The core business of medical education is to educate trainees so that they become competent professionals. To optimize this process, many educational interventions are developed and studied for their effectiveness. Such studies can be considered complex because their outcomes may be influenced by many variables, such as teacher quality, the setting, and trainees’ previous experiences and levels of commitment. Moreover, the number of interacting components within intervention studies is generally high.\(^1\)\(^-\)\(^3\) It is of great importance, therefore, that all relevant components be described explicitly\(^4\) so that educational interventions can be replicated, study outcomes can be understood and interpreted, and reliable conclusions can be drawn regarding the effectiveness of (aspects of) interventions. Furthermore, clear descriptions of educational interventions help medical educators translate successful interventions to local settings.\(^4\)\(^-\)\(^6\)

Studies published on educational interventions in the field of medical education often lack thorough descriptions, however.\(^7\)\(^-\)\(^11\) The main shortcomings that have been identified include incomplete information and a lack of uniformity in the descriptions of both the research methods and the educational interventions themselves.\(^9\)\(^,\)\(^12\)\(^,\)\(^13\) In a review focusing on how studies of educational interventions for evidence-based practice (EBP) describe these interventions, Phillips et al\(^13\) demonstrated that there were many inconsistencies in the descriptions. Because educational interventions affect medical education programs, which in turn affect health care, improving completeness and uniformity in the descriptions of educational interventions is important.\(^14\) To improve such descriptions, it would be helpful to know what problems currently exist. In this review, we therefore systematically analyzed published descriptions of educational interventions to identify potential gaps and limitations.

To identify the limitations of published intervention descriptions, we first needed to identify criteria for thorough descriptions. While several checklists had already been designed to report on educational interventions,\(^5\)\(^,\)\(^8\)\(^,\)\(^9\)\(^,\)\(^10\)\(^,\)\(^12\)\(^,\)\(^13\) none seemed fully appropriate for describing educational interventions in medical education because their scope was either...
content specific or too broad and also included the description of the study methodology. Therefore, based on the educational intervention literature and the aforementioned existing checklists, we established a new checklist of key items that authors should cover to describe interventions in medical education. Using this checklist, we performed a systematic review of medical education intervention studies, with a focus on interventions in postgraduate medical education.

Method
Development of the checklist
We began by developing a list of criteria to assess the descriptions of educational interventions in medical education in January 2015. This checklist, and hence this review, covered the description of the intervention only and did not deal with study design. Our list was based on the Guideline for Reporting Evidence-Based Practice Educational Interventions and Teaching (GREET) checklist developed by Phillips et al. We reformulated the GREET item that was specific to EBP (content/subject) to be applicable to educational interventions in general. We merged two items (planned changes, unplanned changes) and supplemented four items (theory, delivery, schedule, instructors). We incorporated four items (learning objectives, educational strategies, materials, environment) directly into our list. We did not include attendance because the scope of this item was covered by two others (delivery [from GREET] and participants [a new item, described below]).

On the basis of the literature, we identified additional items that were not covered by GREET. According to Morrison et al., Reznich and Anderson, and Windish et al., a first step when developing an intervention should be the assessment of learning needs of future participants. In addition, as discussed by Olson and by Windish et al., knowing the rationale behind the design of an educational intervention provides insight into the main decisions, and participants and context can influence the design. Therefore, we added the following criteria to our checklist: learning needs, intervention development process, context and settings, and participants. We also added items on evaluating the effect of and satisfaction with the intervention: assessment and satisfaction.

The final list of 17 items was completed in November 2015 and discussed during a workshop that month at a Dutch medical education conference. The workshop participants agreed on the importance of all criteria. Subsequently, in January 2016, a pilot survey was conducted, using a subset of 15 educational intervention papers included in our systematic review (described below). Two of the authors (J.M. and N.B.) individually scored each intervention description on all items and discussed their scores. All criteria except assessment proved possible to score. The studies included in the pilot showed that it was often unclear whether the assessment performed was for the study, for the intervention (or part of the intervention), or for both. Assessment was therefore not scored in the current review, although we included this item in the checklist given its overall importance. Furthermore, scoring was difficult on the materials and planned/unplanned changes criteria because it was not always clear which materials were used or whether the intervention was changed during the study. J.M. and N.B. ultimately reached agreement on how to score all criteria. Based on the discussions, a comprehensive description of each item and its scoring was developed.

Our final checklist distinguishes three stages—preparation, intervention, and evaluation—and includes 17 items (Table 1). Each item, except assessment, is scored 0, 1, or 2, depending on the completeness of the description: A score of 2 corresponds to a detailed description, a score of 1 corresponds to an unclear or partial description, and a score of 0 is given when information is lacking entirely. Some criteria contain multiple elements; for example, a complete or detailed description of the context and settings item for a score of 2 includes the educational institute, time of the year, year of the curriculum, and course area/specialty. If one of these elements is missing from the description, the article scores a 1 on this criterion. The scoring guidelines for each item are provided in the last column of Table 1.

Search methods
The MEDLINE (Ovid), Embase (Ovid), and ERIC (Ovid) databases were searched on March 22, 2016, for eligible English-language studies published since January 1, 2014. (The detailed search strategy used for each database is available in Supplemental Digital Appendix 1 at http://links.lww.com/ACADMED/A590.) To narrow the search, the scope was set to studies evaluating the effects of educational interventions during formal teaching (classroom teaching) in postgraduate medical education.

Inclusion and exclusion criteria
We included studies of educational interventions concerning postgraduate trainees (i.e., residents, interns) irrespective of specialty and year of education when residents accounted for at least 50% of the participants and the intervention and/or control session took place during classroom teaching.

Studies concerning physical examination or technical/clinical skills interventions (e.g., palpation instruction) were excluded, as were studies concerning interventions “on the job” and in bedside teaching, because these settings are less controlled than classroom settings and were therefore expected to be more difficult to describe. In addition, articles were excluded if the full text was unavailable in English or if they did not include an evaluation or study the effects of the intervention (i.e., reviews, abstracts, editorials, protocols, innovation reports, conference proceedings). Qualitative studies were excluded because they are generally exploratory and serve a different purpose than quantitative studies, which generally focus on effectiveness.

Study selection
N.B. and J.M. reviewed the titles and abstracts of records identified through the database searches against the inclusion and exclusion criteria. Their independent analysis of a subset (the first 200) of the publications yielded a 97% agreement rate. On the basis of this high level of agreement, the remaining publications were reviewed and selected by either J.M. or N.B. If the title and abstract met the inclusion criteria, or if they did not provide sufficient clarification to determine whether the study met the criteria, the full text was reviewed for eligibility.

The final set of included articles was scored using the 16 scored items on our checklist (Table 1), and data were collected on general study characteristics.
Table 1
Checklist for Thorough Descriptions of Educational Interventions in Medical Education, With Criteria and Scoring

| Criteria by stage | Descriptions                                                                 | Scoring                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------------|-----------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| **Preparation**   |                                                                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Learning needs    | Is the learning need that prompted the intervention described (e.g., gaps between desired and actual practice)? | Score 2: Description of the learning need is clear. It is clear why an intervention on this subject was needed and why it was designed specifically for the participants.  
Score 1: Only a description of the importance of the subject is given without a clear description of why the intervention was needed for these participants.  
Score 0: No information.  

Intervention development process | Is the process for intervention development described (including 1. the basis on which the intervention was developed, 2. who developed the intervention, and 3. how the material was developed)? | Score 2: Clear description of how the intervention was developed regarding the content/subject, and on what basis and by whom.  
Score 1: One or two of the three elements of the development process are described.  
Score 0: No information.  

| **Intervention**  |                                                                             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| Theory            | Is the theory on the educational strategies used in the intervention described (e.g., simulation, team-based learning, role-play, etc.)? | Score 2: The supporting theory regarding the educational strategy (TBL, simulation, web based) is presented.  
Score 1: Educational theory is briefly described (one or two sentences).  
Score 0: No information.  

Learning objectives | Are the learning objectives for all groups involved in the educational intervention described? | Score 2: The specific objectives of the intervention are clearly described.  
Score 1: No clear objectives are described, but a description of capabilities or general objectives is given (e.g., to improve knowledge about a particular subject).  
Score 0: No information.  

Content/subject    | Is the content and the subject of the intervention described? | Score 2: The content and subject are clearly described.  
Score 1: The subject is described, but there is no detailed information about the content, or information is lacking or too general.  
Score 0: No subject or content.  

Participants       | Are the participants (both intervention and controls) described (including 1. the level of knowledge/education of the participants, 2. the numbers, and 3. characteristics [dropouts, male/female, age, religion])? | Score 2: All three criteria are described: number of participants, level of knowledge/education, and other characteristics.  
Score 1: One or two of the three elements of the participants are described.  
Score 0: No information.  

Context and settings | Is the context of the educational intervention described (including 1. educational institute, 2. time of the year, 3. the year of the curriculum, and 4. course area/specialty)? | Score 2: Everything is described: institute, time of the year, the year of the curriculum, and course area.  
Score 1: Some information is lacking.  
Score 0: No information.  

Schedule           | Is the schedule of the intervention described (including 1. the number of sessions, 2. their frequency, 3. timing and duration, and 4. program schedule)? | Score 2: All four criteria are described.  
Score 1: The schedule is incomplete.  
Score 0: No schedule included.  

Materials          | Are the specific educational materials used in the educational intervention described (including 1. materials provided to the learners and 2. those used in the training of educational intervention providers)? | Score 2: Materials are described, and the description is clear.  
Score 1: There is a partial description, but detailed information is lacking.  
Score 0: Materials are not described.  

Educational strategies | Are the teaching/learning strategies used in the educational intervention described (e.g., tutorials, lectures, online modules)? | Score 2: Educational strategies are clear (e.g., interactive session, simulation, lecture).  
Score 1: Strategy remains unclear.  
Score 0: No information.  

Incentives         | Were any incentives or reimbursements related to the intervention provided to the learners, and are these described? Was the intervention mandatory? | Score 2: Whether the participants received an incentive is stated and, if so, the incentive is described. It is clear whether the intervention was mandatory.  
Score 1: Incentives remain unclear, but it is clear whether the intervention was mandatory.  
Score 0: No information.  

(Table continues)
Table 1 (Continued)

| Criteria by stage | Descriptions                                                                 | Scoring                                                                                       |
|-------------------|------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|
| Instructors/teachers* | Are the instructors involved in the educational intervention described (including 1. their professional discipline/expertise and 2. any specific training related to the educational intervention provided for instructors)? | Score 2: Information is available about instructors’ professional discipline/expertise, and whether training was offered (yes/no).<br>Score 1: Information is incomplete, and it remains unclear whether any specific training related to the educational intervention was provided for the instructors or what expertise they might have.<br>Score 0: No information. |
| Delivery          | Are the modes of delivery of the educational intervention described (including 1. group size and 2. the ratio of learners to instructors)? | Score 2: The ratio of learners to instructors is stated.<br>Score 1: The ratio of learners to instructors is not stated.<br>Score 0: No information. |
| Environment       | Are the teaching/learning locations (e.g., conference, university lecture theater, hospital ward, community) where the teaching/learning occurred described? | Score 2: There is a clear description of where (i.e., the environment) the intervention took place.<br>Score 1: The description of the environment is incomplete, or only an aspect of it is described.<br>Score 0: No information. |
| Assessment        | Is it clear whether assessment of the intervention took place (e.g., any type of exam, formative, summative or both, or other assessment details)? | No score: The distinction between assessment of the intervention and the study is difficult to make. |

**Evaluation**

| Planned and unplanned changes* | If the educational intervention required specific adaptation (before or during the intervention) for the learners, is this described (e.g., what, why, when, and how)? | Score 2: Whether the intervention required an adaptation is stated, and any required adaptation is described.<br>Score 1: Not clearly stated or described.<br>Score 0: No information. |
| Satisfaction   | Are the participants’ experiences and level of satisfaction with the intervention described? | Score 2: There is a clear description of the participants’ experiences and level of satisfaction with the intervention.<br>Score 1: There is a brief description of experience/satisfaction (one or two sentences).<br>Score 0: Not described. |

Abbreviation: TBL indicates team-based learning.

*This checklist was developed and used for a systematic review of descriptions of educational interventions in postgraduate medical education. For the distribution of scores across the 105 included studies, see Table 1. For scores for each study, criterion, and stage, see Supplemental Digital Appendix 3 at http://links.lww.com/ACADMED/A591.

*This item was reproduced or adapted from the GREET (Guideline for Reporting Evidence-Based Practice Educational Interventions and Teaching) checklist developed by Phillips et al115 (http://creativecommons.org/licenses/by/4.0). For an explanation of changes, see the Method section.

*The assessment criterion was not scored in the systematic review.

(journal, study design, country of origin, specialty of study participants). After the pilot survey (described above), the remaining articles were divided between J.M. and N.B. for scoring. Articles that referred to supplementary material (in an appendix, online, or in a prior publication) were scored on the basis of all information available, including the supplementary material when it could be obtained. J.M. and N.B. discussed scores frequently to maintain uniformity in scoring practice. Narrative details backing the scores were collected and are available from the corresponding author upon request. As noted above, the included studies were not assessed on the methodology used because the aim of this systematic review was to analyze descriptions of the interventions rather than methodology.

**Analysis**

All supporting quotations and other narrative data, associated scores, and general study characteristics were documented using Microsoft Excel 2011 (Microsoft Corp., Redmond, Washington). Descriptive statistics of study scores and characteristics were calculated using SPSS version 19.17 The analysis was performed from April 2016 to February 2017.

**Results**

The electronic database searches identified 1,826 records (Figure 1). The initial screening of titles and abstracts yielded 122 articles. After review of the full texts, 102 articles were included. Screening the reference lists of these 102 articles for additional eligible studies yielded 5 articles, resulting in a total of 105 articles selected for inclusion in the review.18–122

Throughout the screening process, articles were excluded for the following reasons, in order of incidence:

- No effect study or no evaluation of the intervention (n = 1,063; 58.2%)
- No classroom teaching (n = 331; 18.1%)
- No resident participants or < 50% resident participants (n = 206; 11.3%)
- Qualitative study (n = 50; 2.7%)
- Only title and abstract available (e.g., congress or poster abstracts) (n = 74; 4.1%)
Details on individual articles and the reasons for exclusion are available from the corresponding author upon request.

Characteristics of included studies

The 105 included articles were published in 74 journals, the most prevalent of which were the Journal of Graduate Medical Education (8.6%; n = 9) and Academic Psychiatry (6.7%; n = 7). The studies were conducted in 17 countries, most often in the United States (61.0%; n = 64). The interventions covered 17 specialties; the most frequent were internal medicine (14.3%, n = 15), family medicine (12.4%, n = 13), emergency medicine (10.5%, n = 11), and anesthesiology (10.5%, n = 11). Some studies did not report the specialty of the participants clearly (13.3%, n = 14), and some studies included trainees from multiple specialties (5.7%, n = 6). Almost half of the studies evaluated the intervention by using a pre- and posttest design for one group (48.6%; n = 51). The study characteristics are summarized in Supplemental Digital Appendix 2 at http://links.lww.com/ACADMED/A590.

Assessment of educational innovation descriptions: Scores on the checklist

An overview of the scores for each item, stage, and article is provided in Supplemental Digital Appendix 3 at http://links.lww.com/ACADMED/A591. The distribution of scores for each criterion for the included articles is presented in Table 2. Good practice examples for each item are provided in Supplemental Digital Appendix 4 at http://links.lww.com/ACADMED/A590. Each description of an educational intervention could score between 0 and 32 points based on the criteria checklist (Table 1). None of the articles described all of the criteria in detail. The lowest score of the articles we analyzed was 8 (5 studies29,30,38,61,68,101), and the highest score was 25 (1 study50). The mean score for the 105 included articles was 15.9 (SD 4.1). The range of scores was as follows: 12 articles (11.4%) scored 8–10, 80 articles (76.2%) scored 11–20, and 13 articles (12.4%) scored 21–25.

Twenty-four articles referred to additional material (in an appendix, online, or in a prior publication). We could not find the additional material for 5 articles (mean score: 13.8 [SD 3.1])18,19,44,48,83 For the remaining 19 articles,22,23,25,30,39,41,50,51,74,78,81,91–93,98,104,110,116,119 scoring included the supplementary material. The mean score of these 19 articles was 18.3 (SD 3.7; range 13–25), significantly higher than the mean score of 15.5 (SD 4.0) for the 81 articles without additional information (P = .007; 95% CI on the observed difference = 0.8–4.8).

Preparation (2 criteria). Learning needs were reported in all articles, with 78.1% (n = 82) describing them in detail (score 2). However, 21.9% (n = 23) failed to describe why the intervention was designed specifically for the participants (score 1).

The intervention development process was described in detail by 26.6% (n = 30) of the studies (score 2), as in the following example from Chee et al33:

We designed a teaching skills curriculum for the HMS [Harvard Medical School] Residency Training Program in Ophthalmology following Kern’s 6-step approach to curriculum development. These steps include problem identification, needs assessment of targeted learners, establishment of goals and objectives, design and implementation of educational strategies, and program evaluation.

Intervention (12 criteria). Which educational strategy was applied during the intervention was reported by 79.0% (n = 83) of the studies (score 2). However, 21.4% (n = 55) failed to mention the educational theory behind the strategy (score 0). An illustrative example of a study that did support the applied strategy with theoretical substantiation is Daly et al50:

Priori studies have found that narrow-band imaging (NBI) can be taught to physicians by in-person training and web-based program.
Learning objectives were reported in detail by 19.0% (n = 20) of the studies (score 2), but 43.8% (n = 46) did not report any objectives (of the intervention) at all (score 0). The content/subject of the intervention was described clearly in 77.1% (n = 81; score 2). The number of participants was reported in all studies, except 1. Some information on the context and setting (i.e., the institution, time of the year, year of the curriculum, and course area/specialty) was reported by almost all studies (98.1%, n = 103). However, 65.7% (n = 69) did not describe all of these elements (score 1). A complete description of the schedule (i.e., number of sessions, frequency, timing and duration, and program schedule) was given in 25.7% (n = 27) of the studies (score 2); however, 67.6% (n = 71) did not describe all of these elements (score 1), and 6.7% (n = 7) reported no information on the schedule at all (score 0).

With regard to course materials, 24.8% (n = 26) of the studies did not describe any materials (score 0), and detailed information was missing from another 37.1% (n = 39; score 1). The remaining 38.1% (n = 40) described their materials in more detail (score 2), such as in this example from Haspel et al:

PowerPoint lecture. Participants were given access to the activities at the workshop electronically (initially Google Docs and later Google Forms). We have also made all teaching materials and an instructor handbook available online at no cost and are planning train-the-trainer sessions to facilitate broader dissemination for local implementation.

Whether the participants received any incentives was not described in 74.3% (n = 78) of the studies (score 0). Whether the intervention was mandatory was mentioned in 16.2% (n = 17; score 1). Only 9.5% (n = 10) of the studies reported whether participants received some other kind of incentive (score 2), as in this example from Woodworth et al:

The only reward that subjects received for study participation was a DVD containing the educational video and interactive simulation, which was given to them at the conclusion of the study.

Information about the instructors/teachers was reported in 18.1% (n = 19) of the studies in detail (score 2), but 28.6% (n = 30) did not report any information about the instructors at all (score 0). In more than half of the studies (53.3%, n = 56), information was missing on the instructors’ backgrounds or whether they received any specific training for the intervention (score 1).

Information on delivery (i.e., group size and the ratio of learners to instructors) was described in detail in 22.9% (n = 24) of the studies (score 2), but not at all in 42.9% (n = 45; score 0). The environment was not described in 84.8% (n = 89) of the studies (score 0). Only 5.7% (n = 6) of the studies reported where the intervention took place (score 2), as in this example from Sawatsky et al:

Lectures are given by the same faculty member twice at two separate locations, a university-based hospital and a VA [Veterans Affairs] hospital. Both locations are set up with tables in rows, with several chairs at each table.

### Evaluation (2 criteria).

**Planned/unplanned changes.** Planned/unplanned changes, or whether the intervention required some kind of adaption before or during the intervention, were not reported by 82.9% (n = 87) of the studies (score 2), but not at all in 42.9% (n = 45; score 0). The environment was not described in 84.8% (n = 89) of the studies (score 0). Only 5.7% (n = 6) of the studies reported where the intervention took place (score 2), as in this example from Sawatsky et al:

Lectures are given by the same faculty member twice at two separate locations, a university-based hospital and a VA [Veterans Affairs] hospital. Both locations are set up with tables in rows, with several chairs at each table.

**Satisfaction.** Whether the participants received any incentives was not described in 74.3% (n = 78) of the studies (score 0). Whether the intervention was mandatory was mentioned in 16.2% (n = 17; score 1). Only 9.5% (n = 10) of the studies reported whether participants received some other kind of incentive (score 2), as in this example from Woodworth et al:

The only reward that subjects received for study participation was a DVD containing the educational video and interactive simulation, which was given to them at the conclusion of the study.

Information about the instructors/teachers was reported in 18.1% (n = 19) of the studies in detail (score 2), but 28.6% (n = 30) did not report any information about the instructors at all (score 0). In more than half of the studies (53.3%, n = 56), information was missing on the instructors’ backgrounds or whether they received any specific training for the intervention (score 1).

Information on delivery (i.e., group size and the ratio of learners to instructors) was described in detail in 22.9% (n = 24) of the studies (score 2), but not at all in 42.9% (n = 45; score 0). The environment was not described in 84.8% (n = 89) of the studies (score 0). Only 5.7% (n = 6) of the studies reported where the intervention took place (score 2), as in this example from Sawatsky et al:

Lectures are given by the same faculty member twice at two separate locations, a university-based hospital and a VA [Veterans Affairs] hospital. Both locations are set up with tables in rows, with several chairs at each table.

### Discussion

Thorough and uniform descriptions of educational interventions are helpful for medical educators when comparing interventions and translating good practices into their own local programs. We therefore developed a checklist for thorough descriptions of educational

---

**Table 2**

| Criteria by stage | Score,a no. (%) of 105 studies |
|------------------|--------------------------------|
|                  | 0  | 1  | 2  |
| **Preparation**  |    |    |    |
| Learning needs   | 0  | 23 (21.9) | 82 (78.1) |
| Intervention development process | 51 (48.6) | 24 (22.9) | 30 (28.6) |
| **Intervention** |    |    |    |
| Theory           |    |    |    |
| Learning objectives |    |    |    |
| Content/subject  | 2 (1.9) | 22 (21.0) | 81 (77.1) |
| Participants     | 1 (1.0) | 66 (62.9) | 38 (36.2) |
| Context and settings | 2 (1.9) | 69 (65.7) | 34 (32.4) |
| Schedule         | 7 (6.7) | 71 (67.6) | 27 (25.7) |
| Materials        | 25 (24.8) | 39 (37.1) | 40 (38.1) |
| Educational strategies | 7 (6.7) | 15 (14.3) | 83 (79.0) |
| Incentives       | 78 (74.3) | 17 (16.2) | 10 (9.5) |
| Instructors/teachers | 30 (28.6) | 56 (53.3) | 19 (18.1) |
| **Evaluation**   |    |    |    |
| Planned/unplanned changes | 87 (82.9) | 5 (4.8) | 13 (12.4) |
| Satisfaction     | 56 (53.3) | 12 (11.4) | 37 (35.3) |

*aFor scores for each study, criterion, and stage, see Supplemental Digital Appendix 3 at http://links.lww.com/ACADMED/A591.

*bScoring: 0 = no information; 1 = unclear or partial description; 2 = detailed/elaborate description. (For score definitions for each criterion, see Table 1.) Bold indicates that ≥ 50% of the studies reported a given criterion at the same score level. Percentages may not add up to 100% because of rounding.

One checklist criterion, assessment, was not scored in the systematic review; see the Method section for explanation and Table 1.
interventions in medical education. Using this list, we performed a systematic evaluation of the 105 studies included in our review of educational interventions in postgraduate medical education. Our results indicate that descriptions of educational interventions can be improved.

The only criterion covered by all articles included in our review was the learning needs that prompted the educational intervention, with 78.1% of the studies providing detailed descriptions. Other criteria that were frequently described in detail were the content/subject (77.1%) and educational strategies (79.0%) of the intervention. This finding is in line with Phillips and colleagues’ review of the literature describing educational interventions for EBP. Participants, context and settings, schedule, and instructors/teachers were frequently reported, but many studies failed to provide a complete description for these criteria. On the other hand, our evaluation also indicated that other criteria on our checklist were generally not reported; more than 75% of studies did not describe theory, incentives, environment, or planned and unplanned changes. With respect to the incentives and planned and unplanned changes criteria, it could be that some interventions had no incentives or changes. Such absence is clearly information that would be of interest to readers. Using a checklist, like ours, would aid in reporting all relevant information for a thorough description. Our results show that there is no uniformity in descriptions of educational interventions during classroom teaching for residents. Although not specific to educational interventions, lack of uniformity in descriptions of medical education research has been reported in the literature.5,7,9,12

One reason why certain criteria on our list are not covered by many studies might be that authors feel they must provide a detailed description of their intervention when their article is restricted by a word limit. As a solution, authors could describe the educational intervention in detail in a separate (design) paper, or they could submit additional material. In our review, articles that referred to additional material scored better than average (independent t test, mean difference = 2.8; P = .007; 95% CI = 0.8–4.8). However, this additional information should be easy to find, which was not always the case. On the basis of our review, we recommend describing the educational intervention in a table or figure. This would ensure that medical educators and researchers can find all of the needed information easily and help them more swiftly adapt the intervention for their local sites.

Other reasons why certain criteria are not covered may be that authors are either unaware of all of the criteria (because of the lack of a standard) or disagree regarding their importance. Because the criteria we included in our checklist are based on the literature, and because they were considered to be important during discussions at the Dutch medical education conference, it seems more likely that authors are unaware of them. We hope that this review will enhance the awareness of the importance of a description that reports on all relevant criteria.

Our results showed that more than half of the studies failed to describe the theory supporting the educational strategies. In addition, the learning objectives were reported in detail by less than one-fifth of the included studies. Without addressing these key criteria, it is unclear how one can measure whether the intervention was successful and why study results were obtained.

As we noted above, we decided not to score one criterion, assessment, because our pilot survey showed that it is often not possible to determine whether an included assessment was developed to study the effectiveness of an intervention, was an integral part of the intervention, or both. However, to assess the effectiveness of an intervention, it is of paramount importance to know whether the intervention includes a formative or summative assessment. We therefore recommend including a clear description of the assessment and its purpose in educational intervention studies.

We did not evaluate the quality of the interventions performed; we only evaluated the descriptions of the interventions. The quality of the description of an intervention does not necessarily correlate with the quality of the intervention itself. For instance, many studies lacked a description of the intervention development process; however, the interventions were likely well thought out and set up systematically. Furthermore, a complete description of an intervention does not necessarily mean the intervention was successful.

A limitation of this systematic review is that the sample of studies was limited to reports on educational interventions for residents and classroom teaching. The type of intervention may have some influence on the level and pattern of description. Classroom teaching is a relatively regulated type of intervention that is probably more easily described than many other types of educational interventions. However, there are educational interventions that are even more regulated (e.g., e-learning) or that would require special criteria (e.g., skill training). Future research could include other settings to validate our criteria list. In addition, research could explore the level of agreement on the importance of the selected criteria so that a common standard can be developed. Our checklist provides a departure point for such a standard, that, to the best of our knowledge, previously did not exist.

In conclusion, this systematic review confirms that there is room for improvement in descriptions of educational interventions during classroom teaching in postgraduate medical education. We found that key information on interventions was frequently missing and that there was a lack of uniformity of descriptions across studies. This suggests a need for a common standard for describing educational interventions. This review informs authors and readers on the essential criteria for describing educational interventions in medical education. We encourage other researchers to validate, complement, and use our criteria list, as this should lead to more complete, comparable, and replicable descriptions of interventions.

Acknowledgments: The authors are sincerely grateful to Suzanne van Rhijn for the assistance with the search of articles.

Funding/Support: Stichting BeroepsOpleiding Huisartsen (SBOH) and ZonMw; project number: 12010095425. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
Other disclosures: None reported.

Ethical approval: Reported as not applicable.

Previous presentations: The final list was discussed during a workshop at a Dutch medical education conference.50

J.G. Meinema is a PhD student, Department of General Practice/Family Medicine, Academic Medical Center–University of Amsterdam, Amsterdam, the Netherlands; ORCID: https://orcid.org/0000-0003-3706-1360.

N. Buwalda is a PhD student, Department of General Practice/Family Medicine, Academic Medical Center–University of Amsterdam, Amsterdam, the Netherlands; ORCID: https://orcid.org/0000-0003-2635-2912.

F.S. van Etten-Jamaladin is clinical librarian, Academic Medical Center–University of Amsterdam, Amsterdam, the Netherlands.

M.R.M. Visser is senior researcher, Department of General Practice/Family Medicine, Academic Medical Center–University of Amsterdam, Amsterdam, the Netherlands.

N. van Dijk is professor of general practice/family medicine, Academic Medical Center–University of Amsterdam, Amsterdam, the Netherlands.

References

1 Sullivan GM. Getting off the “gold standard”: Randomized controlled trials and education research. J Grad Med Educ. 2011;3:285–289.
2 Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: The new Medical Research Council guidance. Int J Nurs Stud. 2013;50:587–592.
3 Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ. 2015;350:h1258.
4 Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol. 2008;27:379–387.
5 Reznich CB, Anderson WA. A suggested outline for writing curriculum development journal articles: The ICDRMD format. Teach Learn Med. 2001;13:4–8.
6 Olson CA. Reflections on using theory in research on continuing education in the health professions. J Contin Educ Health Prof. 2013;33:151–152.
7 Cook DA, Beckman TJ, Bordage G. Quality of reporting of experimental studies in medical education: A systematic review. Med Educ. 2007;41:737–745.
8 Cook DA, Bordage G, Schmidt HG. Description, justification and clarification: A framework for classifying the purposes of research in medical education. Med Educ. 2008;42:128–133.
9 Patricio M, Julião M, Fareleira F, Young M, Norman G, van Carneiro A. A comprehensive checklist for reporting the use of OSCEs. Med Teach. 2009;31:112–124.
10 Cohen ER, McGaghie WC, Wayne DB, Lineberry M, Yudkowsky R, Barsuk JH. Recommendations for Reporting Mastery Education Research in Medicine (ReMERM). Acad Med. 2015;90:1509–1514.
11 Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scales RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review. Med Teach. 2009;31:270–28.
12 Windish DM, Reed DA, Boonyasai RT, Chakrabarti C, Bass EB. Methodological rigor of quality improvement curricula for physician trainees: A systematic review and recommendations for change. Acad Med. 2009;84:1677–1689.
13 Phillips AC, Lewis LK, McEvoy MP, et al. A systematic review of how studies describe educational interventions for evidence-based practice: Stage 1 of the development of a reporting guideline. BMC Med Educ. 2014;14:152.
14 Morrison JM, Sullivan F, Morey E, Jolly B. Evidence-based education: Development of an instrument to critically appraise reports of educational interventions. Med Educ. 1999;33:890–893.
15 Phillips AC, Lewis LK, McEvoy MP, et al. Development and validation of the guideline for reporting evidence-based practice educational interventions and teaching (GREET). BMC Med Educ. 2016;16:237.
16 Congres NVMO Nederlandse Vereniging voor Medisch Onderwijs [Netherlands Association for Medical Education]; November 2015; Egmond aan Zee, The Netherlands.
17 SPSS Statistics for Windows [computer program]. Version 19.0. Armonk, NY: IBM Corp.; 2010.
18 Acosta A, Azzalin A, Emmons CJ, Shuster JJ, Jay M, Lo MC. Improving residents’ clinical approach to obesity: Impact of a multidisciplinary didactic curriculum. Postgrad Med J. 2014;90:630–637.
19 Arora KS. A trial of a reproductive ethics and law curriculum for obstetrics and gynaecology residents. J Med Ethics. 2014;40(12):854–856.
20 Azim A, Singh P, Bhatia P, et al. Impact of an educational intervention on errors in death certification: A single intervention study guided the intensive care unit of a tertiary care teaching hospital. J Anaesthesiol Clin Pharmacol. 2014;30:78–81.
21 Bank I, Snell L, Bhanji F. Pediatric crisis resource management training improves emergency medicine trainees’ perceived ability to manage emergencies and ability to identify teamwork errors. Pediatr Emerg Care. 2014;30:879–883.
22 Barbas AS, Haney JC, Henry BV, Heffin MT, Lagoo SA. Development and implementation of a formalized geriatric surgery curriculum for general surgery residents. Gerontol Geriatr Educ. 2014;35:380–394.
23 Becker TK, Skiba JF, Sozener CB. Using simulation education with deliberate practice to teach leadership and resource management skills to senior resident code leaders. J Grad Med Educ. 2014;6:463–469.
24 Carek PJ, Dickerson LM, Stanek M, et al. Education in quality improvement for practice in primary care during residency training and subsequent activities in practice. J Grad Med Educ. 2014;6:50–54.
25 Carrière C, Bias M, Lafitte S, Grenier N, Revel P, Janvier G. Goal-directed ultrasound in emergency medicine: Evaluation of a specific training program using an ultrasonic stethoscope. Eur J Emerg Med. 2015;22:419–425.
26 Cartier RA 3rd, Skinner C, Laselle B. Perceived effectiveness of teaching methods for point of care ultrasound. J Emerg Med. 2014;47:86–91.
27 Chau D, Bensalem-Owen M, Fahy BG. The impact of an interdisciplinary electroencephalographic educational initiative for critical care trainees. J Crit Care. 2014;29:1107–1110.
28 Chee YE, Newman LR, Loewenstein JI, Klock CE. Improving the teaching skills of residents in a surgical training program: Results of the pilot year of a curricular initiative in an ophthalmology residency program. J Surg Educ. 2015;72:890–897.
29 Christianson MS, Washington CI, Stewart KI, Shen W. Effectiveness of a 2-year menopause medicine curriculum for obstetrics and gynecology residents. Menopause. 2016;23:275–279.
30 Corbelli J, Bonnema R, Rubio D, Comer D, McNeil M. An effective multimodal curriculum to teach internal medicine residents evidence-based breast health. J Grad Med Educ. 2014;6:721–725.
31 Daly C, Vennalapati P, Soudagar S, Hornung B, Sharma P, Gupta N. Randomized controlled trial of self-directed versus in-classroom teaching of narrow-band imaging for diagnosis of Barrett’s esophagus-associated neoplasia. Gastrointest Endosc. 2016;83:101–106.
32 diLellis P, Verderio P, Waard MW, Bottema B, Lagro-Janssen T. The effect of gender medicine education in GP training: A prospective cohort study. Perspect Med Educ. 2014;3:343–356.
33 Dolan BM, Yialamas MA, McMahon GT. A randomized educational intervention trial to determine the effect of online education on the quality of resident-delivered care. J Grad Med Educ. 2015;7:376–381.
34 Dreyer J, Hannay I, Lane R. Teaching the management of surgical emergencies through a short course to surgical residents in East/ Central Africa delivers excellent educational outcomes. World J Surg. 2014;38:830–838.
Duello K, Louh I, Greig H, Dawson N. Residents’ knowledge of quality improvement: The impact of using a group project curriculum. Postgrad Med J. 2015;91:431–435.

Feinstein RE. Violence prevention education program for psychiatric outpatient departments. Acad Psychiatry. 2014;38:639–646.

Ferrero NA, Bortsov AV, Arora H, et al. Simulator training enhances resident performance in transesophageal echocardiography. Anesthesiology. 2014;120:149–159.

Ferrell NJ, Melton B, Banu S, Coverdale J, Valdez MR. The development and evaluation of a trauma curriculum for psychiatry residents. Acad Psychiatry. 2014;38:611–614.

Fok MC, Wong RY. Impact of a competency based curriculum on quality improvement among internal medicine residents. BMC Med Educ. 2014;14:252.

Freundlich RE, Newman JW, Tremper KK, et al. The impact of a dedicated research education module for anesthesiology residents. Anesthesiol Res Pract. 2015;2015:6:23959.

Galatisatos P, Rios R, Daniel Hale W, Colburn JL, Christmas C. The lay health educator program: Evaluating the impact of this community health initiative on the medical education of resident physicians. J Relig Health. 2015;54:1188–1196.

Garcia-Rodriguez JA, Donnon T. Using comprehensive video-module instruction as an alternative approach for teaching IUD insertion. Fam Med. 2016;48:15–20.

Gorgas DL, Greenberger S, Bahner DP, Way JL, Christmas C. The lay health educator program: Evaluating the impact of this community health initiative on the medical education of resident physicians. J Relig Health. 2015;54:1188–1196.

Grillo EU, Koenig MA, Gunter CD, Kim Haspel RL, Ali AM, Huang GC. Using a team-based curriculum on quality improvement: The impact of using a group project curriculum. Postgrad Med J. 2014;91:431–435.

Hubert V, Duwat A, Deransy R, Mahjoub A, Loughland C, Kelly B, Ditton-Phare P, et al. Orthopaedic trainees retain knowledge after a partner abuse course: An education study. Clin Orthop Relat Res. 2015;473:2415–2422.

Majague S, Hanley K, Shepard J, Stewart P. Teaching multimorbidity management to GP trainees: A pilot workshop. Educ Prim Care. 2015;26:410–415.

Malone GP, Vale Arismendez S, Schneegans Warzinski S, et al. South Texas residency screening, brief intervention, and referral to treatment (SBIRT) training: 12-month outcomes. Subst Abus. 2015;36:272–280.

Matiinpur M, Sedghi I, Monajemi A, Jafari F, Montaz HE, Ali Seif Rabiei M. Clinical reasoning and improvement in the quality of medical education. Shiraz E Med J. 2014;15(4):1–4.

McCallister JW, Gustin JL, Wells-De Gregorio S, Way DV, Mastronarde JG. Communication skills training curriculum for pulmonary and critical care fellows. Ann Am Thorac Soc. 2015;12:520–525.

McDonald PL, Straker HO, Schlumpf KS, Plack MM. Learning partnership: Students and faculty learning together to facilitate reflection and higher order thinking in a blended course. J Asynchronous Learn Netw. 2014;18(4):73–93.

Mehdi Z, Ross A, Reedy G, et al. Simulation training for geriatric medicine. Clin Teach. 2014;11(5):387–392.

Mitchell JD, Mahmood F, Wong Y, et al. Teaching concepts of transesophageal echocardiography via web-based modules. J Ultrasound Med. 2013;32(5):402–409.

Nasr Esfahani M, Behzadipour M, Jalali Nadoushan A, Shariat SV. A pilot randomized controlled trial of the effectiveness of inclusion of a distant learning component into empathy training. Med J Islam Repub Iran. 2014;28:65.

Ngamruengphong S, Horsley-Silva JL, Hines SL, Pungpapong S, Patel TC, Keaveny AP, Trivedi AA, Azzam P, Douaihy A, Conti T. Teaching brief motivational interviewing: A randomized controlled trial of the effectiveness of inclusion of a distant learning component into empathy training. Med J Islam Repub Iran. 2014;28:65.

Ngaumoengpang S, Horsley-Silva JL, Hines SL, Pungpapong S, Patel TC, Keaveny AP, Trivedi AA, Azzam P, Douaihy A, Conti T. Teaching brief motivational interventions for diabetes mellitus education. Fam Med. 2016;48:167–193.
Review

85 Ogilvie E, Vlahou A, Edsell M, et al. Simulation-based teaching versus point-of-care teaching for identification of basic transoesophageal echocardiography views: A prospective randomised study. Anaesthesia. 2015;70:330–335.

86 O'Sullivan O, Iohom G, O’Donnell BD, Shorten GD. The effect of simulation-based training on initial performance of ultrasound-guided axillary brachial plexus blockade in a clinical setting—A pilot study. BMC Anesthesiol. 2014;14:110.

87 Park CS, Stojiljkovic L, Milicic B, Lin BF, Dör R. Training induces cognitive bias: The case of a simulation-based emergency airway curriculum. Simul Healthc. 2014;9:85–93.

88 Prorok JC, Stolee P, Cooke M, McKinley CA, Lee L. Evaluation of a dementia education program for family medicine residents. Can Geriatr J. 2015;18:57–64.

89 Ramanathan R, Duane TM, Kaplan BJ, Farquhar D, Kasirajan V, Ferrada P. Using a root cause analysis curriculum for practice-based learning and improvement in general surgery residency. J Surg Educ. 2015;72:2286–2293.

90 Ramassamy R, Williams A, Clark EM, Kelley Ramsingh D, Alexander B, Le K, Williams S, Sawatsky AP, Berlacher K, Granieri R. Ultrasound-guided axillary brachial plexus block in a clinical setting—A pilot study. BMC Anesthesiol. 2014;14:110.

91 Schram P, Harris SK, Van Hook S, et al. Implementing adolescent screening, brief intervention, and referral to treatment (SBIRT) education in a pediatric residency curriculum. Subst Abus. 2015;36:332–338.

92 Semler MW, Keriwala RD, Clune JK, et al. A randomized trial comparing didactics, demonstration, and simulation for teaching teamwork to medical residents. Ann Am Thorac Soc. 2015;12:512–519.

93 Shahriari M. Case based teaching at the bed side versus in classroom for undergraduates and residents of pediatrics. J Adv Med Educ Prof. 2014;2:135–136.

94 Shariff U, Kullar N, Haray PN, Dorudi S, Balsubramanian SP. Multimedia educational tools for cognitive surgical skill acquisition in open and laparoscopic colorectal surgery: A randomized controlled trial. Colorectal Dis. 2015;17:441–450.

95 Shayne M, Culakova E, Milano MT, Dhakal S, Constone LS. The integration of cancer survivorship training in the curriculum of hematology/oncology fellows and radiation oncology residents. J Cancer Surviv. 2014;8:167–172.

96 Sherrer R, Wan Y, Dong H, et al. Positive impact of integrating histology and physiology teaching at a medical school in China. Adv Physiol Educ. 2014;38:330–338.

97 Shersheva M, Kim JH, Karr C, et al. Motivational interviewing workshop in a virtual world: Learning as avatars. Fam Med. 2014;46:251–258.

98 Slott W, Blankenstein AH, Schweitzer BP, Deliens L, van der Horst HE. Effectiveness of the “availability, current issues and anticipation” (ACA) training programme for general practice trainees on communication with palliative care patients: A controlled trial. Patient Educ Couns. 2014;95:83–90.

99 Smith BW, Slack MB. The effect of cognitive debiasing training among family medicine residents. Diagnosis (Berl). 2015;2:117–121.

100 Smith CC, McCormick I, Huang GC. The clinician–educator track: Training internal medicine residents as clinician–educators. Acad Med. 2014;89:888–891.

101 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

102 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

103 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

104 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

105 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

106 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

107 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

108 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

109 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

110 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

111 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

112 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

113 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

114 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.

115 Smith ME, Navaratnam A, Jablenska L, Ziganshin A, Sozinov A, Sadigh M. Training young Russian physicians in Uganda—A unique program for introducing global health education in Russia. Ann Glob Health. 2014;80(3):182–183.