Unlimited exposure. The patient mix of GP trainees and their trainers: gaps, disparities, and active steering

de Jong, J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
A systematic review of the relationship between patient mix and learning in work-based clinical settings

Jip de Jong, Mechteld Visser, Nynke van Dijk, Cees van der Vleuten, Margreet Wieringa-de Waard

Submitted for publication
Abstract

**Background** Clinical workplace-based learning has been the means to becoming a medical professional for many years. The importance of an adequate patient mix for an optimal learning process is based on educational theory, and recognized by national and international accreditation standards. The relationship between patient mix and learning in work-based curricula as yet remains unclear.

**Aim** To review research addressing the relationship between patient mix and learning in work-based clinical settings.

**Method** The search was conducted across Medline, Embase, Web of Science, ERIC and the Cochrane Library from the start date of the database to July 2011. Original quantitative studies on the relationship between patient mix and learning were included. Methodological quality was assessed and results were extracted by two reviewers using pre-specified forms.

**Results** A total of 10,420 studies were screened on title and abstract. Of these, 298 articles were included for full-text analysis, which resulted in the inclusion of 22 papers. The quality of the included studies, scored with the MERSQI, ranged from 8.0 to 14.5. A positive relationship was found between patient mix and self-reported outcomes evaluating the progress in competence as experienced by the trainee, such as self-confidence and comfort level. Patient mix was also found to correlate positively with self-reported outcomes evaluating the quality of the learning period, such as self-reported learning benefit, experienced effectiveness of the rotation, or the instructional quality. Variables, such as supervision and learning style, might mediate this relationship. A relationship between patient mix and formal assessment has never been demonstrated.

**Conclusion** Patient mix is positively related to self-reported learning outcome, most evidently the experienced quality of the learning programme.
Introduction

Clinical workplace-based learning has played the leading role in educating medical professionals for many years. The importance of an adequate case or patient mix at that workplace for an optimal learning process is intuitively felt by many professionals and is recognized by several national and international accreditation standards. The World Federation for Medical Education emphasized this in its Global Standards for Quality Improvement, for Postgraduate Medical Education. It states ‘Training locations must have a sufficient number of patients and an appropriate case-mix to meet training objectives. The training must expose the trainee to a broad range of experience in the chosen field of medicine and, when relevant, include both inpatient and outpatient (ambulatory) care and on-duty activity. The number of patients and the case-mix should allow for clinical experience in all aspects of the chosen specialty, including training in promotion of health and prevention of disease.’

The idea that much experience is needed to become competent doctors fits theoretical frameworks. According to Ericsson, medical expertise develops by ‘deliberate practice’. He argues that expert performance is different from everyday performance, as it continues to improve as a function of more experience, coupled with deliberate practice. Expert performance is reached by actively acquiring and refining a cognitive mechanism to support continued learning and improvement. Becoming a medical expert thus requires engagement in practice and appropriate reflection, which can be stimulated by feedback from coaches or trainers. In addition, other frameworks for medical expertise emphasize the importance of clinical experience for learning, such as theories of cognitive structures and dual processing. The essence of these theories is that first conscious, intentional learning (deliberate practice) must be established before routines are automated. These automated routines are the basis of adequate medical handling. Within these frameworks, the experience needed is provided by an adequate patient mix, so patient mix is an important training condition. Following deliberate practice theory, a well-supervised patient mix can be assumed to substantially improve medical competence.

This review was carried out in order to evaluate whether this theory could also be confirmed by empirical evidence. Our primary aim was to systematically review research addressing the relationship between patient mix and learning in workplace-based clinical settings. Our secondary aim was to address the influence of additional variables (e.g. supervision and learning style) on this relationship.
For this purpose, ‘patient mix’ was defined as the number of patients and the types of medical problems presented to learners.\textsuperscript{11} It thus is regarded to consist of a number of patients presenting a certain diversity of diseases (Figure 1). Numerous publications in the medical educational field have followed this ‘volume and diversity’ approach.\textsuperscript{12-14}

Figure 1. Patient mix model

Two-dimensional approach of patient mix. Consider a training programme in a large group of patients with the same diagnoses (A). Because there is no diversity, this cannot be considered to be a ‘patient mix’, but merely the training of a single skill or restricted clinical problem. Situation C expresses a situation in which a large diversity of skills and/or symptoms and diagnoses are theoretically possible, but very few or even no patients are present. This can be regarded as a patient mix, but an extremely meagre one – simply because there are no or very few patients. The line labelled ‘D’ expresses (an arbitrary) ‘cutoff point’ where the diversity is rich enough to start calling the population ‘patient mix’. All points labelled B are considered to be a patient mix.
Methods

ELIGIBILITY CRITERIA (LIST 1)
As we aimed to assess the strength of the relationship between patient mix and learning, only quantitative studies were included which were conducted with medical students/trainees at any level of the formal medical training/career. Patient mix volume, i.e. the quantity of patients encountered and the diversity of skills and/or symptoms and diagnoses had to be described. No simple cut-off for the width of this diversity could be given (Figure 1), but studies on the exposure to one restricted clinical problem or skill were excluded as they only described the volume of that skill or problem, and no diversity. Learning outcome had to be explicitly assessed. The relationship between patient mix and learning had to be quantified by statistical analysis.

INFORMATION SOURCES AND SEARCH STRATEGY
The search was conducted across five sources relevant to education in a clinical context: Medline, Embase, Web of Science, ERIC and the Cochrane Library. The search ran from the start date of the database to July 2011 and was not limited by language, geography, or research methodology. The search strategy was composed by a clinical librarian. The search strategy had to be able to find a ‘golden standard set’ in PubMed. This set contained thirty-eight articles, previously rated as being relevant to the review subject by the authors. The strategy was then translated to the search systems of the other databases.

STUDY SELECTION
Two authors (MW and JJ, or MV and JJ) individually and independently screened the titles and abstracts of all articles using the inclusion and exclusion criteria. Citations that were selected by one author but not by the other author were discussed in order to achieve negotiated consensus on inclusion or exclusion. In case of doubt or persisting disagreement in this phase, the article was included. The full text of all the potentially relevant articles was retrieved. The full-text articles were screened, again independently, by two authors, using the same criteria and were again compared. In case of disagreement, a decision on inclusion or exclusion was once more reached by negotiated consensus. Most studies that were excluded did not have an adequate description of patient mix or did not statistically address the relationship between
patient mix and learning. At each screening phase, each citation was marked as ‘yes’, ‘maybe’ or ‘no’. Inter-observer agreement of the screening phase was measured by Cohen’s Kappa (linearly weighted). Manual searches were conducted across the citations of the papers that were coded, resulting in 17 more citations. These were screened by two authors, but none of them were included.

List 1. Inclusion and exclusion criteria

### Inclusion criteria
- Empirical, quantitative, educational studies with actual patients (no simulations).
- Study population: studies conducted with medical students/trainees at any level of the formal medical training/career.
- Patient mix, clinical encounters, or clinical experience in workplace-based learning had to be described. Patient mix had to be described in some detail, thereby addressing the volume as well as the diversity. Studies on the exposure to a restricted clinical problem or skill were excluded. Medical subspecialties were not excluded beforehand, as long as the patient mix was diverse.
- Learning outcome measures had to be described by self-reported measures, assessment by trainers, preceptors, or others, or by objective structured clinical examinations (OSCEs), multiple choice, or other written exercises.
- The relationship between patient mix and learning had to be quantified by statistical analysis.
- Studies in all languages were included.

### Exclusion criteria
- Qualitative studies.
- Dental and veterinary curricula, any paramedical curricula, nursing curricula, physician assistant curricula, nurse practitioner curricula, dietetic curricula.
- Theoretical medical curricula (not work-based).
- Complementary/alternative medicine.
DATA EXTRACTION
A detailed data extraction form was developed using the Best Evidence in Medical Education (BEME)\textsuperscript{15} standard coding sheet and published reviews\textsuperscript{16,17} as a basis. All selected papers were coded by the authors in pairs (MW and JJ, or MV and JJ). This form contained a general description of the study design and participants, including the training level and the specialty training area (Table 1). In addition, patient mix instruments (e.g. electronic logbooks and questionnaires), volume/diversity descriptions (e.g. top 10 skills or diagnoses lists), learning outcome measures, and the relationship found between patient mix and learning were recorded. We also documented the highest level of the Kirkpatrick hierarchy\textsuperscript{18} on which learning outcomes were assessed. If additional variables were studied in relation to learning outcome (e.g. learning style and supervision), these were also recorded.

QUALITY OF THE STUDIES
To obtain an overview of the quality of the included studies, and thereby the validity of the outcomes, we assessed them with the recently developed Medical Education Research Study Quality Instrument (MERSQI).\textsuperscript{19} This instrument was chosen because, to our knowledge, it is the only instrument fitted for observational studies considering medical education. Two authors (JJ and MV) independently scored the quality of the included papers. In case of disagreement on item scores, a decision was reached by negotiated consensus.

DATA ANALYSES
The various ways in which patient mix was operationalized were categorized in equivalent approaches of volume and diversity descriptions. Learning outcomes were divided into self-reported outcomes and outcomes using formal assessments. The relationship between patient mix and learning is described in sections based on different learning outcomes, as this allowed for homogeneous reporting of results.
Results

SEARCH RESULTS
The search resulted in 11,098 titles. After removal of duplicates, 10,420 studies were reviewed based on title and abstract. Of these studies, 298 were identified as potentially relevant and included for full-text analysis, which resulted in the inclusion of 22 papers (Figure 2). The studies identified had insufficient homogeneous or quantitative data to allow meta-analysis or other formal synthesis. During screening of titles and abstracts, the inter-observer agreement kappa (linearly weighted) was 0.34 (MV–JJ) respectively 0.32 (MW–JJ). Table 1 provides a summary of study descriptions and outcomes. This table forms the basis for the inferences from the studies in the following paragraphs.
METHODOLOGICAL QUALITY OF STUDIES

MERSQI sum-scores ranged from 8.0 to 14.5 (median 11.75). Ten studies reported on the internal structure of their outcome instruments by Cronbach’s Alpha (Table 1) or principal component analysis. Response rates, if presented, varied from 43% to 100%. Data analyses were appropriate in all but one study, and all were beyond the descriptive level. Four studies reported outcome at satisfaction level, whereas 10 studies measured knowledge and/or skills. Less than half of the studies (n=8) measured outcomes at the behavioural level. None of the included studies explicitly measured patient or health-care outcome.

TYPES OF STUDIES

In six studies, the mutual dependence of factors related to learning was addressed in a path analysis or structural equation modelling (SEM) (Table 1). Eight studies compared the patient mix of training sites and their contribution to learning. In three of these studies, similar sites were compared, three others compared academic vs. non-academic sites and two compared inpatients and outpatients. Four studies evaluated the learning effects of an intervention: the introduction of a rotation, a skill-training programme, identification of 10 pre-selected complaints, and a new internship. Two studies compared groups of medical students at a different phase of their training. In four studies, the groups and sites were homogeneous and no interventions were studied.

OPERATIONALIZATION OF PATIENT MIX

Patient mix was measured with various instruments (see Table 1), including questionnaires (n=11), interviews (n=1), and logbooks (n=13), the latter hand-written (n=9), electronic (n=2), or unspecified (n=2). Patient mix was mostly described as the variety of encountered skills and/or diagnoses. Skills were usually technical procedures, such as intubation or suturing. In some studies (n=6), the patient volume was the most pronounced patient mix characteristic. The diversity of the patient mix in these studies was often additionally addressed by one or two variables, but the reports lacked a detailed insight into the diversity of diagnoses. In most other studies, the distribution of encountered diagnoses and medical skills was presented. Several authors presented a top 10 or 20 of the conditions the students meet most frequently. This method was also used to compare the patient mix of different sites.
Learning outcome measures can be divided into self-reported outcomes and formal assessments. Self-reported outcomes are used in ten studies, in five of these, the self-estimated competence was measured as self-confidence or comfort level. In five other studies, the quality of the learning experience or the educational profit of the experience at issue is asked for; such as the effectiveness, the learning benefit, or instructional quality of the rotation. Formal assessments were more diverse. Usually, knowledge or skills were tested; sometimes including clinical performance. Methods used included multiple choice examinations (MCQ) and other written examinations, clinical assessments, oral examinations, and OSCEs.

The relationship was usually tested by comparing the volumes of diagnoses or skill-diversity, or by determining the loading of one patient mix variable to an outcome variable in a path analysis.

In four studies, a positive relationship was found between patient mix and self-reported outcomes evaluating the progress in competence as experienced by the trainee, such as self-confidence and comfort level. By contrast, one study found no difference in confidence between residents in a traditional inpatient rotation and a new one in which experience in ambulatory settings was introduced. Patient mix was also found to correlate with self-reported outcomes evaluating the quality of the learning period, such as self-reported learning benefit, experienced effectiveness of the rotation, or the instructional quality.

In one trial, students in an intervention group who encountered significantly more often 10 chief prerequisite complaints than the control group (31.8% vs 6%) outperformed the control group on a general knowledge examination ($p=0.014$). In six other studies, however, no relationship between patient mix and scores on MCQs or other written examinations was found.
OSCE
No association was found between patient mix and performance in four of the five studies using OSCE assessment.\textsuperscript{20,34,36–38} Fung et al. suggested that the time allotted for students to complete clerkships may not be sufficient to expose them to the number of patients needed to generate a significant effect on clinical performance.\textsuperscript{37} In one study, the OSCE scores even seemed lower in students who attended a higher number of outpatient clinics than those attending fewer outpatient clinics, although experience with emergency admissions and obtaining feedback on these seemed to improve OSCE performance.\textsuperscript{36} The authors concluded that the clinical skills were enhanced by an increased volume of some, but not all, clinical experience. Jolly et al. found that students scored higher on OSCEs if they examined patients on their own, if the objectives (presumably the objectives of the rotation, not reported) had been made clear, and a higher number of clinics were attended.\textsuperscript{38}

ORAL EXAMINATION, CLINICAL ASSESSMENT
Wimmers et al. found that an increased number of patient encounters did not directly lead to improved clinical performance as assessed by supervisors in 227 medical students,\textsuperscript{41} as was the case in two other studies.\textsuperscript{36,39} They did, however, find a strong relationship between number of patients and number of diseases encountered ($r=0.89$).\textsuperscript{41} Ahmed and Hughes, in contrast, found that exposure rates did correlate with the assessment grades awarded by clinical supervisors, but not with a written exercise (quiz) score.\textsuperscript{35} Also Greenberg and Getson found a weak positive correlation between number of patients seen and the students’ clinical performance.\textsuperscript{27}

VARIABLES POTENTIALLY RELEVANT TO THE RELATIONSHIP BETWEEN PATIENT MIX AND LEARNING
Martin et al. found that trainees with a deep, strategic, and well-organized learning style reported significantly higher clinical exposure (combined score for three areas of clinical activity). The well-organized style was also associated with OSCE performance.\textsuperscript{20} McManus et al. additionally found that the amount of knowledge gained from clinical experience was related to strategic and deep learning styles\textsuperscript{39} as was success in a final examination: positive and significant correlations were found for deep and strategic learning, whereas surface learning correlated negatively.
In the path analysis presented by Wimmers et al., supervision quality loaded on patient mix volume and on clinical competence. Hoifoidt et al., however, did not find supervision to load on patient mix volume or on (subjective) learning benefit while in the model of Van der Zwet et al., supervision loaded on both patient mix and instructional quality. In the study by Dolmans et al., a relationship was found between supervision and the effectiveness of a rotation. Also, a significant two-way interaction was found between patient mix and supervision; the latter more strongly influenced the effectiveness of the rotation than patient mix did. In another study, OSCE score also seemed to be ‘modified’ by the quality of the feedback.

Ahmed and Hughes found indications of a relationship between professionalism and both patient mix and composite assessment grades at the end of an attachment. Hoifoidt et al. described that amount of experience affected the quality of the learning environment which itself was related to the learning benefit.

In the study by Yu et al., the overall quality of the surgical clerkship, as perceived by students, was related to the number of cases seen, although no difference in learning outcome was found. Jolly et al. found that six of 43 questionnaire variables correlated with OSCE score. Two of these six can be considered to be related to the learning climate, namely ‘whether students examined patients on their own’ and ‘whether objectives were made clear’.

Discussion

In most studies dealing with the relationship between patient mix and student self-assessment (self-confidence, comfort level), indications of a relationship were found. The indications of positive relationships were stronger regarding the quality of the learning experience (learning benefit, instructional quality, or effectiveness of a rotation). Supervision quality seems to be a mediating factor, which was repetitively found to improve patient- or education-related outcome. This can be regarded to be consistent with the theory of deliberate practice.

The relationship between patient mix and learning outcome was not corroborated with formal assessment outcomes. All but one study dealing with MCQ or other written examinations failed to find any relationship between patient mix and MCQ or written examinations. All the studies relating patient mix to OSCE score
found no association, or under some conditions even a negative association. In one study, a correlation was found between exposure rate and clinical assessment grades, whereas three other studies did not find such a relationship.

The patient mix (also called ‘clinical exposure’ or ‘case mix’) in the articles we reviewed was mostly presented without definition. We found studies describing skills, diagnoses, treatments, or general ideas about patient mix, within different specialties and measured by logbook or questionnaires and presented differently, making the patient mix descriptions extremely heterogeneous. The heterogeneity we found is particularly interesting. In the light of the emphasis, adequate patient mix has gained in the diverse accreditation standards of several countries and internationally. Due to the heterogeneity, we had difficulty in finding a proper cut-off point for the number of diagnoses or skills that need to be engaged to fulfil the diversity inclusion criterion. This heterogeneity indicates the need for a discussion on the value of the concept. For instance, Berlowitz et al. stated that patient mix should describe how patients are distributed along characteristics that may affect specific outcomes of interest; he thereby stresses that the concept of patient mix in itself is not relevant. The fact that the patient mixes described in the reviewed studies are so diverse, may be partly because they are related to different outcomes in the different settings at different stages of education. It often seemed that the presented patient mix depended on the instrument the authors had to their disposal and not on study-specific operationalization of the desired patient mix of the attachment. Besides the relationship with the outcome, more clarity about the relationship between the diversity and volume aspects of patient mix might be strived for. In this review, we found operationalizations of patient mix that were fairly different in that respect, allowing for very few inferences between studies.

The learning outcome measures were classified into self-reported assessment and formal assessment. The precise description of the used formal assessment methods in the studies was often meagre; example questions or exercises were not found. The reliability of clinical assessments is questionable; subjectivity can be a problem. Pulito et al. found that direct observations of trainees interacting with patients occur too infrequently. Students prepare for assessments and their results may reflect their preparation more than their real competence. Terms such as OSCE or MCQ might suggest that similar instruments were used in different studies. However, the precise content and the number of stations or questions were found to differ, if mentioned at all.
We formulated six possible explanations that could explain why we found so little evidence for the relationship between patient mix and the results of formal assessment.

1. **Patient mix does not contribute to medical competence development.**
   This idea is highly unlikely, although theoretically possible. The positive relationship between patient mix and self-assessment outcome (compared with formal assessment) is not per se an indication of a relationship between patient mix and learning. The largely absent relationship between patient mix and formal assessment might indicate that ‘clinical experience without training increases confidence but not competence’.
   The idea that one becomes automatically more competent with increasing experience can be illusory.

2. **The relationship between patient mix and learning is more complex and many other variables play a role (such as supervision quality, learning style, learning environment, or professionalism).**
   Based on educational theory, the importance of other variables in the relationship could be expected. In several studies, supervision was found to be strongly related to learning outcome. It may, therefore, be seen as an important mediator. Supervision quality was, however, not described or measured in the majority of the included studies, and the potentially mediating effect may have been overlooked. This may have been the case with other variables related to learning as well.

3. **The time span covered in most studies was too short.**
   Current educational theories assume a general problem-solving ability, but case-specific knowledge is considered of predominant importance. This means that competences do not transfer easily, which implies that long exposure in many domains is essential for doctors to become fully competent. The time-span covered by most of the included studies, may have been too short to find positive results.

4. **The patient mix is inadequately measured.**
   Patient mix is usually described by encountered skills or diagnoses and in terms of volume and diversity. Other potentially relevant descriptors are the complexity in relation to the stage of learning and the learning value or benefit of cases. These aspects, with exceptions, are not usually described, so the validity of the instruments might have been imperfect.
Furthermore, in several studies, the patient mix was aggregated per training site and comparison was made between sites, not between students. Maybe this is a too coarse comparison to establish the relationship between patient mix and learning which can also be regarded as a limitation of our inclusion and exclusion criteria.

5. **The validity of formal assessment is insufficient.**
In a systematic review, Hamdy et al. found only mild to moderate correlations between measurements obtained in medical schools and future performance in medical practice.\(^5\) The clinical validity of OSCEs was also questioned by McManus. OSCEs, MCQs and other assessments may not be appropriate for determining the specific contribution of patient mix on learning.

Self-reported outcome instruments are usually designed especially for the study, whereas the formal assessments used are commonly part of the standard assessment procedures. These are not tailored to the study question and may suffer from bias due to the preparation of students for assessments.\(^5\) The self-report instruments might therefore be a more appropriate fit for the research questions.

6. **The quality of the studies was insufficient.**
The majority of the included studies had a single-group, post-test-only design which may be considered inferior to pre–post test designs (n=4) or trials (n=2). Several studies were merely a comparison of training sites or an evaluation of a new curriculum.\(^3\)\(^2\);\(^3\)\(^5\)

**Limitations**
Composing an efficient and sufficient search strategy is complex. Despite our attempts to sharpen the patient mix definition to an accurate and workable one, we made pragmatic choices. We were not able to formulate the exact border between ‘some disease (or skills)’ and ‘patient mix’. This resulted in low inter-rater agreement. A substantial number of papers were included or excluded based on negotiated consensus. This happened more in the beginning of the review process (Phase 2) than later (Phase 3). A minority of the studies included were intended to specifically explore the relationship between patient mix and learning for general educational theory purposes. Most of the studies concerned merely an evaluation of a programme or curriculum change, or were a comparison between training sites.

Many studies had to be excluded because they lacked a statistical analysis of the relationship between patient mix and learning outcome.
FUTURE DIRECTIONS
This systematic review emphasizes the problem with the description of ‘patient mix’. Despite its attention in international accreditation standards, the concept itself seems poorly defined. Educational research would benefit from a standardized approach in patient mix descriptions; volume can always be measured, but diversity should be explicated in relation to the outcome. Future studies should aim at addressing which parts of patient mix (e.g. volume and diversity) contribute to learning and which parts do not. A theoretical framework accounting for other relevant parameters in the relationship between patient mix and learning, such as supervision and learning style, may be helpful, and instead of using the standard assessment procedures, research question tailored, objective outcome should be developed.

Nearly all studies at hand ‘accepted’ the patient mix that was presented to the participants as a given factor. Interventions on the patient mix were indirect (curriculum change) active influencing of the patient mix was not found. It would be interesting to see what the effect of tailoring the patient mix to the specific learning goals and needs of individual students would yield.

To avoid bias due to preparation for an examination, study outcome (assessments) should be unobtrusive; for instance, assessments based on a random selection of routinely made video recordings could be considered. If, at second best, a traditional approach is chosen – similar to the designs we found, triangulation should be strived for. This can be reached by measuring study-tailored self-assessment, including the quality of the learning experience and self-confidence or alike, combined with formal assessment derived from knowledge assessment, and assessment of clinical competence. These studies are preferably done in multi-institution trials. An inquiry into the detailed aspects of patient mix, and the contribution of these aspects to learning is desirable. This may be done in a qualitative design; trainees and clinical teachers may be interviewed about their ideas of minimal or optimal patient volumes and spread of diagnosis diversity and their benefit for learning.

CONCLUSIONS
Patient mix, defined in terms of volume and diversity, is related to self-reported learning outcome, most evidently the experienced quality of the learning programme. A relationship between patient mix and the results of formal assessment has rarely been demonstrated. Not only supervision in particular, but also learning style seem mediating variables of the relationship between patient mix and learning.
Table 1

<table>
<thead>
<tr>
<th>Author</th>
<th>Year of publication</th>
<th>Country</th>
<th>Design</th>
<th>Time span</th>
<th>Speciality</th>
<th>Nr. of participants analysed and Educational level participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmed, 1999</td>
<td></td>
<td>UK</td>
<td>Single group, post test</td>
<td>7 weeks</td>
<td>Paediatrics</td>
<td>226 Students</td>
</tr>
<tr>
<td>Boots, 2008</td>
<td>Australia</td>
<td></td>
<td>Pre-post test design with two groups</td>
<td>10 weeks</td>
<td>Internal medicine</td>
<td>230 Students and 174 Interns</td>
</tr>
<tr>
<td>Chatenay, 1996</td>
<td>Canada</td>
<td></td>
<td>Single cohort randomized to attend one out of four training sites. Post test only</td>
<td>10 weeks</td>
<td>Surgery</td>
<td>109 Students</td>
</tr>
<tr>
<td>Dolmans, 2002</td>
<td>NL</td>
<td></td>
<td>Observational cohort study</td>
<td>3-12 weeks per rotation</td>
<td>8 disciplines</td>
<td>1208 Residents</td>
</tr>
<tr>
<td>Duke, 2011</td>
<td>Canada</td>
<td></td>
<td>Non equivalent control group, pre-post test</td>
<td>4 weeks</td>
<td>Family medicine</td>
<td>79 Residents</td>
</tr>
<tr>
<td>Fung, 2007</td>
<td>USA</td>
<td></td>
<td>Retrospective single group, post test</td>
<td>Not specified (three clerk-ships)</td>
<td>Inpatient internal, ambulatory and family medicine</td>
<td>166 Third year students</td>
</tr>
<tr>
<td>Greenberg, 1999</td>
<td>USA</td>
<td></td>
<td>Single group, post test</td>
<td>8 weeks</td>
<td>Paediatrics</td>
<td>118 Students</td>
</tr>
<tr>
<td>Gruppen, 1993</td>
<td>USA</td>
<td></td>
<td>Single group, pre-post test</td>
<td>1 month</td>
<td>Internal medicine</td>
<td>43 Third year students</td>
</tr>
<tr>
<td>Hoifodt, 2004</td>
<td>Norway</td>
<td></td>
<td>Cross-sectional</td>
<td>4 months</td>
<td>Psychiatry</td>
<td>85 Preregistration house officers</td>
</tr>
<tr>
<td>Jacobson, 1998</td>
<td>USA</td>
<td></td>
<td>Observational cohort study</td>
<td>I2 weeks</td>
<td>Internal medicine</td>
<td>43 Students</td>
</tr>
<tr>
<td>Jolly, 1996</td>
<td>UK</td>
<td></td>
<td>Single group, post test</td>
<td>4 + 16 weeks</td>
<td>Pathology, then medicine and surgery</td>
<td>152 Clinical students</td>
</tr>
<tr>
<td>Lampe, 2008</td>
<td>USA</td>
<td></td>
<td>Non-randomized trial</td>
<td>6 months</td>
<td>Emergency medicine</td>
<td>37 Senior medical students</td>
</tr>
<tr>
<td>PM instrument</td>
<td>PM description (Volume/diversity)</td>
<td>Learning Instrument (self-reported measure is specified, Cronbach’s alpha if reported)</td>
<td>Relation patient mix and learning</td>
<td>MERSQI</td>
<td>Highest Kirkpatrick level</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------------------------</td>
<td>--------</td>
<td>--------------------------</td>
<td></td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Change in exposure from year 1 to year 2 in 42 conditions and 20 skills</td>
<td>MCQ† and other written exam (0.67-0.71) General and clinical competence assessment</td>
<td>Clinical experience did not relate to MCQ/written exam score</td>
<td>I3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Encounters (ordinal scale) of 15 specified skills</td>
<td>Questionnaire, self-confidence</td>
<td>There may be a relation between skill exercise and confidence</td>
<td>9.5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Put volume of elective/ ER admissions, operations, scrubs, outpatient clinics and procedures.</td>
<td>MCQ† (0.67), OSCE* (0.48), Clinical performance</td>
<td>Skills were enhanced by increased volume of some but not all clinical experience. Complex relation between feedback and OSCE* performance. Quality of feedback seems to mediate this relationship</td>
<td>I4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Questionnaire</td>
<td>One variable composed of 3 questionnaire items (sufficient-patients, diagnostic variety and patients independently dealt with)</td>
<td>Self-reported effectiveness questionnaire</td>
<td>Self-perceived effectiveness depends on Patient mix and supervision. Supervision more strongly influences effectiveness when patient mix is limited</td>
<td>I0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(Unspecified) Logbook</td>
<td>- Percentage of students that attended clinics of 18 subspecialty</td>
<td>MCQ† knowledge test Self-confidence questionnaire Oral and written feedback assessment</td>
<td>No difference between ambulatory and inpatient sites</td>
<td>I2.5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Electronic (PDA**) logbook</td>
<td>Number of patients in 6 diagnostic categories</td>
<td>OSCE* (0.34-0.65)</td>
<td>No relation between patient exposure and OSCE* score</td>
<td>I2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Number of patients in 4 diagnostic domains</td>
<td>Clinical performance, case presentation and NBME†</td>
<td>No relation between volume and exam score</td>
<td>II.5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Number of patients, Percentage of students that encountered 20 diagnostic categories Top 15</td>
<td>Written exam</td>
<td>No correlation between the students levels of experience and knowledge</td>
<td>I2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Nr of subjects having experience in 12 psychiatric skills</td>
<td>Questionnaire, Subjective learning benefit (0.84)</td>
<td>Subjective learning benefit was related to amount of experience, competence and formal teaching programme. Supervision and previous experience had no impact on subjective learning</td>
<td>II.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Top 10, Percentage of encounters of 6 diagnostic categories and 1 skill</td>
<td>Categorization of self-reported learning points</td>
<td>In and outpatient encounters differed. Learning differences between in- and outpatient apply to pathophysiology, evaluation/work-up and patient education/counselling.</td>
<td>8.5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Questionnaire responses handling patient volume and skills</td>
<td>OSCE* (0.69)</td>
<td>No relation between clinical experience and educational outcome</td>
<td>I3.5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Electronic (PDA**) logbook</td>
<td>Nr of students that met the pre-specified target complaints.</td>
<td>MCQ† and written exam</td>
<td>Group seeing a required number of representative patients showed better knowledge</td>
<td>I2.5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Author, Year of publication</td>
<td>Country</td>
<td>Design</td>
<td>Time span</td>
<td>Speciality</td>
<td>Nr. of participants analysed and Educational level participants</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------</td>
<td>-----------</td>
<td>------------</td>
<td>---------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Martin, 2000 UK</td>
<td></td>
<td>Single group, post test</td>
<td>1 year</td>
<td>Medicine and Surgery</td>
<td>194 (150 returned learning style form) Students</td>
<td></td>
</tr>
<tr>
<td>Mcleod, 1997 Canada</td>
<td></td>
<td>Observational cohort study</td>
<td>8 week</td>
<td>Internal medicine</td>
<td>40 Residents and 29 clinical clerks</td>
<td></td>
</tr>
<tr>
<td>McManus, 1998 UK</td>
<td></td>
<td>Prospective study of two cohorts assessed at application to med school and at the end of their final year</td>
<td>About 5 years</td>
<td>Undergraduate curriculum, not otherwise specified</td>
<td>684 Students (1st cohort 301/ 2nd 383)</td>
<td></td>
</tr>
<tr>
<td>Nomura, 2008 Japan</td>
<td></td>
<td>Pre-post test design with non-equivalent control group</td>
<td>2 years</td>
<td>Multidisciplinary</td>
<td>2474 before +1166 after Postgraduate 'residents' without clinical experience</td>
<td></td>
</tr>
<tr>
<td>O’Hara, 2002 USA</td>
<td>USA</td>
<td>Single group, post test</td>
<td>4 weeks + 4 days</td>
<td>Women health care in a family medicine clerkship</td>
<td>445 Students</td>
<td></td>
</tr>
<tr>
<td>Saywell, 2002 USA</td>
<td>USA</td>
<td>Single group, post test</td>
<td>4 weeks</td>
<td>Musculoskeletal medicine in a family medicine clerkship</td>
<td>445 Third year students</td>
<td></td>
</tr>
<tr>
<td>Schwiebert, 1993 USA</td>
<td>USA</td>
<td>Single group, post test</td>
<td>1 month</td>
<td>Family medicine</td>
<td>185 Third year students</td>
<td></td>
</tr>
<tr>
<td>Wimmers, 2006 NL</td>
<td>NL</td>
<td>Single group, post test</td>
<td>12 weeks</td>
<td>Internal medicine</td>
<td>152 Students</td>
<td></td>
</tr>
<tr>
<td>Yu, 2011</td>
<td>New Zealand</td>
<td>Single group, post test</td>
<td>6 weeks</td>
<td>Surgery</td>
<td>166 Fourth year students</td>
<td></td>
</tr>
<tr>
<td>vd Zwet, 2010 NL</td>
<td>NL</td>
<td>Single group, post test</td>
<td>10 weeks</td>
<td>General Practice</td>
<td>284 Fifth year students</td>
<td></td>
</tr>
</tbody>
</table>

* Objective Structured Clinical Examination
† Multiple Choice Questions
‡ National Board of Medical Examiners
** Personal Digital Assistant
<table>
<thead>
<tr>
<th>PM instrument and Hand written logbook</th>
<th>PM description (Volume/diversity)</th>
<th>Learning Instrument (self-reported measure is specified, Cronbach’s alpha if reported)</th>
<th>Relation patient mix and learning</th>
<th>MERQI</th>
<th>Highest Kirkpatrick level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire and Hand written logbook</td>
<td>Nr of patients in 7 diagnostic categories</td>
<td>Clinical assessment, Critical Appraised Topic and OSCE* (0.69-0.74)</td>
<td>Heterogeneity of clinical exp from sites did not translate into heterogeneity of learning outcomes</td>
<td>13.5</td>
<td>2</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Based on factor analysis: Patient mix variable composed of 3 questionnaire items: number of patients, patient variety and quality of patient contacts</td>
<td>Questionnaire, instructional quality</td>
<td>Supervision and patient mix load on instructional quality</td>
<td>10.0</td>
<td>1</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>- Nr of patients - Nr of different diseases encountered</td>
<td>Written exam, oral exam</td>
<td>Slight differences between university and private practice in patient mix, but no difference in results on oral and written exam</td>
<td>11.5</td>
<td>2</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>- Proportion of students that encountered each of 20 diagnoses ordered as a top 20. - Number of patients</td>
<td>Combined clinical performance assessment and oral examination (0.67)</td>
<td>An increased nr. of patient encounters did not (directly) lead to improved competence. Quality of supervision indirectly had impact on student learning and the nr. of patient encounters</td>
<td>13.0</td>
<td>3</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>- Volume of patient by ages - Volume and percentage of Top 10 diagnosis by age group</td>
<td>Questionnaire comfort level</td>
<td>Relationship between experience and comfort level between some diagnostic categories</td>
<td>8.0</td>
<td>1</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>- Volume of patient by ages - Volume and percentage of Top 10 diagnosis by age group</td>
<td>Questionnaire comfort level</td>
<td>Relationship between experience and comfort level between some diagnostic categories</td>
<td>8.0</td>
<td>1</td>
</tr>
<tr>
<td>Interview, (unspecified) logbook, Questionnaire</td>
<td>- Patient numbers of 20 diagnoses, ordered as top 20. - (Top) 15 skills, % patient encounters in which skill was relevant</td>
<td>Relevance for learning Questionnaire / interview</td>
<td>Inpatient based experience is better than ambulatory care experience for learning</td>
<td>8.5</td>
<td>2</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Single experience score based on 15/20 (1st /2nd cohort) unspecified acute conditions, 10/10 surgical operations and 17/29 practical procedures</td>
<td>MCQ and other written exam Clin performance (0.87, 0.88)</td>
<td>No association between clinical experience and exam score. Study habits predict examination performance</td>
<td>14.5</td>
<td>3</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Nr of students with ‘no experience’ with specified diagnosis (grouped) in 22 (sub) specialties</td>
<td>Questionnaire self confidence</td>
<td>Clinical experience and confidence levels improved, especially at university hospitals</td>
<td>10.5</td>
<td>2</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>- Volume of patient by ages - Volume and percentage of Top 10 diagnosis by age group</td>
<td>Questionnaire comfort level</td>
<td>Relationship between experience and comfort level between some diagnostic categories</td>
<td>8.0</td>
<td>1</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>- Hand written logbook</td>
<td>- Total nr of patients, outpatients and emergencies.</td>
<td>OSCE* (0.70)</td>
<td>No association between clinical experience and OSCE* score. Positive association clinical experience with learning style.</td>
<td>12.0</td>
</tr>
</tbody>
</table>

- PM = Performance Measures
- MERQI = Medical Education Research Quality Index
REFERENCES

1. Liaison Committee on Medical Education (LCME). Functions and Structure of a Medical School. Standards for Accreditation of Medical Education Programs Leading to the M.D. Degree. http://www.lcme.org/functions2010jun.pdf
15. BEME Collaboration. http://www2.warwick.ac.uk/fac/med/beme/


Appendix 1. Search strategies

1. PUBLMED

Set 1: patient mix

((("Diagnosis-Related Groups"[MeSH] OR diagnosis related group*[tiab] OR case mix*[tiab]
OR casemix*[tiab] OR diagnosis cluster*[tiab] OR patient distribution*[tiab] OR clinical exposure*[tiab]
OR selected conditions*[tiab] OR disease management*[tiab] OR clinical method*[tiab] OR diagnosis cluster*[tiab]
OR distribution patients*[tiab])))) AND

Set 2: learning

OR “Clinical Competence”[MeSH] OR clerkship*[tiab] OR trainee*[tiab] OR training*[tiab]
OR (work*[tiab] OR resident*[tiab] OR residency*[tiab] OR (“work”[MeSH Terms] OR work*[tiab])) AND
(based*[tiab]) AND (“learning”[MeSH] OR learn*[tiab]))) AND

Set 3: population

(“Hospitals, Teaching”[MeSH] OR teaching hospital*[tiab] OR “Specialties, Medical/education”[MeSH]
OR “Primary Health Care”[MeSH] OR student*[tiab] OR practice*[tiab]))

2. EMBASE

Set 1: patient mix:

case mix/ or case mix*[ti,ab.] or casemix*[ti,ab.] or diagnosis related group/ or diagnosis related group*

Set 2: learning:

curriculum/ or curricul*[ti,ab.] or exp Medical Education/ or medical education*[ti,ab.] or exp Clinical Competence/ or clerkship*[ti,ab.] or trainee*[ti,ab.] or training*[ti,ab.] or work based learning*[ti,ab.]

Set 3: population:

exp Teaching Hospital/ or teaching hospital*[ti,ab.] or exp medicine/ or exp Primary Health Care/ or
student*[ti,ab.] or practice*[ti,ab.]
3. COCHRANE LIBRARY

#1 (case mix*):ti,ab,kw
#2 (casemix*):ti,ab,kw
#3 (diagnosis related group*):ti,ab,kw
#4 MeSH descriptor Diagnosis-Related Groups explode all trees
#5 (clinical exposure*):ti,ab,kw
#6 (clinical encounter*):ti,ab,kw
#7 (clinical experience*):ti,ab,kw
#8 (patient mix*):ti,ab,kw
#9 (logbook*):ti,ab,kw
#10 (consultation):ti,ab,kw
#11 (selected conditions):ti,ab,kw
#12 (diseases management):ti,ab,kw
#13 (clinical method*):ti,ab,kw
#14 (diagnosis cluster*):ti,ab,kw
#15 (distribution patient*):ti,ab,kw
#16 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR
#7 OR #8 OR #9 OR #10
OR #11 OR #12 OR #13 OR #14 OR #15)
#17 MeSH descriptor Curriculum explode all trees
#18 (curricul*):ti,ab,kw
#19 MeSH descriptor Education, Medical

4. ERIC

1 exp "Case Method (Teaching Technique)"/
2 exp Clinical Experience/ or clinical exposure.mp.
3 clinical encounter.mp.
4 logbook*.ti,ab.
5 exp Patients/
6 (case mix* or casemix*).ti,ab.
7 1 or 2 or 3 or 4 or 5 or 6
8 exp Curriculum/
9 curricul*.ti,ab.
10 exp Medical Education/
11 medical education.ti,ab.
12 clinical competence.mp.
13 exp "Clinical Teaching (Health Professions)"/
14 clerkship*.ti,ab.
15 exp Trainees/
16 trainee*.ti,ab.
17 residen*.ti,ab.
18 work based learning.mp.
19 11 or 9 or 17 or 12 or 15 or 14 or 8 or 18 or 16 or
10 or 13
20 exp Medical Education/
21 teaching hospital*.ti,ab.
22 exp Primary Health Care/
23 exp Medical Students/
24 22 or 21 or 23 or 20
25 24 and 7 and 19
26 *medical education/
27 22 or 21 or 26 or 23
28 27 and 7 and 19
5. WEB OF SCIENCE
Title=(“family practice” OR “general pract*” OR “family medicine” OR “primary care” OR “internal medicine” of psychiatr* OR “hospital*” or surgery) AND Title=(curriculum or training* OR trainee* OR clerks* OR residen* OR education* OR learn* OR medical student* OR internship* OR work based learning) AND Title=(“case mix” OR “casemix” OR “experience*” or disease* OR logbook* OR “patient mix” OR examination* OR patient* OR diagnos* OR condition*)