME Guide No. 23. *Med Teach.* 2012;34(6):e421–44. doi: <http://dx.doi.org/10.3109/0142159X.2012.680939>. PubMed.
- 42 Yammine K, Violato C. The effectiveness of physical models in teaching anatomy: a meta-analysis of comparative studies. *Adv Health Sci Educ Theory Pract.* 2016;21(4):883–95. doi: <http://dx.doi.org/10.1007/s10459-015-9644-7>. PubMed.
- 43 Collins JP, Harden JM. AMEE Medical Education Guide No. 13: real patients, simulated patients and simulators in clinical examinations. *Med Teach.* 1998;20(6):508–21. doi: <http://dx.doi.org/10.1080/01421599880210>.
- 44 Cleland JA, Abe K, Rethans J-J. The use of simulated patients in medical education: AMEE Guide No 42. *Med Teach.* 2009;31(6):477–86. doi: <http://dx.doi.org/10.1080/01421590903002821>. PubMed.
- 45 Jones R, Higgs R, de Angelis C, Prideaux D. Changing face of medical curricula. *Lancet.* 2001;357(9257):699–703. doi: [http://dx.doi.org/10.1016/S0140-6736\(00\)04134-9](http://dx.doi.org/10.1016/S0140-6736(00)04134-9). PubMed.
- 46 Trachsel S, Schaufelberger M, Feller S, Küng L, Frey P, Guttormsen Schär S. Evaluation eines neuen Mentoring-Programms für Medizinstudierende in der hausärztlichen Grundversorgung: Erfahrungen von Studierenden und Lehrärzten. *GMS J Med Educ.* 2010;27(3):Doc42. doi: <http://dx.doi.org/10.3205/zma000679>.
- 47 Barrows HS. Simulated (Standardized) Patients and other Human Simulations. Chapel Hill, N.C.: Health Sciences Consortium; 1987.
- 48 Johnston JL, Lundy G, McCullough M, Gormley GJ. The view from over there: reframing the OSCE through the experience of standardised patient raters. *Med Educ.* 2013;47(9):899–909. doi: <http://dx.doi.org/10.1111/medu.12243>. PubMed.
- 49 Schmitz FM, Schnabel KP, Stricker D, Fischer MR, Guttormsen S. Learning communication from erroneous video-based examples: A double-blind randomised controlled trial. *Patient Educ Couns.* 2017;100(6):1203–12. doi: <http://dx.doi.org/10.1016/j.pec.2017.01.016>. PubMed.
- 50 Schmitz FM, Schnabel KP, Bauer D, Bachmann C, Woermann U, Guttormsen S. The learning effects of different presentations of worked examples on medical students’ breaking-bad-news skills: A randomized and blinded field trial. *Patient Educ Couns.* 2018;101(8):1439–51. doi: <http://dx.doi.org/10.1016/j.pec.2018.02.013>. PubMed.
- 51 Smee S. Skill based assessment. *BMJ.* 2003;326(7391):703–6. doi: <http://dx.doi.org/10.1136/bmj.326.7391.703>. PubMed.
- 52 Standards for educational and psychological testing. Washington, DC: Amer Educ Res Assoc; 2014. 230.
- 53 Epstein RM. Assessment in medical education. *N Engl J Med.* 2007;356(4):387–96. doi: <http://dx.doi.org/10.1056/NEJMra054784>. PubMed.
- 54 Ten Cate O, Hart D, Ankel F, Busari J, Englander R, Glasgow N, et al.; International Competency-Based Medical Education Collaborators. Entrustment Decision Making in Clinical Training. *Acad Med.* 2016;91(2):191–8. doi: <http://dx.doi.org/10.1097/ACM.0000000000001044>. PubMed.
- 55 Ten Cate O. Entrustment as Assessment: Recognizing the Ability, the Right, and the Duty to Act. *J Grad Med Educ.* 2016;8(2):261–2. doi: <http://dx.doi.org/10.4300/JGME-D-16-00097.1>. PubMed.
- 56 van der Vleuten CP, Schuwirth LW, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, et al. A model for programmatic assessment fit for purpose. *Med Teach.* 2012;34(3):205–14. doi: <http://dx.doi.org/10.3109/0142159X.2012.652239>. PubMed.
- 57 Mak-van der Vossen M. ‘Failure to fail’: the teacher’s dilemma revisited. *Med Educ.* 2019;53(2):108–10. doi: <http://dx.doi.org/10.1111/medu.13772>. PubMed.
- 58 Yepes-Rios M, Dudek N, Duboyce R, Curtis J, Allard RJ, Varpio L. The failure to fail underperforming trainees in health professions education: A BEME systematic review: BEME Guide No. 42. *Med Teach.* 2016;38(11):1092–9. doi: <http://dx.doi.org/10.1080/0142159X.2016.1215414>. PubMed.
- 59 Norcini J, Anderson B, Bollela V, Burch V, Costa MJ, Duvivier R, et al. Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach.* 2011;33(3):206–14. doi: <http://dx.doi.org/10.3109/0142159X.2011.551559>. PubMed.
- 60 Guttormsen S, Perruchoud A. Die eidgenössische Prüfung ist keine Lehrveranstaltung. *Schweiz Arzteztg.* 2016;97(18–19):679. doi: <http://dx.doi.org/10.4414/saez.2016.04676>.
- 61 Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ.* 2004;38(9):1006–12. doi: <http://dx.doi.org/10.1111/j.1365-2929.2004.01932.x>. PubMed.
- 62 Barrows HS, Abrahamson S. The programmed patient: a technique for appraising student performance in clinical neurology. *J Med Educ.* 1964;39(8):802–5. PubMed.
- 63 Vu NV, Barrows HS. Use of standardized patients in clinical assessments: recent developments and measurement findings. *Educ Res.* 1994;23(3):23–30. doi: <http://dx.doi.org/10.3102/0013189X023003023>.
- 64 Newlin-Canzone ET, Scerbo MW, Gliva-McConvey G, Wallace AM. The cognitive demands of standardized patients: understanding limitations in attention and working memory with the decoding of nonverbal behavior during improvisations. *Simul Healthc.* 2013;8(4):207–14. doi: <http://dx.doi.org/10.1097/SIH.0b013e31828b419e>. PubMed.
- 65 Guttormsen S, Beyeler C, Bonvin R, Feller S, Schirlo C, Schnabel K, et al. The new licencing examination for human medicine: from concept to implementation. *Swiss Med Wkly.* 2013;143:w13897. doi: <http://dx.doi.org/10.4414/smw.2013.13897>. PubMed.
- 66 Kane M. The Argument-based approach to validation. *School Psych Rev.* 2013;42(4):448–57. doi: <http://dx.doi.org/10.1080/02796015.2013.12087465>.
- 67 apa.org. Washington: American Psychological Association; 2020. Available from: <https://www.apa.org>.
- 68 Baig LA, Beran TN, Vallevand A, Baig ZA, Monroy-Cuadros M. Accuracy of portrayal by standardized patients: results from four OSCE stations conducted for high stakes examinations. *BMC Med Educ.* 2014;14(1):97. doi: <http://dx.doi.org/10.1186/1472-6920-14-97>. PubMed.
- 69 Sanko JS, Shekhter I, Gattamorta KA, Birnbach DJ. Development and psychometric analysis of a tool to evaluate confederates. *Clin Simul Nurs.* 2016;12(11):475–81. doi: <http://dx.doi.org/10.1016/j.ecns.2016.07.006>.
- 70 Smith C, O’Byrne C. Using an exam-readiness tool to ensure quality of standardized/simulated patient role portrayal in high-stakes simulation assessments. *Clear Exam Rev.* 2017;27(1).
- 71 Wind LA, Van Dalen J, Muijtjens AMM, Rethans J-J. Assessing simulated patients in an educational setting: the MaSP (Maastricht Assessment of Simulated Patients). *Med Educ.* 2004;38(1):39–44. doi: <http://dx.doi.org/10.1111/j.1365-2923.2004.01686.x>. PubMed.
- 72 Preusche I, Schmidts M, Wagner-Menghin M. Twelve tips for designing and implementing a structured rater training in OSCEs. *Med Teach.* 2012;34(5):368–72. doi: <http://dx.doi.org/10.3109/0142159X.2012.652705>. PubMed.
- 73 Tavares W, Ginsburg S, Eva KW. Selecting and simplifying: rater performance and behavior when considering multiple competencies. *Teach Learn Med.* 2016;28(1):41–51. doi: <http://dx.doi.org/10.1080/10401334.2015.1107489>. PubMed.
- 74 Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *BMJ.* 1975;1(5955):447–51. doi: <http://dx.doi.org/10.1136/bmj.1.5955.447>. PubMed.

- 75 Sayer M, Bowman D, Evans D, Wessier A, Wood D. Use of patients in professional medical examinations: current UK practice and the ethico-legal implications for medical education. *BMJ*. 2002;324(7334):404–7. doi: <http://dx.doi.org/10.1136/bmj.324.7334.404>. PubMed.
- 76 Newble DI. Using real patients in professional medical exams. Suggestions would make examinations with real patients impractical. *BMJ*. 2002;324(7347):1217a. doi: <http://dx.doi.org/10.1136/bmj.324.7347.1217a>. PubMed.
- 77 Harden RM, Lilley P, Patricio M. *The Definitive Guide to the OSCE: The Objective Structured Clinical Examination as a performance assessment*. 1st ed. Elsevier, New York: Churchill Livingstone; 2015.
- 78 Hill AE, Davidson BJ, Theodoros DG. A review of standardized patients in clinical education: Implications for speech-language pathology programs. *Int J Speech Lang Pathol*. 2010;12(3):259–70. doi: <http://dx.doi.org/10.3109/17549500903082445>. PubMed.
- 79 Barrows HS. An overview of the uses of standardized patients for teaching and evaluating clinical skills. *AAMC. Acad Med*. 1993;68(6):443–51, discussion 451–3. doi: <http://dx.doi.org/10.1097/00001888-199306000-00002>. PubMed.
- 80 Girzadas DV, Jr, Antonis MS, Zerth H, Lambert M, Clay L, Bose S, et al. Hybrid simulation combining a high fidelity scenario with a pelvic ultrasound task trainer enhances the training and evaluation of endovaginal ultrasound skills. *Acad Emerg Med*. 2009;16(5):429–35. doi: <http://dx.doi.org/10.1111/j.1553-2712.2009.00399.x>. PubMed.
- 81 Garg A, Haley H-L, Hatem D. Modern moulage: evaluating the use of 3-dimensional prosthetic mimics in a dermatology teaching program for second-year medical students. *Arch Dermatol*. 2010;146(2):143–6. doi: <http://dx.doi.org/10.1001/archdermatol.2009.355>. PubMed.
- 82 Rethans J-J, Gorter S, Bokken L, Morrison L. Unannounced standardized patients in real practice: a systematic literature review. *Med Educ*. 2007;41(6):537–49. doi: <http://dx.doi.org/10.1111/j.1365-2929.2006.02689.x>. PubMed.
- 83 Tamblin R, Abrahamowicz M, Dauphinee D, Wenghofer E, Jacques A, Klass D, et al. Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. *JAMA*. 2007;298(9):993–1001. doi: <http://dx.doi.org/10.1001/jama.298.9.993>. PubMed.
- 84 Gamble A, Bearman M, Nestel D. A systematic review: Children & Adolescents as simulated patients in health professional education. *Adv Simul (Lond)*. 2016;1(1):1. doi: <http://dx.doi.org/10.1186/s41077-015-0003-9>. PubMed.
- 85 Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Med Teach*. 2007;29(9-10):855–71. doi: <http://dx.doi.org/10.1080/01421590701775453>. PubMed.
- 86 Sohrmann M, Berendonk C, Nendaz M, Bonvin R; Swiss Working Group For Profiles Implementation. Nationwide introduction of a new competency framework for undergraduate medical curricula: a collaborative approach. *Swiss Med Wkly*. 2020;150:w20201. doi: <http://dx.doi.org/10.4414/smw.2020.20201>. PubMed.
- 87 Yudkowsky R, Park YS, Hyderi A, Bordage G. Characteristics and implications of diagnostic justification scores based on the new patient note format of the USMLE Step 2 CS Exam. *Acad Med*. 2015;90(11, Suppl):S56–62. doi: <http://dx.doi.org/10.1097/ACM.0000000000000900>. PubMed.
- 88 Wright MC, Taekman JM, Endsley MR. Objective measures of situation awareness in a simulated medical environment. *Qual Saf Health Care*. 2004;13(Suppl 1):i65–71. doi: <http://dx.doi.org/10.1136/qshe.2004.009951>. PubMed.
- 89 Harder N, Lemoine J, Harwood R. Psychological outcomes of debriefing healthcare providers who experience expected and unexpected patient death in clinical or simulation experiences: A scoping review. *J Clin Nurs*. 2020;29(3-4):330–46. doi: <http://dx.doi.org/10.1111/jocn.15085>. PubMed.
- 90 Lilot M, Evain J-N, Bauer C, Cejka J-C, Faure A, Balança B, et al. Relaxation before debriefing during high-fidelity simulation improves memory retention of residents at three months: a prospective randomized controlled study. *Anesthesiology*. 2018;128(3):638–49. doi: <http://dx.doi.org/10.1097/ALN.0000000000002045>. PubMed.
- 91 Gropper MA, Miller RD, Eriksson LI, Fleisher LA, Wiener-Kronish JP, Cohen NH, et al. *Miller's Anesthesia, 2-Volume Set E-Book*. Elsevier Health Sciences; 2019.
- 92 Kolbe M, Weiss M, Grote G, Knauth A, Dambach M, Spahn DR, et al. TeamGAINS: a tool for structured debriefings for simulation-based team trainings. *BMJ Qual Saf*. 2013;22(7):541–53. doi: <http://dx.doi.org/10.1136/bmjqs-2012-000917>. PubMed.
- 93 Brown C, Cleland J, Walsh K. The costs of medical education assessment. *Med Teach*. 2016;38(2):111–2. doi: <http://dx.doi.org/10.3109/0142159X.2015.1105946>. PubMed.
- 94 Nordgren LD, Johnson T, Kirschbaum M, Peterson ML. Medical errors: excess hospital costs and lengths of stay. *J Healthc Qual*. 2004;26(2):42–8, quiz 48–9. doi: <http://dx.doi.org/10.1111/j.1945-1474.2004.tb00484.x>. PubMed.
- 95 Wass V, Jones R, Van der Vleuten C. Standardized or real patients to test clinical competence? The long case revisited. *Med Educ*. 2001;35(4):321–5. doi: <http://dx.doi.org/10.1046/j.1365-2923.2001.00928.x>. PubMed.
- 96 Yudkowsky R. Should we use standardized patients instead of real patients for high-stakes exams in psychiatry? *Acad Psychiatry*. 2002;26(3):187–92. doi: <http://dx.doi.org/10.1176/appi.ap.26.3.187>. PubMed.
- 97 Wallenstein J, Heron S, Santen S, Shayne P, Ander D. A core competency-based objective structured clinical examination (OSCE) can predict future resident performance. *Acad Emerg Med*. 2010;17(Suppl 2):S67–71. doi: <http://dx.doi.org/10.1111/j.1553-2712.2010.00894.x>. PubMed.
- 98 Durning SJ, Cation LJ, Markert RJ, Pangaro LN. Assessing the reliability and validity of the mini-clinical evaluation exercise for internal medicine residency training. *Acad Med*. 2002;77(9):900–4. doi: <http://dx.doi.org/10.1097/00001888-200209000-00020>. PubMed.
- 99 Al Ansari A, Ali SK, Donnon T. The construct and criterion validity of the mini-CEX: a meta-analysis of the published research. *Acad Med*. 2013;88(3):413–20. doi: <http://dx.doi.org/10.1097/ACM.0b013e318280a953>. PubMed.
- 100 Schwartz VS, Rothpletz-Puglia P, Denmark R, Byham-Gray L. Comparison of standardized patients and real patients as an experiential teaching strategy in a nutrition counseling course for dietetic students. *Patient Educ Couns*. 2015;98(2):168–73. doi: <http://dx.doi.org/10.1016/j.pec.2014.11.009>. PubMed.
- 101 Carvalho IP, Pais VG, Silva FR, Martins R, Figueiredo-Braga M, Pedrosa R, et al. Teaching communication skills in clinical settings: comparing two applications of a comprehensive program with standardized and real patients. *BMC Med Educ*. 2014;14(1):92. doi: <http://dx.doi.org/10.1186/1472-6920-14-92>. PubMed.