The remediation challenge: theoretical and methodological insights from a systematic review

Jennifer Cleland,[1] Heather Leggett,[2] John Sandars,[2] Manuel J Costa,[3] Rakesh Patel,[2,4] & Mandy Moffat,[1]

OBJECTIVES Remediation is usually offered to medical students and doctors in training who underperform on written or clinical examinations. However, there is uncertainty and conflicting evidence about the effectiveness of remediation. The aim of this systematic review was to synthesise the available evidence to clarify *how* and *why* remediation interventions may have worked in order to progress knowledge on this topic.

METHODS The MEDLINE, EMBASE, CINAHL (Cumulative Index to Nursing and Allied Health Literature), ERIC (Educational Resources Information Centre), Web of Science and Scopus databases were searched for papers published from 1984 to April 2012, using the search terms 'remedial teaching', 'education', 'medical', 'undergraduate'/or 'clinical clerkship'/or 'internship and residency', 'at risk' and 'struggling'. Only studies that included an intervention, then provided retest data, and reported at least one outcome measure of satisfaction, knowledge, skills or effects on patients were eligible for inclusion. Studies of practising doctors were excluded. Data were abstracted independently in duplicate for all items. Coding differences were resolved through discussion.

RESULTS Thirty-one of 2113 studies met the review criteria. Most studies were published after 2000 (n = 24, of which 12 were published from 2009 onwards), targeted medical students (n = 22) and were designed to improve performance on an immediately subsequent examination (n = 22). Control or comparison groups, conceptual frameworks, adequate sample sizes and long-term follow-up measures were rare. In studies that included long-term follow-up, improvements were not sustained. Intervention designs tended to be highly complex, but their design or reporting did not enable the identification of the active components of the remedial process.

CONCLUSIONS Most remediation interventions in medical education focus on improving performance to pass a re-sit of an examination or assessment and provide no insight into what types of extra support work, or how much extra teaching is critical, in terms of developing learning. More recent studies are generally of better quality. Rigorous approaches to developing and evaluating remediation interventions are required.

*Medical Education 2013: **47**: 242–251*
doi:10.1111/medu.12052

Discuss ideas arising from this article at www.mededuc.com 'discuss'

[1]Division of Medical and Dental Education, University of Aberdeen, Aberdeen, UK
[2]Leeds Institute of Medical Education, University of Leeds, Leeds, UK
[3]Life and Health Sciences Research Institute (ICVS), School of Health Sciences, University of Minho, Braga, Portugal
[4]Department of Medical and Social Care Education, Leicester Medical School, University of Leicester, Leicester, UK

Correspondence: Jennifer Cleland, Division of Medical and Dental Education, University of Aberdeen, West Wing Polwarth Building, Foresterhill, Aberdeen AB25 2AZ, UK. Tel: 00 44 1224 437257; E-mail: jen.cleland@abdn.ac.uk

INTRODUCTION

Determining fitness for practice in medicine and the health care professions requires students to fulfil predetermined criteria laid down by the relevant professional body and interpreted by each individual educational institution providing pre-registration education. The objective of these criteria is to ensure safe practice. However, a small proportion of medical and health care students perform poorly on clinical or academic examinations. This is of concern to all medical educators. Research has identified that measures of attainment in medical degree examinations can predict subsequent performance on licensing examinations and clinical competence after medical school.[1,2] In addition, weak medical students go on to become weak doctors,[3,4] further highlighting the importance of early identification of underperformance. These issues are not unique to the USA. Our own studies show that UK medical students on an undergraduate course who failed clinical or written examinations in the early years continued to perform poorly until and in their final examinations.[5,6]

However, the complex patterns of assessment in medicine mean that struggling students may continue with little guidance or support[7] and supervising clinicians are often reluctant to fail underperformance.[8–10] Thus, students' learning problems remain unaddressed, leading to repeated failure and underperformance.[5,6,11] Timely intervention for poor performance has the potential to enable the individual to deal with adverse learning and behaviour patterns promptly before these cause problems in clinical practice.

Most medical schools have remediation processes, developed usually on the basis of, for example, staff availability and interest, and the nature of students' difficulties, and flexible enough to be tailored to student needs. Most remediation processes consist of three steps: identification or diagnosis; a remediation intervention, and retesting.[12–14] However, these remediation processes place substantial time demands on faculty staff.[7,13] Furthermore, faculty members report uncertainty about the efficacy of remediation interventions.[13] That this uncertainty is well founded is supported by recent evidence. Pell *et al.*[15] carried out a retrospective analysis of the longitudinal attainment of 125 students who underperformed in objective structured clinical examinations (OSCEs). They found that a remediation intervention (assessment-focused revision followed by re-assessment) was not effective in terms of subsequent examination performance: in fact, the relative performance of the student group under study declined across serial OSCEs. Although these data were derived from one medical school only, the remediation intervention described is typical.[12–14]

These findings are at odds with those of a number of studies that have evaluated performance on written or clinical examinations before and after a specific remediation intervention[7,14–38] and concluded that the intervention was effective. However, single studies 'are limited in the generalisability of the knowledge they produce about concepts, populations, settings and times' and 'frequently illuminate only one part of a larger explanatory puzzle'.[39] In addition, they are subject to limitations imposed by time-, sample- and context-specificity, which can undermine their applicability, relevance and usefulness in other contexts.[40] Given the conflicting evidence, resource implications and staff concerns that refer to remediation, it is critical to carry out a research synthesis of the literature on remediation interventions in medicine in order to identify consistencies, variability and generalisability across studies. The cumulative knowledge to be gained by doing so is essential to inform the quality (and hence effectiveness in terms of the long-term gain of producing doctors who are fit for practice) of remediation research and processes.[41]

A thematic review of the remediation literature was published in 2009.[42] It focused on 13, mostly North American (n = 11), studies of below-standard performance of students, trainees or practising doctors. This review was descriptive and study quality was assessed solely in terms of Kirkpatrick's hierarchy of evaluation.[43,44] By contrast, our aim was to clarify *how* or *why* remediation interventions may have worked[45] in order to progress knowledge on this topic, particularly given that much work in this area has been published since Hauer *et al.* published their review.[42]

To achieve this, our specific research questions were:

- Can the various interconnecting parts of remedial interventions for medical students and doctors in training be delineated to identify precisely what works and why?[46]
- What theoretical frameworks are used in studies of remedial interventions for medical students and doctors in training?

Only by addressing these questions can we build a cumulative understanding of causal mechanisms,

design more effective interventions and apply them appropriately across different groups and settings.[47]

METHODS

Study eligibility and selection

The types of evidence to be included in the review included reports (scientific articles, published letters) assessing the effectiveness of a faculty-led remediation intervention to address underperformance on clinical or written examinations in medical students or trainee doctors. Only reports which included an intervention and provided retest data for the students or doctors to assess whether they had achieved the required standard to enable progress, and reported at least one qualitative or quantitative outcome measure of satisfaction, knowledge, skills or effects on patients were eligible for inclusion. Exclusion criteria denied the inclusion of studies in practising doctors.

We searched the MEDLINE, EMBASE, CINAHL (Cumulative Index to Nursing and Allied Health Literature), ERIC (Educational Resources Information Centre), Web of Science and Scopus databases for articles published from 1984 to September 2011, using pre-specified search terms (Appendix S1). The search was extended to April 2012 to identify any articles published after the main search concluded. To ensure comprehensiveness, we also hand-searched the reference lists of the full-text articles assessed for eligibility (Fig. 1) and used the PubMed e-mail alert system to identify articles with 'remediation' in the title. We also corresponded personally with researchers known to be active in this area of research to identify unpublished work and work in submission.

Data extraction

We used the TREND (*t*ransparent *r*eporting of *e*valuations with *n*on-randomised *d*esigns) checklist[48] to guide data extraction. This was selected as it emphasises the reporting of theories used and descriptions of intervention and comparison conditions, and research design, in evaluation studies that use non-randomised designs, and so was felt to be appropriate for review given the nature of the research studies on the topic. We also collected data on reported process outcomes. Using these parameters, we developed and refined a data

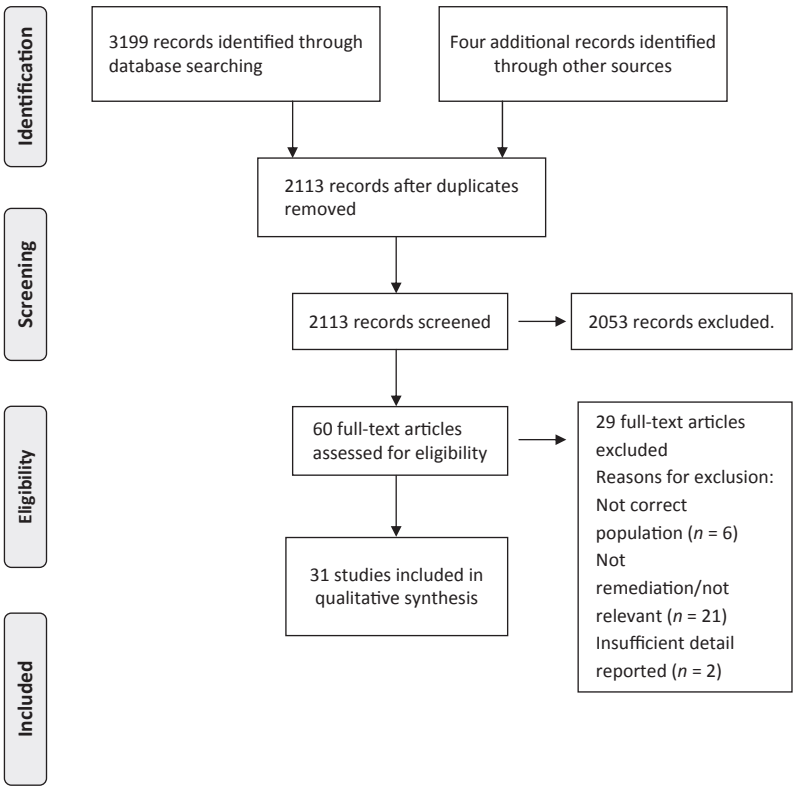

Figure 1 Flow diagram of records identified, included and excluded, and reasons for exclusions, April 2012

Table 1 Summary of findings across 31 studies of remediation interventions in medical education

Articles, *n*	31					
Articles per country, *n* (%)	USA (*n* = 20, 65%)	USA offshore (*n* = 3, 10%)	UK (*n* = 5, 16%)	Canada (*n* = 1, 3%)	New Zealand (*n* = 1, 3%)	Netherlands (*n* = 1, 3%)
Students per study	Median: 23 (total across all studies: 1614*)					
Level of training*	Junior medical students (*n* = 950)	Senior medical students (*n* = 368)	Mixed junior and senior students (*n* = 165)	Residents (*n* = 125)	Not noted (*n* = 6)	
Focus according to need*	Knowledge (*n* = 971)	Skills (*n* = 320)	Knowledge and skills (*n* = 258)	Knowledge, skills and attitudes (*n* = 28)	Information not provided (*n* = 37)	
Theoretical basis of the remediation intervention, studies, *n*	None stated (*n* = 23)	Learner-centred approaches (*n* = 2)	Kolb cycle (*n* = 1)	Multi-faceted (*n* = 4)		

None stated = 23 Self-regulatory (cognitive) theory = 4 Learner-centred approaches = 2 Kolb cycle = 1 Self-assessment = 1
* Based on 29 reports because two did not report participant numbers
Table S1 [online] shows full details of the studies reviewed

abstraction form through iterative pilot testing and revision.

We abstracted data independently in duplicate for all items (coded by HL and MM). Any discrepancies were discussed and agreed upon. The information was then entered into an Excel database.

Data synthesis and analysis

This was a two-stage review in which the making of a descriptive map of the research was followed by a narrative empirical synthesis. A qualitative synthesis of the data was selected as this method was deemed appropriate in the context of our research questions (Evidence for Policy and Practice Information and Co-ordinating Centre [EPPI-Centre; http://eppi.ioe.ac.uk/cms][49]).

Reviews are driven by the questions they seek to answer and by reviewers who may interpret the issues from different ideological and theoretical perspectives. Those involved in the present review are either medical educators with a recognised interest in identifying and addressing the issue of remediation (JC, JS, MJC, RP) or social scientists working in the field of medical education (MM, HL). We explicitly acknowledged our subjectivity and used the principles of cooperative enquiry (i.e. discussing findings, and critically reflecting and expanding on them[50]) on an ongoing basis to address this within the group. This was accomplished by maintaining an audit trail of the developing interpretation via face-to-face meetings, and telephone and e-mail correspondence.

RESULTS

Of the 2113 studies found, we identified 31 studies eligible for inclusion through duplicate review of titles, abstracts and full texts, and personal communication with authors (Fig. 1).

Table 1 summarises information on location and setting, target population, study design, number of participants, focus of the intervention, as well as theoretical framework, summary of the intervention and outcome measures. (A full descriptive map of the reviewed studies is available in Table S1 [online]. This includes an analysis of quality in terms of whether or not the information provided about the

intervention was sufficient to identify the critical components.)

Remediation in medical education appears to be an increasingly important area of research: the majority (n = 23, 74%) of studies were published after 2000.[7,14–17,19–22,24,27–31,33,34,37,38,51–54] and 11 of these were published from 2009 onwards.[14–16,20,29,31,34,38,52–54] We also identified one unpublished study conducted since 2009.[55]

Most of the studies (n = 24, 77%) had been conducted in North America or in offshore medical schools.[14,16–19,23–34,36–38,51–54] Five studies were from the UK,[7,15,20–22] one was from New Zealand[34] and one was from the Netherlands.[55] Non-English language studies were not excluded by the review criteria, but none were identified.

The majority of studies targeted medical students (n = 21, 68%).[7,14,15,19–22,24–26,29,31–33,35,36,38,51–54] Of these, a number targeted early years (Years 1 and 2) students[25,26,38,53,54] or included students from these year groups.[22,36] However, Year 4 students represented the year group most commonly targeted or included in interventions.[7,14,15,19–22,24,29,33,51] In the UK study settings, Year 4 students have one further year of study before pre-registration (i.e. the medical school programmes in the studies reported are 5 years in length), whereas in the USA and Canada, Year 4 is the final year of study before licensure.

The nine studies of doctors in training were from North America.[16–18,23,27,28,30,34,37] The most common specialty under study was surgery.[17,27,28,30,34]

Studies tended to target knowledge (n = 16, 52%)[16,17,21,23,26–28,30–32,35,37,38,53,54,55] or skills (n = 10, 32%)[14,15,18,19,22,24,33,34,51,52] deficits. Five (16%) targeted both knowledge and skills deficits[7,20,21,23,36] and one study did not specify what was specifically targeted in terms of knowledge, skills or attitudes.[25]

Most studies included very small sample numbers, ranging from six to 377, with a median of 23. Only one study stated a sample size calculation.[55]

Most studies were either prospective before-and-after studies (n = 18)[7,14,17–19,22–24,26,29,30,32–34,37,38,51,52] or retrospective before-and-after studies (n = 7).[15,16,20,25,27,28,35] Other study designs used were retrospective case series (n = 1),[31] qualitative or survey methodology (n = 3)[21,53,54] and a parallel-group randomised controlled trial (n = 1).[55] It was unclear what design was used in one study.[36] Nine (29%) prospective before-and-after studies looked at one cohort or year group only.[14,19,24,26,33,34,37,51,52] Seven studies (23%) included a comparison group of students who had not failed the index examination.[15–17,20,26,37,38] One study used a historical control group of students from previous cohorts who had failed the index examination and repeated the first semester without receiving the remedial intervention.[38] Two studies used a two-stage intervention in which all underperforming students received one course, but the 'intervention' group received more input (e.g. a remediation programme with or without faculty staff feedback[14] and 'standard academic support' with or without a study skills programme.[55] The study[55] randomly assigned students to the intervention and control groups, whereas that by White *et al.*[14] assigned students to groups according to their level of underperformance. One study used a comparison group of students who had failed the index examination but did not receive the intervention.[26]

Most (n = 22, 71%) studies focused solely on performance on a specific, subsequent examination (e.g. resits, re-taking a standard examination such as the US Medical Licensing Examination, or the next standard examination in a programme).[14–22,24,26–30,32–35,37,51,52] Indeed, many studies overtly focused on examination technique and content boosting for a specific examination (which the participants were required to re-sit or re-take to progress their education or training). For example, Aeder *et al.*[16] stated that 'the last 40 minutes of the class were devoted to going over applicable questions published by the American Academy of Paediatrics' and Shokar[37] stated that 'examination techniques were taught to and practised by the group'. Statements of this nature were identified in eight studies, all of which were carried out in North America.[16,17,26–28,30,32,37] Remediation was noted to have no significant impact on subsequent examination performance in one of these studies.[37]

The remaining eight studies (26%) took a more holistic perspective to remediation and either included the provision of personal support or took a broader approach to learning and teaching skills and knowledge.[7,23,25,31,36,38,53,54] Of these, seven reported a long-term positive effect of the intervention in that most of those who received remediation progressed in their studies.[23,25,31,36,38,53,54] By contrast, Pell *et al.*[15,55] reported no longer-term improvement following the intervention and indeed Pell *et al.*[15] found

that student performance deteriorated after passing the target examination.

Only eight studies (26%) referred explicitly to educational theory[14,18,19,30,38,53,54,55] but note that three[38,53,54] of these studies were linked. The theories used included a learner-centred approach,[18,19] self-assessment in learning,[14] cognition and metacognition[38,53,54,55] and the Kolb learning cycle.[30] Only Winston *et al.*[53] explicitly – and extensively – mapped out the components of their intervention onto learning theory (note that Stegers-Jager *et al.* [2012] based their syllabus and materials on those of Winston *et al.*, 2010[53]).

Studies were assessed for complexity.[46] An intervention was defined as complex if it utilised several interacting components. Dimensions of complexity can include, for example, the number of and interactions between components, number and variability of outcomes and the degree of flexibility or tailoring of the intervention permitted. The majority of the studies reviewed did not report in detail what they did, why and for how long. Most interventions reported a variety of activities (e.g. tutorials, directed reading, skills practice, feedback, examination practice, case presentations), but few studies clearly justified their approaches on the basis of appropriate theory (with the notable exceptions of [14,18,19,30,53]). Many studies simply stated, for example, that the programme elements included mentoring or tutorials without describing the content or format of, or rationale for, these tutorials or mentoring sessions.[17,34,35] It was rare for authors to even consider the issue of complexity (but see [27,32,54]). For example, Winston *et al.*[54] considered the role of the teacher in the process of an intervention.

Where reported, intervention length varied widely and included periods ranging from a single full week,[22] to over 4 weeks (three studies[29,33,37]), 5 weeks[55], 6 weeks (three studies[18,26,32]) and 1 year (five studies[7,16,23,25,27]). Intervention intensity was not reported in terms of theoretical rationale: where any rationale was provided, it was pragmatic.

The evaluation of process as well as outcome is recommended in evaluating the effectiveness of interventions in other areas.[46] Twelve studies (39%) collected data on the perceived usefulness of and satisfaction with the remedial intervention from recipients.[7,18,19,21,23,25,27,33,36,51,53,54] Four of these studies collected similar data from faculty staff.[7,21,23,54] Sanfey *et al.*[34] and Winston *et al.*[38] stated that they believed that changing their remediation from voluntary to mandatory may have motivated residents to improve their practice.

DISCUSSION

This review established that the addressing of underperformance in medical students or doctors in training is an active area of primary research, but the majority of studies identified would be classed as being of low quality according to the criteria for grading quality of evidence.[39] The evidence comes predominantly from uncontrolled before-and-after studies with small samples and few process or long-term outcome measures, which may not convincingly distinguish intervention effectiveness from background effects or the Hawthorne effect.[56,57] As a consequence, we cannot delineate precisely what works, and why, in remedial interventions for medical students and doctors.

The issue of complexity is clear. The designs and methods of the studies reviewed, and the lack of detail reported on the precise nature of many of the interventions, do not allow us to identify which components of the process actually made a difference. We do not know, for example, if underperforming students do better on retesting purely because their motivation is increased (fright at having failed first time around may be a great motivator) and if their improvement has absolutely nothing to do with the remediation intervention.

Generally, interventions tended to represent 'more of the same', such as additional or intensive knowledge or skills teaching. If the original teaching did not help students to learn appropriately, there seems little reason to assume that 'more of the same' will do so a second time around, even if this is delivered in small groups by senior faculty staff. It may be that the critical factor in (short-term) improved performance is individual analysis of performance and feedback,[58,59] which were inherent in many of the studies reviewed. However, without more rigorous study designs, this cannot be assumed.

Most interventions were tailored to improve performance to the standard required to pass a re-sit or re-take rather than to support the development of effective lifelong learning skills. The ethics of supporting students to progress to the next stage of training only to continue to perform poorly (e.g. [6,15,38]) are, at best, questionable. It is also debatable whether scarce faculty resources should be used to support progression without improvement, which

may take weak students further towards registration as potentially weak doctors,[3,4] when the evidence suggests that faculty members find it harder to fail senior students.[8]

Moreover, we do not know what types of extra support work, or how much extra teaching is critical. The process measure of 'satisfaction', where employed, does not add much to our understanding of barriers, facilitators or what precisely works. However, those studies that explicitly attempted to tease out what makes a difference[14,38] are a welcome addition to the literature, hinting as they do that particular subgroups of students (e.g. those who attend when the intervention is not mandatory,[38] those who are slightly better performers[55], those who can accurately self-assess,[14] and possibly those taught by experienced teachers[54]) may respond best to remediation. One recent study, not included in this review, suggests that organisational factors, such as the implementation of an academic dismissal policy, may also be influential.[60] These findings are unsurprising. Firstly, it has long been recognised that reasons for poor performance are myriad; that is, poor performers are not a homogeneous group. Secondly, evidence from complex interventions in clinical areas has recently highlighted that these have different effects in different subgroups.[61] In identifying differences in subgroups and hinting at potential teacher influences, these studies raise just as many questions as they answer – the list of possible active components is long, if not endless[46] – but examining this in detail will progress knowledge on this topic and provide further information on how best to use scarce faculty resources.

We also wished to examine the extent to which interventions were underpinned by theory. The review indicates that there is a growing trend for studies to use theory to conceptualise their interventions. The most widely used theoretical framework was broadly cognitive, using self-regulation, metacognition and reflection, and the giving and receiving of feedback.[14,38] This seems very appropriate: education research has indicated quantitative and qualitative differences in regulation processes and activities between weak and strong learners (e.g. [62]). Highly self-regulated learners are academically more successful than those students with low levels of skill in self-regulated learning (SRL) or those who lack regulation in their learning (e.g. [63]). We cannot assume that students can self-regulate when they enter medical school: indeed, an exploratory study suggests that differences in SRL in successful and unsuccessful medical student learners are identifiable.[64] Thirty years of education research has identified that explicit training in SRL techniques is effective[65,66] in terms of improving learning outcomes for students. We suggest that this framework could make important contributions to traditional medical training assessment frameworks that have been used to identify and remediate strugglers (see also[67–69]). The provision of 'learning to learn' courses for Year 1 medical students may provide an effective approach to helping students at an early stage of their medical careers to identify their SRL approach and make changes that might reduce their chances of future underperformance.[70] Further research is indicated.

Most remediation efforts are targeted at learners in the latter years of medical school. However, early remediation interventions have the potential to stop the cycle of underperformance that is characteristic of many struggling students. Struggling students have low self-efficacy beliefs and negative feelings about learning that directly influence their motivation to persist with difficult learning tasks.[71] These students need to experience success as soon as they are identified as struggling so that they can feel a sense of control over their learning and performance. Moreover, given that the aim of a remediation intervention should be to ensure safe practice, developing defensible systems for identifying and addressing underperformance may empower medical schools to advise students who fail, and fail again even after intervention, to seek alternative careers.

To the best of our knowledge, this is the first systematic review of the literature on remediation in medical education and the first review in this area to have focused on *how* or *why* remediation may have worked.[39] We included published letters and unpublished literature in this review, but not conference abstracts, and therefore our findings may be subject to publication bias and the under-reporting of negative findings (although this hypothesis is not supported by the data presented). The process of reviewing was strengthened by numerous checks by several authors. Our findings differed from those of a thematic review reported in 2009,[42] partly as a result of the many studies published in recent years, most of which were conducted in the setting of undergraduate medical education, and as a result of the different focus of this review. A potential weakness of this study is that we did not systematically search the literature to explore how professions allied to medicine tackle underperformance. Although the evidence suggests that self-reported satisfaction is unlikely to be a particularly useful measure, we took the decision to

include the studies (n = 4) which reported student or faculty staff views of the remediation intervention or an aspect of the intervention as their only outcome measure in order to provide a fuller picture of the literature on this topic. We did not report on the psychometric adequacy of the assessments used to identify underperformance and re-assess performance and thus we cannot comment on the reliability or validity of these measures. We excluded studies involving practising doctors for pragmatic reasons, but are aware that this area (e.g. effectiveness of continuing medical education, role of maintenance of certification and maintenance of licensure programmes) is also in need of research.

The data presented here and in the wider literature suggest that changing the focus of teaching, learning and assessment to include assessment of the actual processes of learning would progress work in this area. Doing so should enable the accurate diagnosis of underperformance and the early identification of the reasons for its occurrence, as well as providing a sound theoretical basis for remediation that focuses on 'learning to learn' rather than 'examination coaching'. Although we acknowledge the roles of local context[46] and individual factors in underperformance, multi-institution approaches to remediation would provide insight into the generalisability of interventions and allow more rigorous, controlled study designs, which are necessary to determine cause–effect relationships. Adopting complex intervention models, such as those recommended by the Medical Research Council (MRC) in the UK would enable the identification and evaluation of the key components of any intervention, progressing knowledge of what does and does not work. Finally, and critically, moving the focus of remediation interventions to one of 'learning to learn' shifts the outcome measures from immediate examination performance to performance as a doctor in training.

In summary, this review has established the current state of knowledge on the conceptual and methodological bases of remedial interventions in medical education. In doing so, it has identified that all is not doom and gloom: more recent studies are of much better quality, particularly in terms of the explicit use of pedagogic or theoretical frameworks to underpin intervention design, often with the inclusion of longer-term follow-up. This review has identified new research questions in this area and will serve as a tool to stimulate debate among those involved in identifying and addressing underperformance in medical education.

Contributors: JS conceived the original idea for the study. HL and MM carried out the searches and HL prepared the evidence map, under the supervision of JC and JS. Using the evidence map for guidance, JC analysed the data and prepared the first draft of the manuscript. RP and MJC commented on the evidence mapping exercise. All authors contributed to the critical revision of the paper and approved the final manuscript for publication.
Acknowledgements: the authors thank Kalman Winston for his review of the manuscript before submission, and Karen Stegers-Jager and Kalman Winston for permission to include at the time unpublished articles in the review.
Funding: none.
Conflicts of interest: none.
Ethical approval: not applicable.

REFERENCES

1 Gonnella JS, Erdmann JB, Hojat M. An empirical study of the predictive validity of number grades in medical school using three decades of longitudinal data: implications for a grading system. *Med Educ* 2004;**38** (4):425–34.
2 Hojat M, Gonnella JS, Veloski JJ, Erdmann JB. Is the glass half full or half empty? A re-examination of the associations between assessment measures during medical school and clinical competence after graduation. *Acad Med* 1993;**68** (2 Suppl):69–76.
3 Challis A, Flett G, Batstone M. An accident waiting to happen? A case for medical education. *Med Teach* 1999;**21** (6):582–5.
4 Papadakis MA, Teherani A, Banach MA, Knettler TR, Rattner SL, Stern DT, Veloski JJ, Hodgson CS. Disciplinary action by medical boards and prior behaviour in medical school. *N Engl J Med* 2005;**353** (25):2673–82.
5 Cleland J, Arnold R, Chesser A. Failing finals is often a surprise for the student but not the teacher: identifying difficulties and supporting students with academic difficulties. *Med Teach* 2005;**27** (6):504–8.
6 Cleland JA, Milne A, Sinclair H, Lee AJ. Cohort study on predicting grades: is performance on early MBChB assessments predictive of later undergraduate grades? *Med Educ* 2008;**42** (7):676–83.
7 Sayer M, Chaput de Saintonge M, Evans D, Wood D. Support for students with academic difficulties. *Med Educ* 2002;**36** (7):643–50.
8 Cleland JA, Knight LV, Rees CE, Tracey S, Bond CM. Is it me or is it them? Factors that influence the passing of underperforming students. *Med Educ* 2008;**42** (8):800–9.
9 Dudek NL, Marks MB, Regehr G. Failure to fail: the perspectives of clinical supervisors. *Acad Med* 2005;**80** (10 Suppl):84–7.
10 Speer AJ, Solomon DJ, Fincher RM. Grade inflation in internal medicine clerkships: results of a national survey. *Teach Learn Med* 2000;**12** (3):112–6.

11 Tooth D, Tonge K, McManus IC. Anxiety and study methods in pre-clinical students: causal relation to examination performance. *Med Educ* 1989;**23** (5): 416–21.
12 Frellsen SL, Baker EA, Papp KK, Durning SJ. Medical school policies regarding struggling medical students during the internal medicine clerkships: results of a national survey. *Acad Med* 2008;**83** (9):876–81.
13 Hauer KE, Teherani A, Irby DM, Kerr KM, O'Sullivan PS. Approaches to medical student remediation after a comprehensive clinical skills examination. *Med Educ* 2008;**42** (1):104–12.
14 White CB, Ross PT, Gruppen LD. Remediating students' failed OSCE performances at one school: the effects of self-assessment, reflection, and feedback. *Acad Med* 2009;**84** (5):651–4.
15 Pell G, Fuller R, Homer M, Roberts T. Is short-term remediation after OSCE failure sustained? A retrospective analysis of the longitudinal attainment of underperforming students in OSCE assessments. *Med Teach* 2012;**34** (2):146–50.
16 Aeder L, Fogel J, Schaeffer H. Paediatric board review course for residents 'at risk'. *Clin Pediatr (Phila)* 2010;**49** (5):450–6.
17 Borman KR. Does academic intervention impact ABS qualifying examination results? *Curr Surg* 2006;**63** (6):367–72.
18 Brooks R, Robb A, Tabak D. A remedial course on interviewing skills. *Acad Med* 1996;**71** (5):560.
19 Chou CL, Chang A, Hauer KE. Remediation workshop for medical students in patient–doctor interaction skills. *Med Educ* 2008;**42** (5):537–8.
20 Cleland J, Mackenzie RK, Ross S, Sinclair HK, Lee AJ. A remedial intervention linked to a formative assessment is effective in terms of improving student performance in subsequent degree examinations. *Med Teach* 2010;**32** (4):185–90.
21 Denison AR, Currie AE, Laing MR, Heys SD. Good for them or good for us? The role of academic guidance interviews. *Med Educ* 2006;**40** (12):1188–91.
22 Dowell J, Dent JA, Duffy R. What to do about medical students with unsatisfactory consultation skills? *Med Teach* 2006;**28** (5):443–6.
23 Edeiken BS. Remedial programme for diagnostic radiology residents. *Invest Radiol* 1993;**28** (3):269–74.
24 Faustinella F, Orlando PR, Colletti LA, Juneja HS, Perkowski LC. Remediation strategies and students' clinical performance. *Med Teach* 2004;**26** (7):664–5.
25 Ficklin FL, Hazelwood JD, Carter JE, Shellhamer RH. The reduced load as a remedial programme to increase retention of first-year medical students. *J Med Educ* 1985;**60** (5):406–8.
26 Frierson HT Jr. Combining test-taking intervention with course remediation: effects on National Board subtest performance. *J Natl Med Assoc* 1987;**79** (2):161–5.
27 Gregg SC, Eisenberg D, Duffy AJ, Longo WE. Design, management, and critical evaluation of a surgical basic/clinical science curriculum: the role of an educational chief resident. *J Surg Educ* 2008;**65** (1):36–42.
28 Harthun NL, Schirmer BD, Sanfey H. Remediation of low ABSITE scores. *Curr Surg* 2005;**62** (5):539–42.
29 Klamen DL, Williams RG. The efficacy of a targeted remediation process for students who fail standardised patient examinations. *Teach Learn Med* 2011;**23** (1): 3–11.
30 Kosir MA, Fuller L, Tyburski J, Berant L, Yu M. The Kolb learning cycle in American Board of Surgery In-Training Exam remediation: the Accelerated Clinical Education in Surgery course. *Am J Surg* 2008;**196** (5):657–62.
31 Laatsch L. Evaluation and treatment of students with difficulties passing the Step examinations. *Acad Med* 2009;**84** (5):677–83.
32 Magarian GJ, Campbell SM. A tutorial for students demonstrating adequate skills but inadequate knowledge after completing a medicine clerkship at the Oregon Health Sciences University. *Acad Med* 1992;**67** (4):277–8.
33 Majdan JF, Berg K, Berg D, Veloski J, Rattner S. *Development of a Course for Students Who Fail a High-Stakes OSCE*. Pittsburgh, PA: Society of General Internal Medicine 2008:B5.3.
34 Sanfey H, Ketchum J, Bartlett J, Markwell S, Meier AH, Williams R, Dunnington G. Verification of proficiency in basic skills for postgraduate year 1 residents. *Surgery* 2010;**148** (4):759–66; discussion 766–7.
35 Schwartz PL, Loten EG. Effect of remedial tutorial help on students who fail in-course assessments. *Acad Med* 1998;**73** (8):913.
36 Segal SS, Giordani B, Gillum LH, Johnson N. The Academic Support Program at the University of Michigan School of Medicine. *Acad Med* 1999;**74** (4):383–5.
37 Shokar GS. The effects of an educational intervention for 'at-risk' residents to improve their scores on the In-training Exam. *Fam Med* 2003;**35** (6):414–7.
38 Winston KA, van der Vleuten CP, Scherpbier AJ. An investigation into the design and effectiveness of a mandatory cognitive skills programme for at-risk medical students. *Med Teach* 2010;**32** (3):236–43.
39 Cook TD, Cooper HM, Cordray DS, Hedges LV, Hartmann H, Light RJ, Louis TA Mosteller F eds. *Meta-analysis for Explanation: A casebook*. New York, NY: Russell Sage 1992.
40 Davies P. The relevance of systematic reviews to educational policy and practice. *Oxford Rev Educ* 2000;**26** (3–4):365–78.
41 Eva K. Broadening the debate about quality in medical education research. *Med Educ* 2009;**43**:294–6.
42 Hauer KE, Ciccone A, Henzel TR, Katsufrakis P, Miller SH, Norcross WA, Papadakis MA, Irby DM. Remediation of the deficiencies of physicians across the continuum from medical school to practice: a thematic review of the literature. *Acad Med* 2009;**84** (12):1822–32.

43 Kirkpatrick DL. Evaluation of training. In: Craig R, Mittel I, eds. *Training and Development Handbook*. New York, NY: McGraw Hill 1967;87–112.
44 Yardley S, Dornan T. Kirkpatrick's levels and education 'evidence'. *Med Educ* 2012;**46** (1):97–106.
45 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** (2):128–33.
46 Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths Fea. Designing and evaluating complex interventions to improve health care. *BMJ* 2007;**334**:455–9.
47 Michie S, Abraham C. Interventions to change health behaviours: evidence-based or evidence-inspired? *Psychol Health* 2004;**19** (1):29–49.
48 des Jarlais DC, Lyles C, Crepaz N. Improving the reporting quality of non-randomised evaluations of behavioural and public health interventions: the TREND statement. *Am J Public Health* 2004;**94**: 361–6.
49 Oakley A. Resistances to new technologies of evaluation: education research in the UK as a case study. *Evid Policy* 2006;**2** (1):61–86.
50 Reason P. The practice of cooperative inquiry. *Syst Pract Action Res* 2002;**15**:169–76.
51 Chang A, Chou CL, Hauer KE. Clinical skills remedial training for medical students. *Med Educ* 2008;**42** (11):1118–9.
52 DeMuth R, Phillips J, Wagner D. An OSCE remediation experience focused on diagnostic reasoning. *Med Teach* 2011;**33** (7):592–3.
53 Winston KA, van der Vleuten CPM, Scherpbier AJJA. At-risk medical students: implications of students' voice for the theory and practice of remediation. *Med Educ* 2010;**44** (10):1038–47.
54 Winston KA, van der Vleuten CP, Scherpbier AJ. The role of the teacher in remediating at-risk medical students. *Med Teach* 2013;**34**:e732–42.
55 Stegers-Jager KM, Cohen-Schotanus J, Themmen AP. *Motivation, learning strategies, participation and medical school performance.. Med Educ* 2012;**46**:678–88.
56 Cook TD, Campbell DT. *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Chicago, IL: Rand McNally 1979.
57 Lipsey MW, Wilson DB. The efficacy of psychological, educational and behavioural treatment. Confirmation from meta-analysis. *Am Psychol* 1993;**48**:1181–209.
58 Ende J. Feedback in clinical medical education. *JAMA* 1983;**250** (6):777–81.
59 Sinclair HK, Cleland JA. Undergraduate medical students: who seeks formative feedback? *Med Educ* 2007;**41** (6):580–2.
60 Stegers-Jager KM, Cohen-Schotanus J, Splinter TAW, Themmen APN. Academic dismissal policy for medical students: effect on study progress and help-seeking behaviour. *Med Educ* 2011;**45**:987–94.
61 Bucknall CE, Miller G, Lloyd SM, Cleland J, McCluskey S, Cotton M, Stevenson RD, Cotton P, McConnachie A. Glasgow supported self-management trial (GSuST) for patients with moderate to severe COPD: randomised controlled trial. *BMJ* 2012;**344**:1060.
62 Schunk DH, Zimmerman BJ. *Self-regulated Learning: From Teaching to Self-reflective Practice*. New York, NY: Guilford 1998.
63 Heikkila A, Lonka K. Studying in higher education: students' approaches to learning, self-regulation, and cognitive strategies. *Stud High Educ* 2006;**31** (1):99–117.
64 Cleary TJ, Sandars J. Assessing self-regulatory processes during clinical skill performance: a pilot study. *Med Teach* 2011;**33** (7):368–74.
65 Dignath C, Buttner G. Components of fostering SRL among students. A meta-analysis of intervention studies at primary and secondary school level. *Metacogn Learn* 2008;**3**:231–64.
66 Hattie J, Biggs J, Purdie N. Effects of learning skills interventions on students learning: a meta-analysis. *Rev Educ Res* 1996;**66**:99–136.
67 Durning SJ, Cleary TJ, Sandars J, Hemmer P, Kokotailo P, Artino AR. Perspective: viewing 'strugglers' through a different lens: how a self-regulated learning perspective can help medical educators with assessment and remediation. *Acad Med* 2011;**86** (4):488–95.
68 Sandars J, Cleary TJ. Self-regulation theory: applications to medical education: AMEE Guide No. 58. *Med Teach* 2011;**33** (11):875–86.
69 White CB, Gruppen LD. Self-regulation learning in medical education. In: Swanwick T, ed. *Understanding Medical Education*. Chichester: Wiley-Blackwell 2010;271–82.
70 Sandars J. Pause 2 Learn: developing self-regulated learning. *Med Educ* 2010;**44** (11):1122–3.
71 Artino AR Jr, Hemmer PA, Durning SJ. Using self-regulated learning theory to understand the beliefs, emotions, and behaviours of struggling medical students. *Acad Med* 2011;**86** (10 Suppl):35–8.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Table S1. Full details of the studies included in the present review.

Appendix S1. Search terms used to identify published materials for the current review.

Received 4 May 2012; editorial comments to authors 22 June 2012; accepted for publication 24 July 2012