Evidence-based medical leadership development: a systematic review

Oscar Lyons, Robynne George, Joao R Galante, Alexander Maft, Thomas Fordw, Jan Frich, Jaason Matthew Geerts

ABSTRACT
Health systems invest significant resources in leadership development for physicians and other health professionals. Competent leadership is considered vital for maintaining and improving quality and patient safety. We carried out this systematic review to synthesise new empirical evidence regarding medical leadership development programme factors which are associated with outcomes at the clinical and organisational levels. Using Ovid MEDLINE, we conducted a database search using both free text and Medical Subject Headings. We then conducted an extensive hand-search of references and of citations in known healthcare leadership development reviews. We applied the Medical Education Research Study Quality Indicator (MERSQI) and the Joanna Briggs Institute Critical Appraisal Tool to determine study reliability, and synthesised results using a meta-aggregation approach. 117 studies were included in this systematic review. 28 studies met criteria for higher reliability studies. The median critical appraisal score according to the MERSQI was 8.5/18 and the median critical appraisal score according to the JBI was 3/10. There were recurring causes of low study quality scores related to study design, data analysis and reporting. There was considerable heterogeneity in intervention design and evaluation design. Programmes with internal or mixed faculty were significantly more likely to report organisational outcomes than programmes with external faculty only (p < 0.049). Project work and mentoring increased the likelihood of organisational outcomes. No leadership development content area was particularly associated with organisational outcomes. In leadership development programmes in healthcare, external faculty should be used to supplement in-house faculty and not be a replacement for in-house expertise. To facilitate organisational outcomes, interventions should include project work and mentoring. Educational methods appear to be more important for organisational outcomes than specific curriculum content. Improving evaluation design will allow educators and evaluators to more effectively understand factors which are reliably associated with organisational outcomes of leadership development.

INTRODUCTION
Health systems invest significant resources in leadership development for physicians and other health professionals. Competent leadership is considered vital for team effectiveness, for clinical and financial performance and for maintaining and improving quality and patient safety. Clinical leadership development involves activities to promote leadership competencies among clinicians, while medical leadership development refers to activities centred on doctors.

Research suggests that medical leadership development can improve outcomes at individual, organisational and clinical levels. Evidence backing medical leadership development activities has, however, been variable in quality. There has been a particular lack of research and evaluation that goes beyond individual learner feedback and subjective outcomes. One systematic review of 45 studies evaluating leadership development interventions for doctors found that effective interventions were characterised by the use of multiple learning methods, including seminars and group work, alongside action learning projects in multidisciplinary teams. These findings were echoed in a recent study by Geerts et al., who emphasised that plans need to be in place for transferring learning from the intervention into the working environment.

We undertook this systematic review to synthesise recent empirical evidence regarding medical leadership development programme factors associated with outcomes at the clinical and organisational levels. We specifically investigated links between aspects of programme design, delivery and evaluation and improved outcomes. Given the variable quality of studies highlighted in previous reviews, we applied two validated critical appraisal instruments to isolate higher reliability findings. This review is the first to apply both instruments in order to identify and synthesise the highest quality empirical evidence in medical leadership development.

METHODS
The design of this review was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and the Best Evidence in Medical Education (BEME) guide for systematic reviews. Our methods were based on the review conducted by Frich et al, with methodological changes drawn from other reviews. Following the BEME recommendations for systematic reviews, we hand-searched references and citations of known reviews extensively to supplement our database search. In line with recommendations from Geerts et al. and Rosenman et al., we assessed study quality using the Medical Education Research Study Quality Indicator (MERSQI), which is designed to measure the methodological quality of quantitative medical education research studies. We added...
the Joanna Briggs Institute (JBI) Critical Appraisal Checklist, which is designed for meta-aggregation of qualitative research and is well-established in healthcare research.

**Search strategy**

We began this review by re-examining the data set identified in the review of leadership development for physicians by Frich et al. With assistance from a specialist librarian at the University of Oxford, we then based our search strategy on Frich et al's review. Using the Ovid MEDLINE database, we conducted a search using both free text and Medical Subject Headings. The full search terms are listed in the online supplemental material. This search identified 501 unique publications. We then conducted an extensive hand-search of references and of citations in known healthcare leadership development reviews using Web of Science and Google Scholar. This identified an additional 107 studies for possible inclusion, for a total of 608 records for screening (figure 1).

**Inclusion criteria**

We included any peer-reviewed study published in English between January 2000 and January 2020 which:

1. Describes a leadership development intervention (programme, workshop, course and so on).
2. Includes physicians as learners (defined here as any practising doctor post-qualification).
3. Evaluates the leadership development intervention.

Qualitative, quantitative and mixed evaluations were included. We excluded studies where leadership development was a minor focus or where the proportion of physicians was lower than 10% of intervention participants.

**Screening process**

Two members of the review team (OL and TF) independently screened all study titles and abstracts for eligibility. Articles that were approved by either reviewer progressed to full-text review. Two members of the review team independently reviewed for inclusion the full text of all 207 articles that passed the title and abstract screen (TF and RG reviewed half each, OL reviewed all). Where there was disagreement about inclusion, all three reviewers (OL, TF, RG) reached consensus by discussion, with the third reviewer (TF or RG) arbitrating where required.

**Data abstraction**

After screening and reviewing for eligibility, 117 unique studies were included for abstraction and analysis. Data were abstracted and coded for educational setting, methods, content, evaluation methods and outcomes. Outcome data were categorised according to an adapted version of Kirkpatrick’s Framework for evaluation of training programmes (see table 1). One reviewer abstracted and coded all 117 included studies (OL). The second reviewers (RG/JRG/AM/TF) each abstracted and coded at least five studies in full to ensure consistency between reviewers. Data abstraction and coding for all 117 studies was then cross-checked by the second reviewers. Any differences were resolved by consensus, with a third reviewer arbitrating where required. Where possible, statistical tests performed in studies were replicated and checked for accuracy.

**Study quality appraisal**

Previous reviews have shown marked variation in the quality of studies of medical leadership development. To isolate the most reliable evidence linking medical leadership programmes to improved outcomes, two researchers independently critically appraised each included study using the MERSQI and JBI Instruments. Differences in MERSQI and JBI quality score were resolved by consensus, and a third researcher arbitrated where needed.

The MERSQI was applied to all 117 studies. The MERSQI is a validated appraisal tool consisting of 10 items in six domains which relate to design, sampling, type of data collected, validity of evaluation methods, analysis and outcomes. Each domain is
Table 1  Kirkpatrick’s Framework for evaluation of training programmes, with adaptations from Frich et al.8

<table>
<thead>
<tr>
<th>Kirkpatrick level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>Reaction</td>
</tr>
<tr>
<td></td>
<td>Participants’ satisfaction with the learning experience, its organisation, presentation, content, teaching methods and quality of instruction</td>
</tr>
<tr>
<td>Level 2A</td>
<td>Change in attitudes</td>
</tr>
<tr>
<td></td>
<td>Changes in the attitudes or perceptions among participant groups towards leadership, management and/or administration</td>
</tr>
<tr>
<td>Level 2B</td>
<td>Change in knowledge or skills</td>
</tr>
<tr>
<td></td>
<td>For knowledge, this relates to the acquisition of concepts, procedures and principles; for skills, this relates to the acquisition of thinking/problem-solving, psychomotor and social skills</td>
</tr>
<tr>
<td>Level 3A</td>
<td>Behavioural change (self-reported)</td>
</tr>
<tr>
<td></td>
<td>Transfer of learning to the workplace and changes to professional practice, as noted by participants themselves</td>
</tr>
<tr>
<td>Level 3B</td>
<td>Behavioural change (observed)</td>
</tr>
<tr>
<td></td>
<td>Transfer of learning to the workplace and changes to professional practice, as noted by a third party or by promotions</td>
</tr>
<tr>
<td>Level 4a</td>
<td>Results (self-reported)</td>
</tr>
<tr>
<td></td>
<td>Organisational outcomes perceived by respondents and group effectiveness perceived by subordinates</td>
</tr>
<tr>
<td>Level 4b</td>
<td>Results (observed)</td>
</tr>
<tr>
<td></td>
<td>Tangible organisational outcomes, such as reduced costs, improved quality and safety, impact of projects</td>
</tr>
</tbody>
</table>

scored to a maximum of 3, for a total score of 5–18. In line with Geerts et al.,9 studies with scores of 12 or higher were categorised as higher reliability studies (see the Data analysis section). The JBI Checklist for Qualitative Studies was also applied where a study used mixed methods (k=53) or qualitative methods (k=10). Fundamental differences in study design, sampling, evaluation instruments and analysis preclude summative comparison of mixed-methods or qualitative studies to quantitative studies using the MERSQI.16 21 23 24 The JBI Checklist is considered the most appropriate qualitative critical appraisal tool for use in pragmatic meta-aggregation of qualitative research.25 It includes 10 items which regard the study’s research questions, methods, analysis and reporting, for a total score of 0–10. Following recommendations from the JBI Reviewers’ Manual,17 a cut-off score for higher reliability studies was predetermined at 6/10. This score was chosen as studies obtaining six or more points included most key elements of high-quality design.

Data analysis
MERSQI and JBI Scores were used to establish which studies presented more reliable evidence of outcomes. Summary statistics were calculated for all 117 studies. In line with Geerts et al.,9 studies with a final MERSQI Score of 12/18 or higher or studies also analysed separately to isolate the most reliable evidence, as were qualitative and mixed-methods studies which achieved the pre-determined JBI Score of 6/10 or higher. As there was substantial methodological heterogeneity, study characteristics and outcomes were synthesised using a meta-aggregation approach.23 All study quality appraisal scores are presented in the Online supplemental table 1, and full data extraction tables are available on request.

RESULTS
Study reliability (MERSQI and JBI)
Twenty-eight of 117 studies (25%) were categorised as higher reliability. Two studies were categorised as higher reliability by both the MERSQI and the JBI tool.26 27 14 studies (12%) by the MERSQI only and 12 studies (10%) by the JBI tool only. The median critical appraisal score according to the MERSQI was 8.5 (range 5–16 from possible range of 5–18) and the median critical appraisal score according to the JBI was 3 (range 0–9 from possible range of 0–10). Online supplemental table 1 includes the MERSQI and JBI Scores for all included studies.

Study design showed considerable room for improvement, as shown in online supplemental tables 2 and 3. Nearly half the of studies (46%) relied on post-programme evaluations only, and 92% did not include a control group. Of the nine studies that did include control groups, most had substantial methodological flaws in their selection of control groups. One common method for control group recruitment was to use unsuccessful course applicants.28–30 In terms of evaluation design, the median evaluation instrument score was 0 (range 0–3). The majority of studies (59%) did not fulfil any of the MERSQI requirements for evaluation instruments, including reporting questionnaire design, wording and content. Objective outcome measures were used in only a minority of studies, with 60% relying solely on self-reported measures.

Data analysis and reporting likewise showed considerable limitations. Only one in five studies (20%) met criteria for comprehensive analysis and reporting of data. Few studies analysed their data beyond descriptive statistics to consider the generalisability and implications (13%). In many cases, studies omitted basic statistical significance tests.

Many studies did not contain key reporting elements for qualitative research as outlined in the JBI tool (see online supplemental table 3). There was clear congruence between research methodologies chosen and the research objectives and methods employed in 60% of studies. A minority of studies adequately reported their analysis (28%) and interpretation of data (25%), the potential for the researcher to have influenced data collection and interpretation (23%) and the researcher’s cultural or theoretical orientation (15%). Participant voices were clearly represented through quotes in only 16/53 (30%) of mixed-methods studies and 5/10 (50%) of qualitative studies. There was a statement of ethical approval or ethics exemption in only 26 of 63 studies (40%) which used qualitative methods. No study included a statement of philosophical perspective (normally expected for qualitative research).17

Programme design
There was considerable heterogeneity in leadership development intervention design. It was often unclear whether established good practice for development of medical education interventions was followed, as shown in figure 2.31 32 Only 52 studies (44%) reporting having conducted a needs assessment before...
their intervention, and only 20 studies (17%) explicitly reported using an established capability or competency framework to inform leadership programme goals and objectives. There was, however, a plan for training transfer reported or built into 68 of 117 interventions (59%).

The majority of interventions were carried out in a single hospital department (27%), single hospital (22%) or a single university (12%). Just under a quarter (23%) of interventions were conducted in multiple healthcare centres. A further 15% of studies were conducted within a specialty training programme outside healthcare centres.

Most of the studies took place in the USA (67%) or the UK (16%). The remainder of studies were in other European countries (7%), Canada (4%) or Australia (3%), with a single study each from Africa, India, Israel and Qatar.

Programmes ranged in length from 2 hours to 4 years. The median intervention length was 6 months, and the most common length was 1 year (19%). Only 18 interventions (15%) lasted longer than 1 year. Five interventions (4%) were shorter than 1 day.

Programme faculty
Programmes were predominately delivered by either in-house faculty (36%) or a mix of in-house and external faculty (32%). Programmes delivered by mixed faculty were most likely to show organisational outcomes, as shown in figure 3. The professional backgrounds, qualifications and experience of faculty were generally not reported.

Participants
The majority of programmes included doctors only (76%). Physician learners ranged from residents (60%) to full specialists (30%) and academic medical faculty (19%). Only nine studies of 117 involved doctors from more than one category. Behavioural outcomes were reported in a similar percentage of higher reliability studies for each category (85%–92%), while organisational outcomes were more commonly reported in programmes with academic medical faculty (50%) or full specialists (44%) than in programmes with only residents (20%). The 26 studies (24%) reporting multidisciplinary programmes included a combination of nurses (12%), managers (15%) and allied health professionals (9%). Most studies did not report the gender of participants (74%) or the age of participants (87%).

In terms of participant selection criteria, the majority of interventions included participants who volunteered (27%), were nominated (19%) or who applied to the programme (16%). In some cases the application process was highly competitive. Interventions were mandatory in one-fifth of studies (20%). A considerable proportion of all studies (23%) did not report the selection process for their learners, including one quarter (25%) of the studies categorised as higher reliability by MERSQI criteria.

Educational methods
A wide range of educational methods were employed in various combinations across the reviewed studies, as shown in figure 4. Most interventions included lectures (68%) and small group work (61%). Project work was included in the majority of programmes which reported organisational outcomes (68%), but only in a minority of studies which did not report organisational outcomes (33%). Individual or team mentoring was also more prevalent where organisational outcomes were reported (47% vs 23%).

Educational content
Educational content varied considerably among interventions. The most consistent content area was leadership theory (reported in 65% of interventions). The other common content areas were performance management (44%), self-management (41%), change management (39%), communication (36%), teamwork (33%), quality improvement (30%), healthcare policy (27%), healthcare finance (26%) and leadership behaviours (20%). There were no notable educational content differences in higher reliability studies or in studies which reported organisational outcomes (Kirkpatrick level 4).

Evaluation methods
A wide range of evaluation methods were employed across the included studies. Nearly half used quantitative methods only for their evaluation (46%). Of the remainder, most studies used mixed methods (45%), with 10 studies (9%) using purely...
outcomes in higher reliability studies were more likely have included project work (70% vs 44%), mentoring (50% vs 22%), coaching (22% vs 11%) and reflective instruments such as personality type assessments (22% vs 6%) than higher reliability studies that did not report organisational outcomes. Organisational outcomes were reported less frequently in higher reliability studies that included simulation or role play (10% vs 33%).

**DISCUSSION**

The aim of this review was to synthesise recent empirical evidence and explore factors associated with higher level outcomes in physician leadership development.

We found a substantial increase in the number of studies which evaluate medical leadership development interventions compared with previous reviews.6-10 14 15 In many studies, it is still not clear whether best practices for design, delivery and evaluation are being followed.31 It is also not clear whether there are sufficient behavioural and organisational outcomes to justify the considerable and increasing investments in medical leadership development.

Compared with previous reviews, we found an increase in the proportion of studies which report the use of active learning methods such as project work, simulation, discussions and reflections, which are widely accepted to be a vital component of leadership development55 and which were associated in our review with increased Kirkpatrick level 4 outcomes.

No single leadership development content area was particularly associated with improved outcomes. With respect to educational methods, however, there was an association between the inclusion of individual or group project work and of mentoring with organisational outcomes. This may support the established position that educational methods are more important than specific curriculum content for leadership development.58 Simulation and role play were less common in higher reliability studies which reported organisational outcomes that those that did not report organisational outcomes. This unexpected finding could result from these studies being situated in a training environment rather than a working environment. Alternatively, it could result from the evaluation process and study designs rather than from a lack of organisational impact. Studies which included simulation and role play tended to focus their evaluations on objective changes in behaviour at the expense of evaluating organisational outcomes (see online supplemental table 1). Interestingly, lacking a leadership development framework did not seem to impede programmes from reporting organisational

Behavioural and organisational outcomes in higher reliability studies

A full summary of outcomes from all 117 studies is provided in online supplemental table 1.

There was a range of behavioural (Kirkpatrick level 3) and organisational (Kirkpatrick level 4) outcomes demonstrated in higher reliability studies.

Behavioural changes were objectively demonstrated in higher reliability studies through observed changes in behaviour,26 27 39-43 promotions,44 45 increased responsibilities or titles28 46-49 and project completion.30-52 Subjective changes in behaviour included improved communication,39 influence,50 delegation,27 collaboration,28 involvement in service improvement47 and application of skills learnt or improved leadership in general.39 40 54-57 These changes were indicated through interviews, free text questionnaire responses and behavioural self-assessments.

Organisational outcomes in higher reliability studies (Kirkpatrick level 4) were defined prospectively and in most cases were objectively demonstrated through leadership project impact evaluations. Projects achieved a range of outcomes, including reduced waiting times,50 improved patient care,46 50 and cost savings.27 46 47 50 By assessing the financial impact of projects completed during the intervention and relating this to programme costs, one higher reliability study reported a 364% financial return-on-investment (ROI).27 Other objective outcomes included reduced organisational turnover of participants.28 improved departmental working climate,39 reduced sick leave44 and increased promotion of women.45 Organisational outcomes were subjectively indicated through reports of increased staff retention36 and improvement in organisational effectiveness.25 One study reported that ‘intangible benefits’ resulted in a 106% financial ROI.31

Organisational outcomes in higher reliability studies were reported more frequently from programmes delivered by a mix of internal and external faculty than from programmes delivered by only external faculty (83% vs 11%), as shown in figure 2. Organisational outcomes were also more frequently reported from interventions conducted in a whole hospital (57%) or multiple hospitals (40%), compared with interventions conducted in a single specialty (conference or outside-hospital training programme) (33%), single university (25%) or in a single department (0%). There were no notable differences in outcomes related to specific educational content.

Higher reliability studies that reported organisational outcomes were more likely to include project work (70% vs 44%), mentoring (50% vs 22%), coaching (22% vs 11%) and reflective instruments such as personality type assessments (22% vs 6%) than higher reliability studies that did not report organisational outcomes. Organisational outcomes were reported less frequently in higher reliability studies that included simulation or role play (10% vs 33%).
outcomes. This may indicate that programmes which are designed as bespoke solutions to local needs are more likely to achieve organisational impact than pre-packaged approaches to leadership development.

There was an additional association of more senior participant level with organisational outcomes. This may be related to the wider scope of influence or practice of senior physicians compared with resident physicians. It could also indicate that there is a longer post-programme development period before residents are able to have an impact on organisational outcomes. This would align with the finding that programmes which evaluated longer-term outcomes were more likely to report organisational outcomes.

Importantly, our findings indicated that leadership development interventions which used a combination of internal and external faculty were most likely to report organisational outcomes, and those interventions which used external faculty only were least likely. This could have significant implications for procurement and design of leadership development interventions across healthcare, particularly as courses run internally are associated with significantly reduced costs.63 64

As in previous physician leadership development reviews that used critical appraisal instruments,6 7 8 14 we found that studies frequently did not meet criteria for high reliability. Many studies failed to report important methodological features, which restricts readers’ ability to appraise studies and learn from their findings. This was particularly notable in terms of questionnaire design, with fewer than one in 10 studies using validated questionnaires or reporting their questionnaire content in detail. Most studies also did not report or analyse outcome evaluation data comprehensively. Many study designs were biased towards obtaining positive results, particularly in terms of the absence of control groups, having stringent or undisclosed selection criteria, including leading questions on questionnaires and relying solely on self-ratings. This is likely to have resulted in improved reported outcomes. The lack of evaluation quality seems to indicate perfunctory attention paid to evaluation design and precludes confident conclusions from these studies. Future studies could benefit from consulting study quality appraisal checklists such as the MERSQI and JBI in advance, in order to effectively design their evaluations.

This review does indicate that certain recommendations for improved programme evaluation are beginning to be applied into research. Whereas only 29% of the studies reviewed by Frich et al8 included qualitative components, 63 (54%) of the 117 studies included in our review used mixed or qualitative methods. In a nascent and complex field such as medical leadership development research,1 8 9 61 qualitative methods can have value in terms of establishing effective programme design features to achieve desired outcomes,23 25 31 32 as well as helpful nuances of how, for whom, to what extent or in what circumstances interventions are effective or not.9 10 62

Additionally, many studies in this systematic review evaluated outcomes at Kirkpatrick level 3 behavioural change (57%) or level 4 organisational outcomes (24%). This is a significant improvement from previous reviews.23 15 Changes in behaviour (level 3) and organisational outcomes (level 4) are more closely associated with transfer of learning to the working environment than participant reaction (level 1) and learning (level 2).63 65

CONCLUSION
Our review has practical implications for those commissioning, designing and evaluating medical leadership development programmes in healthcare. No specific area of curriculum content and no particular leadership development framework were clearly associated with behavioural or organisational outcomes. While relevance and appropriateness of educational content is important,31 this systematic review has more clear implications for leadership development methods than for specific content. Where possible, interventions should include projects and individual or group mentoring. Transfer of learning from the programme into learners’ daily work and their organisations should be planned into the programme and where possible active learning educational designs should be employed, including opportunities for learners to set their own goals for development. External faculty should be judiciously used to supplement in-house faculty, not as a replacement for in-house expertise.

In terms of evaluation design, efforts should be made to ensure that evaluations are cost-effective and produce data that is useful for both practitioners and researchers.66 67 Effective mixed-methods evaluation strategies should be integrated into evaluation designs. Study quality checklists such as the MERSQI and JBI could be consulted in the programme design phase to help build high quality quantitative and qualitative evaluation methods into programmes. At the minimum, evaluation design should include consideration of assessment at multiple time points, inclusion of control groups and collection of objective data, as well as collection of qualitative data from interviews, focus groups, questionnaires or observations. Programme goals and intended organisational outcomes should be explicitly considered during evaluation design, so that measures of organisational outcomes (including project outcomes) can be incorporated into the evaluation design. Improving study design and building robust evaluation methods into programmes will allow evaluators and educators to more effectively understand factors which are reliably associated with high level programme outcomes. This could both inform the improvement of individual programmes and appraisal tools with cut-off scores for higher reliability studies. To the best of our knowledge, this is the first systematic review of healthcare leadership development interventions to use the JBI critical appraisal tool to critically appraise qualitative studies. The JBI tool enabled us to identify 12 additional higher reliability qualitative and mixed-methods studies which were not identified using the MERSQI. Marked heterogeneity of studies and evaluations precluded a formal meta-analysis, therefore, we adopted a meta-aggregation approach. This enabled us to highlight design components that are correlated with behavioural and organisational outcomes in higher reliability studies.

A substantial majority of studies reported only positive outcomes, which could represent a publication bias, and we limited our review to English language peer-reviewed studies. In line with Frich et al,6 8 our database search was limited to MEDLINE, however, we augmented our database search with an extensive hand-search of reference lists and citations using Web of Science and Google Scholar. The hand-search revealed that many relevant empirical studies were absent from recent reviews despite some of those reviews searching a greater range of research databases. This could indicate flaws in healthcare leadership development literature tagging and filing procedures within medical and educational databases.

Limitations and strengths
This review was limited by the reliability of the studies included. We attempted to control for study reliability using critical appraisal tools with cut-off scores for higher reliability studies. To the best of our knowledge, this is the first systematic review of healthcare leadership development interventions to use the JBI critical appraisal tool to critically appraise qualitative studies. The JBI tool enabled us to identify 12 additional higher reliability qualitative and mixed-methods studies which were not identified using the MERSQI. Marked heterogeneity of studies and evaluations precluded a formal meta-analysis, therefore, we adopted a meta-aggregation approach. This enabled us to highlight design components that are correlated with behavioural and organisational outcomes in higher reliability studies. A substantial majority of studies reported only positive outcomes, which could represent a publication bias, and we limited our review to English language peer-reviewed studies. In line with Frich et al, we database search was limited to MEDLINE, however, we augmented our database search with an extensive hand-search of reference lists and citations using Web of Science and Google Scholar. The hand-search revealed that many relevant empirical studies were absent from recent reviews despite some of those reviews searching a greater range of research databases. This could indicate flaws in healthcare leadership development literature tagging and filing procedures within medical and educational databases.

CONCLUSION
Our review has practical implications for those commissioning, designing and evaluating medical leadership development programmes in healthcare. No specific area of curriculum content and no particular leadership development framework were clearly associated with behavioural or organisational outcomes. While relevance and appropriateness of educational content is important, this systematic review has more clear implications for leadership development methods than for specific content. Where possible, interventions should include projects and individual or group mentoring. Transfer of learning from the programme into learners’ daily work and their organisations should be planned into the programme and where possible active learning educational designs should be employed, including opportunities for learners to set their own goals for development. External faculty should be judiciously used to supplement in-house faculty, not as a replacement for in-house expertise.

In terms of evaluation design, efforts should be made to ensure that evaluations are cost-effective and produce data that is useful for both practitioners and researchers. Effective mixed-methods evaluation strategies should be integrated into evaluation designs. Study quality checklists such as the MERSQI and JBI could be consulted in the programme design phase to help build high quality quantitative and qualitative evaluation methods into programmes. At the minimum, evaluation design should include consideration of assessment at multiple time points, inclusion of control groups and collection of objective data, as well as collection of qualitative data from interviews, focus groups, questionnaires or observations. Programme goals and intended organisational outcomes should be explicitly considered during evaluation design, so that measures of organisational outcomes (including project outcomes) can be incorporated into the evaluation design. Improving study design and building robust evaluation methods into programmes will allow evaluators and educators to more effectively understand factors which are reliably associated with high level programme outcomes. This could both inform the improvement of individual programmes and appraisal tools with cut-off scores for higher reliability studies. To the best of our knowledge, this is the first systematic review of healthcare leadership development interventions to use the JBI critical appraisal tool to critically appraise qualitative studies. The JBI tool enabled us to identify 12 additional higher reliability qualitative and mixed-methods studies which were not identified using the MERSQI. Marked heterogeneity of studies and evaluations precluded a formal meta-analysis, therefore, we adopted a meta-aggregation approach. This enabled us to highlight design components that are correlated with behavioural and organisational outcomes in higher reliability studies.

A substantial majority of studies reported only positive outcomes, which could represent a publication bias, and we limited our review to English language peer-reviewed studies. In line with Frich et al, we database search was limited to MEDLINE, however, we augmented our database search with an extensive hand-search of reference lists and citations using Web of Science and Google Scholar. The hand-search revealed that many relevant empirical studies were absent from recent reviews despite some of those reviews searching a greater range of research databases. This could indicate flaws in healthcare leadership development literature tagging and filing procedures within medical and educational databases.

Limitations and strengths
This review was limited by the reliability of the studies included. We attempted to control for study reliability using critical appraisal tools with cut-off scores for higher reliability studies. To the best of our knowledge, this is the first systematic review of healthcare leadership development interventions to use the JBI critical appraisal tool to critically appraise qualitative studies. The JBI tool enabled us to identify 12 additional higher reliability qualitative and mixed-methods studies which were not identified using the MERSQI. Marked heterogeneity of studies and evaluations precluded a formal meta-analysis, therefore, we adopted a meta-aggregation approach. This enabled us to highlight design components that are correlated with behavioural and organisational outcomes in higher reliability studies. A substantial majority of studies reported only positive outcomes, which could represent a publication bias, and we limited our review to English language peer-reviewed studies. In line with Frich et al, we database search was limited to MEDLINE, however, we augmented our database search with an extensive hand-search of reference lists and citations using Web of Science and Google Scholar. The hand-search revealed that many relevant empirical studies were absent from recent reviews despite some of those reviews searching a greater range of research databases. This could indicate flaws in healthcare leadership development literature tagging and filing procedures within medical and educational databases.

CONCLUSION
Our review has practical implications for those commissioning, designing and evaluating medical leadership development programmes in healthcare. No specific area of curriculum content and no particular leadership development framework were clearly associated with behavioural or organisational outcomes. While relevance and appropriateness of educational content is important, this systematic review has more clear implications for leadership development methods than for specific content. Where possible, interventions should include projects and individual or group mentoring. Transfer of learning from the programme into learners’ daily work and their organisations should be planned into the programme and where possible active learning educational designs should be employed, including opportunities for learners to set their own goals for development. External faculty should be judiciously used to supplement in-house faculty, not as a replacement for in-house expertise.

In terms of evaluation design, efforts should be made to ensure that evaluations are cost-effective and produce data that is useful for both practitioners and researchers. Effective mixed-methods evaluation strategies should be integrated into evaluation designs. Study quality checklists such as the MERSQI and JBI could be consulted in the programme design phase to help build high quality quantitative and qualitative evaluation methods into programmes. At the minimum, evaluation design should include consideration of assessment at multiple time points, inclusion of control groups and collection of objective data, as well as collection of qualitative data from interviews, focus groups, questionnaires or observations. Programme goals and intended organisational outcomes should be explicitly considered during evaluation design, so that measures of organisational outcomes (including project outcomes) can be incorporated into the evaluation design. Improving study design and building robust evaluation methods into programmes will allow evaluators and educators to more effectively understand factors which are reliably associated with high level programme outcomes. This could both inform the improvement of individual programmes and
contribute to the medical leadership literature as a whole. It is only through more considered and thorough evaluation of physician leadership development programs that we will be able to invest the movement they represent.

Twitter Oscar Lyons @oscarlyonsnz, Jan Frich @J_Frich and Jaason Matthew Geerts @jaasongeerts

Acknowledgements We would like to thank Tatjana Petrinic, University of Oxford Health Care/Healthcare Librarian, for her invaluable assistance and advice in the search process.

Contributors OL, RG and JRG planned the review. OL, RG and TF screened studies for inclusion. OL, RG, JRG, AM and TF abstracted and coded studies. OL, RG, JRG, AM, TF, JF and JMG contributed to analysis, writing and editing the manuscript.

Funding Oscar Lyons was supported during this work by a Rhodes Scholarship, a Goodger and Schorstein Research Scholarship (University of Oxford) and the Shirccliffe Fellowship (Universities New Zealand).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/ or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID ids Oscar Lyons http://orcid.org/0000-0001-5809-7173
Jan Frich http://orcid.org/0000-0001-9079-7508
Jaason Matthew Geerts http://orcid.org/0000-0001-6672-3859

REFERENCES


26 Orme D, Campbell C. How leadership training saves money ‘service line leadership’ at Nottingham University Hospitals. Leader 2019;3:29–36.


Review


<table>
<thead>
<tr>
<th>Source</th>
<th>First Author, Year</th>
<th>Setting</th>
<th>Learner Number</th>
<th>Learner Type</th>
<th>Intervention Length</th>
<th>Intervention Description</th>
<th>Teaching Methods</th>
<th>Educational Content</th>
<th>Main Findings by Kirkpatrick Level</th>
<th>JM Score</th>
<th>NMRDQ Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wolfe, 2020</td>
<td>Two US ICUs</td>
<td>10</td>
<td>3 Physicians</td>
<td>7 Nurses</td>
<td>8 months</td>
<td>6 externally provided modules, total 23.5 hrs</td>
<td>Learning activities, small group skill practice, problem-solving sessions, feedback and reinforcement of skills assignment, assessments, feedback</td>
<td>Leadership, communication, coordination, problem solving/conflict management, team culture</td>
<td>3a. Leaders reported increased satisfaction with their own communication and leadership skills 3b. Communication skills of ICU nurses and physician leaders improved significantly in simulation (hours 37 to 79.68). Relationship skills remained high (77.9/100). 4b. Reported increased problem solving between groups and decreased personal stress in use of the ICU.</td>
<td>n/a</td>
<td>16</td>
</tr>
<tr>
<td>Persons, 2019</td>
<td>Single US hospital residency</td>
<td>14</td>
<td>Residents</td>
<td>Emergency Medicine</td>
<td>4 days</td>
<td>4 days of simulation scenarios</td>
<td>Introductory didactic presentation followed by a series of skill stations and structured debriefs</td>
<td>Crew Resource Management (CRM)</td>
<td>2b. Both teams showed overall gradual improvement in CRM skills compared to the preceding teams, suggesting that observational learning of CRM was effective in this setting. Project work, interactive seminars, one-on-one project mentoring. 3b. Eight individuals accomplished 100% of their projects, 20/27 completed at least half of the project.</td>
<td>n/a</td>
<td>14</td>
</tr>
<tr>
<td>Cooper, 2001</td>
<td>Single advanced life support course</td>
<td>35</td>
<td>Mixed variety, doctors, and nurses</td>
<td>2-day resuscitation course</td>
<td>75 mini leadership development seminar courses</td>
<td>Lectures, videos and discussion groups, home reading</td>
<td>Importance of leadership, behaviours of effective leadership, introduction/intervention</td>
<td>2a. Aspiration to higher leadership position inside an academic health centre decreased; 3b. Participants more likely to hold regional or national leadership titles and to have taken on new leadership titles. No significant difference in promotions. 4b. Sick leave increased by 6.9 days per year. Fewer in intervention group compared to reference group (1.3 days vs 8.2 days).</td>
<td>3a. Self-reported significant improvement in effectiveness as a leaders and power and influence. 3b. 100% of participants completed their projects. 4b. Skills of projects achieved level 3 (moderate improvement in process measure) with 22% of those attaining level 4 (significant improvement in outcomes measure).</td>
<td>n/a</td>
<td>15</td>
</tr>
<tr>
<td>Malling, 2009</td>
<td>Single educational region in Denmark</td>
<td>28</td>
<td>Consultants (responsible for education)</td>
<td>6 months</td>
<td>Two three-day residential modules and follow-up day</td>
<td>Residential modules and follow-up day.</td>
<td>Pedagogical knowledge, organization of specialist training, educational evaluation and quality assurance, planning specialist training in the department, supervision of supervisors, implementation strategies, personal development, leadership in specialist training, research in medical education</td>
<td>2b. Significant improvements in self-reported knowledge and skills 2b. Technical, administration and human skills feedback did not improve or differ from the control group. 2a. No significant differences in self-reported well-being, self-esteem, mental energy, influence, authority, participation, feedback, goal clarity or efficiency.</td>
<td>n/a</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>van Vliet, 2004</td>
<td>University hospitals in Sweden</td>
<td>52</td>
<td>Specialists, senior physicians, heads of departments</td>
<td>1 year</td>
<td>Three programs, including mentor programs, management networks, and letters held across 1 year; no details on number, duration, or allocation to programs</td>
<td>Mentoring/Networking/Lectures</td>
<td>NI</td>
<td>NI</td>
<td>2a. No significant differences in self-reported skills development, self-efficacy, mental or physical ability or work-related exhaustion</td>
<td>n/a</td>
<td>14</td>
</tr>
<tr>
<td>Fawcett, 2008</td>
<td>Single US hospital</td>
<td>131</td>
<td>Assistant/associate/Full professors</td>
<td>9 months</td>
<td>4 x 1.5 day sessions over 8 months</td>
<td>Interactive teaching methods based on adult learning principles, an active learning projects</td>
<td>Personal development as a leader managing people and relationships, technical and administrative skills, management finance and accounting understanding the organizational system</td>
<td>3b. Technical, administration and human skills feedback did not improve or differ from the control group. 3b. Large but not significant increases in all objective measures of leadership, problem solving, decision making, teamwork, coordination, problem solving/conflict management, team work.</td>
<td>3a. Self-reported significant improvement in effectiveness as a leaders and power and influence. 3b. No significant difference in senior management positions between program and control. 4b. No significant difference in self-reported well-being, self-esteem, mental energy or work-related exhaustion.</td>
<td>n/a</td>
<td>15</td>
</tr>
<tr>
<td>Levine, 2008</td>
<td>Single US academic medical centre</td>
<td>67</td>
<td>Residents (Chief residents, medicine and surgery)</td>
<td>1 year</td>
<td>Two-day offsite immersion training, project work</td>
<td>Small group discussions, evidence-based mini-lectures, interactive seminars, overview of project meetings</td>
<td>Management of complex older patients, geriatric medicine, Crew Resource Management (CRM),crew resource management</td>
<td>3b. Self-reported significant improvement in effectiveness as a leaders and power and influence. 3b. Participants more likely to hold regional or national leadership titles and to have taken on new leadership titles. No significant difference in promotions. 3a. Self-reported significant improvement in effectiveness as a leaders and power and influence. 3b. Eight individuals accomplished 100% of their projects, 26/27 completed at least half of the project.</td>
<td>3a. Aspiration to higher leadership position inside an academic health centre decreased; 3b. Significant improvements in self-reported knowledge and skills. 3b. No significant difference in senior management positions between program and control. 4b. Skills of projects achieved level 3 (moderate improvement in process measure) with 22% of those attaining level 4 (significant improvement in outcomes measure).</td>
<td>n/a</td>
<td>15</td>
</tr>
<tr>
<td>Hopkins, 2008</td>
<td>Single US hospital network</td>
<td>113</td>
<td>Senior medical leaders and academic faculty, 18 laboratory, 94 doctors</td>
<td>9 months</td>
<td>Six 1.5 day sessions spaced over 9 months</td>
<td>Baseline assessments of their leadership competencies, Multi-source feedback, Myers Briggs Type Indicator and the Thomas-Kilmann Conflict Mode Instrument, feedback materials, assignments, case study, role plays, discussion in small groups, brief reflection and writing assignments, response to video vignettes, brainstorming, and small group problem-solving assignments, with minimal emphasis on didactic lectures.</td>
<td>Management as a leader. Managing and leading teams and organizations, including mentorship.</td>
<td>3b. Self-reported significant improvement in effectiveness as a leaders and power and influence. 3b. Participants rated the course beneficial and meeting their expectations (22-3.2/4). 3a. Self-reported significant improvement in effectiveness as a leaders and power and influence. 3b. Eight individuals accomplished 100% of their projects, 26/27 completed at least half of the project.</td>
<td>3b. Significant improvements in self-reported knowledge and skills. 3b. Significant improvements in self-reported leadership skills. 3b. No significant difference in senior management positions between program and control. 4b. Skills of projects achieved level 3 (moderate improvement in process measure) with 22% of those attaining level 4 (significant improvement in outcomes measure).</td>
<td>n/a</td>
<td>15</td>
</tr>
<tr>
<td>Daniels, 2008</td>
<td>Single US university executive education programme</td>
<td>78</td>
<td>Female academic leadership faculty (assistant or full professor level)</td>
<td>1 year</td>
<td>Executive leadership development program for senior female faculty</td>
<td>Not specified</td>
<td>Not specified</td>
<td>1. Educational leadership competence. The executive program was designed around a four-part program model was slightly shifted but significantly greater than the means in both the control groups. (average 62.7) increased.</td>
<td>3b. Self-reported significant improvement in effectiveness as a leaders and power and influence. 3b. No significant difference in senior management positions between program and control. 4b. Skills of projects achieved level 3 (moderate improvement in process measure) with 22% of those attaining level 4 (significant improvement in outcomes measure).</td>
<td>n/a</td>
<td>13</td>
</tr>
<tr>
<td>Orme, 2019</td>
<td>Single UK hospital trust</td>
<td>625</td>
<td>Consultants and senior healthcare professionals and managers</td>
<td>12 months</td>
<td>12 months with 4 workshop days spread over 6 months</td>
<td>Five face-to-face delivery days, ongoing telephone coaching and the use of a benchmarked 280 degree profile, project work, interactive scenarios. 3 days, from 1 day after 3 months, then 1 day after 6 months, Support via coaching between.</td>
<td>3b. Communication skills of ICU nurse and physician leaders improved significantly in simulation (hours 37 to 79.68). Relationship skills remained high (77.9/100). 4b. Reported increased problem solving between groups and decreased personal stress in use of the ICU. 2b. Both teams showed overall gradual improvement in CRM skills compared to the preceding teams, suggesting that observational learning of CRM was effective in this setting. Project work, interactive seminars, one-on-one project mentoring. 3b. Eight individuals accomplished 100% of their projects, 20/27 completed at least half of the project.</td>
<td>1. Evaluation using the net promotor score gave 92% score (promoters-detractors/total) 2b. Significant improvements in self-reported knowledge and skills. 3b. Significant improvements in self-reported leadership skills. 3b. No significant difference in senior management positions between program and control. 4b. Skills of projects achieved level 3 (moderate improvement in process measure) with 22% of those attaining level 4 (significant improvement in outcomes measure).</td>
<td>6</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>Study (First Author, Year)</td>
<td>Setting</td>
<td>Learner Number</td>
<td>Learner Type</td>
<td>Intervention Length</td>
<td>Intervention Description</td>
<td>Teaching Methods</td>
<td>Educational Context</td>
<td>Main Findings by Repartition Level</td>
<td>AB Score</td>
<td>MERSQI Score</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------</td>
<td>----------------</td>
<td>-------------</td>
<td>---------------------</td>
<td>-------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>---------------------------------</td>
<td>---------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Cole, 2017 (11)</td>
<td>Single US hospital department</td>
<td>10</td>
<td>Residents (Anesthesiology FOGY/R)</td>
<td>2 weeks</td>
<td>2 weeks operating room management and leadership elective rotation</td>
<td>Action learning, literature review, reflective learning</td>
<td>Non-technical skills</td>
<td>Leaders (6) demonstrated BMJ quality improvements were achieved: 3a. Increased re-evaluates and debriefs; 3b. &quot;pitches and actively seeks out information&quot;; 3c. &quot;anticipates changing environment&quot;; 3d. Increase across a range of metrics; task management, clinical decision making, situational awareness, as measured by &quot;recalibration&quot; non-technical skills (CTA) questionnaire; 3e. Statistically significant increase in leadership of national committees, production of national workshops and presentation at national platforms.</td>
<td>n/a</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>Haider MM, 2018 (12)</td>
<td>Single US specialty association, 45 sites</td>
<td>49</td>
<td>PwCaudie academic faculty</td>
<td>10 months</td>
<td>3 sessions focusing on the individual, their training and leadership, followed by team interaction with the participants</td>
<td>&quot;highly interactive format&quot;, peer mentorship, Professional development, leadership training; administrative skill development.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ten-Have, 2013 (13)</td>
<td>Four Dutch (4) in Singel hospital</td>
<td>9</td>
<td>Exposure/intensive care fellow</td>
<td>Central/experienced intensive</td>
<td>23 months</td>
<td>2-day simulation, group feedback on observed interdisciplinary rounds, 2-week post-assessment before or after post-training added, half-day workshop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gilboyse, 2017 (14)</td>
<td>Single Canadian residency program (paediatrics)</td>
<td>29</td>
<td>Residents (Paediatrics, FOGY-POA)</td>
<td>1/2 day</td>
<td>Preparatory session followed by two simultaneous scenarios</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LoPrete, 2000 (15)</td>
<td>Four US residency programs</td>
<td>6</td>
<td>Residents (Family/General)</td>
<td>2 years</td>
<td>60 hours of education in 20 modules</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wunder, 2007 (16)</td>
<td>Single US department of surgery</td>
<td>42</td>
<td>Surgical fellows</td>
<td>6 months</td>
<td>Long week-end didactic study, bi-monthly patient safety related meetings, monthly education conferences, 2-days per year for didactic and group presentations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Higher Quality Studies (AB test)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pedra, 2016 (17)</td>
<td>Single US hospital department</td>
<td>21</td>
<td>Academic surgeons from novice to full professor grade</td>
<td>8 months</td>
<td>1 day per month</td>
<td>Didactic and experiential learning; case studies, team improvement projects, multi-source feedback, debriefing with executive coach</td>
<td>Leadership, team building, business acumen, and health care context</td>
<td>Participants reported high levels of satisfaction with the programme, ranked a 8/10 (10 being excellent use of their time). 2a. Participants &quot;not only enabled but also capable of effective change in their local environment&quot;. 2b. Participants reported increased self-awareness and increased team-building skills, and improvement of leadership knowledge. 3a. Participants reported improved ability to foster collaborative relationships, and general improvement of interactions and networks. 3b. Participants reported increased self-awareness and increased team-building skills, and improvement of leadership knowledge. 4a. Several respondents noted they had planned to leave the NHS and decided to stay after the programme. 4b. Several respondents noted they had planned to leave the NHS and decided to stay after the programme. 5. All interviewees agreed that the experience was valuable. 6. All interviewees agreed that the experience was valuable. 7. All interviewees agreed that the experience was valuable. 8. All interviewees agreed that the experience was valuable. 9. 80. 10.</td>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Thiemermann, 2016 (18)</td>
<td>US Regional healthcare system</td>
<td>21</td>
<td>Physicians across a range of specialties</td>
<td>10 months</td>
<td>2-4 meetings/month (2-3 of additional learning opportunities)</td>
<td>Behavioral style assessment, multi-source feedback, coaching, online discussions, online learning modules, learning project in small groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bergman, 2009 (19)</td>
<td>Single Swedish hospital</td>
<td>53</td>
<td>Managers (3 physicians, 10 nurses, and 11 other health personnel)</td>
<td>7 weeks to 17 months</td>
<td>1. One week intensive course 2. Long-term support group (25-30 participants) 3. Long-term support group (25-30 participants)</td>
<td>Group dynamics, communication, leadership theories</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monohan, 2018 (20)</td>
<td>National UK programme</td>
<td>111</td>
<td>Doctors (seniority and specialty-care) nurses, public health professionals, allied health professionals and managers</td>
<td>3-6 months</td>
<td>3-month placement in a resource-poor country</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tech P., 2019 (21)</td>
<td>Single UK academic healthcare centre</td>
<td>136</td>
<td>Faculty members previously with leadership potential</td>
<td>20 weeks</td>
<td>20 weeks LED mentor (75% Engagement Hour)</td>
<td>Experiential learning NOS</td>
<td>Self-awareness, critical thinking, effective communication, inclusion, collaboration, empowered professionalism</td>
<td>1. All interviewees agreed that the experience was enjoyable; 2a. Increased percentage considering themselves to be leaders after the programme (72% before, 85% after); interviews reported increased confidence. 2b. Increased self-awareness and leadership skills reported in questionnaire. Participants reported increased awareness of the leadership styles of others. 3a. 75% reported using their new skills; 68% reported not being able to use their skills. Interviewees reported changes in the way they relate to others. 4a. Several respondents noted they had planned to leave the NHS and decided to stay after the programme. 4b. Several respondents noted they had planned to leave the NHS and decided to stay after the programme. 5. All interviewees agreed that the experience was enjoyable; 2a. Increased percentage considering themselves to be leaders after the programme (72% before, 85% after); interviews reported increased confidence. 2b. Increased self-awareness and leadership skills reported in questionnaire. Participants reported increased awareness of the leadership styles of others. 3a. 75% reported using their new skills; 68% reported not being able to use their skills. Interviewees reported changes in the way they relate to others. 4a. Several respondents noted they had planned to leave the NHS and decided to stay after the programme. 5. All interviewees agreed that the experience was enjoyable; 2a. Increased percentage considering themselves to be leaders after the programme (72% before, 85% after); interviews reported increased confidence. 2b. Increased self-awareness and leadership skills reported in questionnaire. Participants reported increased awareness of the leadership styles of others. 3a. 75% reported using their new skills; 68% reported not being able to use their skills. Interviewees reported changes in the way they relate to others. 4a. Several respondents noted they had planned to leave the NHS and decided to stay after the programme.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source (First Author, Year)</td>
<td>Setting</td>
<td>Learner Number</td>
<td>Learner Type</td>
<td>Intervention Length</td>
<td>Intervention Description</td>
<td>Teaching Methods</td>
<td>Educational Content</td>
<td>Main Finding by Kirkpatrick &amp; Kirkpatrick level</td>
<td>JBI Score</td>
<td>BMJ Leader Score</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>----------------</td>
<td>--------------</td>
<td>---------------------</td>
<td>--------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>----------------------------------------------</td>
<td>-----------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>Beerman, 2012 (22)</td>
<td>Single Australian residency program (surgeon)</td>
<td>32</td>
<td>Residents (Surgical trainees, medical students, and allied health professionals)</td>
<td>2 days</td>
<td>Two-day case simulation course</td>
<td>Simulation, peer observations, multi-source feedback, role-plays, lectures, videos, medical scenarios</td>
<td>Patient-centred communication, inter-professional communication, teamwork, leadership and professionalism</td>
<td>3a. 84% of participants reported a new leadership position; 64% of female graduates, 43% of underrepresented minority graduates. 1a. All participants cited the course as good or very good. One third of the participants described the simulation scenario as “less than useful”.</td>
<td>8</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>Conway, 2015 (23)</td>
<td>12 US Primary Care residencies (4 locations)</td>
<td>30</td>
<td>Faculty (Residency Directors, Internal Medicine, Pediatrics)</td>
<td>6 months</td>
<td>2-day session with follow-up over 6 months</td>
<td>Didactic small group sessions, workshops, conference calls, and online learning</td>
<td>Leadership change management framework, problem-based learning, case-based seminars, competency assessment, patient-centeredness, and patient-centered medical home principles</td>
<td>Teamwork, patient safety, communications, individual and collective leadership, recognizing difference in perspective between managers and clinicians, and how to speak up to voice concerns, specifically</td>
<td>1. Scores for relevance and quality of simulations on questionnaire and free text comments rated &gt;5/10.</td>
<td>8</td>
<td>6.5</td>
</tr>
<tr>
<td>Cooper, 2011 (24)</td>
<td>Single US academic medical center</td>
<td>108</td>
<td>Physicians, nurses, allied health professionals, administrative managers</td>
<td>1 day</td>
<td>Workshop</td>
<td>Seminar, simulation, review of data from safety climate survey, team project</td>
<td>Early detection of stressors, communication, and conflict resolution</td>
<td>1. Participants were happy with the course</td>
<td>7</td>
<td>10.5</td>
<td></td>
</tr>
<tr>
<td>Agué, 2015 (25)</td>
<td>Single UK pharmacy</td>
<td>8</td>
<td>Specialty trainees (Pharmacists, Research Pharmacists, and Outsourcing Managers)</td>
<td>4 years</td>
<td>Four-year part-time programme to Master’s level with academic and vocational components</td>
<td>Diploma/MSc modules, Action learning sets, Workplace-based projects, Shadowing placements, Leadership development tools (self-reflection)</td>
<td>MNO competency: Leadership, policy, organisational development, governance</td>
<td>1. Participants reported increased self-confidence as leaders (87% of respondents) and willingness to speak up</td>
<td>7</td>
<td>10.5</td>
<td></td>
</tr>
<tr>
<td>McKidd, L, 2015 (26)</td>
<td>National UK programme</td>
<td>345</td>
<td>Residents (Specialty and GP doctors in training)</td>
<td>1 year</td>
<td>Immersive internship outside of practice with direct mentorship from national and healthcare-related organizations</td>
<td>Immersive internship, visits to other host organizations and parliament, learning on leadership and management, and action learning sets</td>
<td>Policy development, project management, research and analysis, writing and publishing, professional networking skills</td>
<td>1. Participants reported taking new approaches to their roles</td>
<td>7</td>
<td>10.5</td>
<td></td>
</tr>
<tr>
<td>Cohen, 2017 (27)</td>
<td>Multiple UK health care organizations (London)</td>
<td>68</td>
<td>Consultants (22), Registrars (35), Grade unidentifiable</td>
<td>1 day</td>
<td>Workshop</td>
<td>Lecture-style presentation: policy leadership simulation</td>
<td>Background of NHS reform and healthcare challenges</td>
<td>1. Participants reported increased self-confidence as leaders (87% of respondents) and willingness to speak up</td>
<td>7</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Ruston, 2010 (28)</td>
<td>Single UK pharmacy</td>
<td>3</td>
<td>Residents (General Practice Specialty Trainees)</td>
<td>4 months</td>
<td>2 days per week for 4 months</td>
<td>Peer learning sets, meetings, project work, reflective diaries</td>
<td>Strategic and contextual issues, commissioning, design and delivery of healthcare services, service evaluation, public health agenda, leadership, management, and partnership skills</td>
<td>1. Participants reported increased self-confidence as leaders (87% of respondents) and willingness to speak up</td>
<td>6</td>
<td>10.5</td>
<td></td>
</tr>
<tr>
<td>Other Included Studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathews, L, 2018 (29)</td>
<td>Single UK hospital</td>
<td>50</td>
<td>Medical facility personnel with supervisory or team leader responsibility</td>
<td>10 months</td>
<td>60 hours of learning over 10 month course. Two one day seminars, one two-day off-site session, seven monthly four hour meetings, reading assignments</td>
<td>Seminars, experiential activities, small group discussions, multi-source feedback, teambuilding activities, reading</td>
<td>Emotional intelligence, leadership behaviour, leadership foundations skills</td>
<td>1. 94% of participants would recommend the programme to others, training rated as worthwhile investment (8/10)</td>
<td>3</td>
<td>11.5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source (First Author, Year)</th>
<th>Setting</th>
<th>Learner Number</th>
<th>Learner Type</th>
<th>Intervention Length</th>
<th>Intervention Description</th>
<th>Teaching Methods</th>
<th>Educational Content</th>
<th>Main Findings by Level of Impact</th>
<th>JB Score</th>
<th>NSQD Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atmara, 2014 [38]</td>
<td>Family medicine residency programme (Glas)</td>
<td>56</td>
<td>Residents (family medicine year 1-4)</td>
<td>5 days</td>
<td>3-day online management workshop</td>
<td>Multiple teaching case studies, small group discussions, team exercises</td>
<td>Professional skills, interprofessional skills, practice-based learning and improvement system-based practice</td>
<td>Increase 1.4% confidence 2. Increase self-assessed risk management, conflict management, communication skills, time management, ability to write algorithms 3. Supervisor-reported increased effective use of health resources, coordination of patient care, patient communication skills</td>
<td>n/a</td>
<td>11/11</td>
</tr>
<tr>
<td>Fereidouni, 2016 [33]</td>
<td>Single US college (MED)</td>
<td>37</td>
<td>Obstetricians (junior fellows); young physicians, senior fellows, unspecified (other)</td>
<td>3 to 5 years</td>
<td>6-day national intensive leadership development for OB/GYN physicians</td>
<td>Interactive skills-building workshops, series of leadership and psychological assessment tools, including a 360-degree analysis, individual and group mentoring, structured networking, independent learning and reflection, individual leadership project</td>
<td>Organizational culture, leadership and empowerment, communication, motivation, advocacy, media, negotiation skills, health policy</td>
<td>Increase of 20% in number of publications, 30% increase in national committee chairs, 50% increase in local committee chairs, increase in number of presentations at national conferences, increase in number of leadership roles</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>Chang, 2019 [32]</td>
<td>Three US healthcare organisations in 24 US states</td>
<td>65</td>
<td>48 Physicists, 8 Radiologists, 2 Dentists, 3 Nurses, 6 Pharmacists, 1 Psychologist, 1 Social Worker, 1 Speech Pathologist</td>
<td>9 months</td>
<td>Ethics</td>
<td>Adaptive leadership making reactive change and generally, mission, vision, goals and strategies evaluation and resilience management, training influence and persuasion strategies, case studies and advice presentation skills development and marketing strategies, and stakeholders evaluation design</td>
<td>Building up business case and budget measuring impact project management</td>
<td>1</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Day, 2010 [36]</td>
<td>Single US specialty association (orthopaedics)</td>
<td>100</td>
<td>Orthopaedic surgeons</td>
<td>1 year</td>
<td>Mentoring</td>
<td>Mentoring by established orthopaedic leader</td>
<td>Unclear</td>
<td>2. Significant increase in L6 leadership domains: knowledge, tolerance for demands of leadership, and leadership positioning 3. Significant increase in the number of national committee (22% pre, 62% post-programme, p &lt; 0.001). 48% of peers advanced in academic rank vs 21% of controls (p &lt; 0.005). 1. Participants reported positive experiences with the program. Attendance was at a rate of 95% or higher. 2. 34% of respondents disagreed that the mentoring process had helped them in their professional growth. 3. 98% increased their commitment to and support of the vision and strategy of the organization</td>
<td>2b. Increase confidence in applying conflict resolution strategies, and with public speaking</td>
<td>n/a</td>
</tr>
<tr>
<td>Korsheva, 2007 [34]</td>
<td>Single Russian medical centre</td>
<td>70</td>
<td>Physicians (32), nurses, and a wide range of administrators</td>
<td>5 months</td>
<td>Five three-day sessions over five months</td>
<td>Lectures, seminars, case studies, experiential exercises, individual assessment, executive coaching, including a 360° assessment, mentoring, team project work</td>
<td>Strategic thinking and personal awareness, Leadership qualities, Leadership best practices, Strategic thinking and personal awareness, Organisational culture, leading and empowering, Professionalism</td>
<td>1. Participants reported almost unanimously (86-100%) for all measures including that the programme made significant impact locally and nationally. 2. Mean satisfaction score of 4.86/5 1. 100% of respondents indicated that they would recommend the course to colleagues. 2. 25% of respondents indicated that they had improved their interpersonal skills relative to their training and practicum. 3. Successful adoption of most projects into the organisation.</td>
<td>1. Participants reported positive experiences with the program. Attendance was at a rate of 95% or higher. 2. 34% of respondents disagreed that the mentoring process had helped them in their professional growth. 3. 98% increased their commitment to and support of the vision and strategy of the organization</td>
<td>3a. Improved relationship-building and networking skills and the development of improved interpersonal skills 3b. No clear effect on career progression. 4b. No significant increase in tangible benefits from project 5. In all elements of the programme were positively rated by participants with ratings becoming more positive as the programme progressed. 2. Increased self-awareness, greater personal resilience and improved motivation. 3. Participants “appear to have gained tremendous personal benefit” which focused on confidence to operate outside of their comfort zones, greater clarity about their leadership role, greater assurance about their own leadership capacity and sense of empowerment. 4. Participants developed an increased understanding of key policy issues and the need to develop skills in the areas of strategic influence. 5. Improved relationship building and networking skills and the development of improved influencing ability.</td>
</tr>
<tr>
<td>Source (First Author, Year)</td>
<td>Setting</td>
<td>Learner Number</td>
<td>Learner Type</td>
<td>Intervention Length</td>
<td>Intervention Description</td>
<td>Teaching Methods</td>
<td>Educational Context</td>
<td>Main Findings by Kirkpatrick Level</td>
<td>ACORN Score</td>
<td>NACDG Score</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------</td>
<td>----------------</td>
<td>--------------</td>
<td>---------------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td>-------------------</td>
<td>-------------------------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Conner, 2017 [44]</td>
<td>Single US integrated health care organization</td>
<td>30</td>
<td>Residents (including Chief Residents)</td>
<td>1 day</td>
<td>8 hrs</td>
<td>Enhanced intellectual engagement, three weeks prior to course</td>
<td>Didactics session to simulate teaching encounters</td>
<td>Landshark. Engagement, and core feedback skills, interpretation of emotional intelligence inventory, interpersonal and communication skills and professionalism</td>
<td>1. 80% agreed that the program met their learning needs. Participants reported continuing the GUTS used. 2. OSTE scores (out of 100) increased from mean = 51.42, SD = 7.80 to mean = 51.22, SD = 4.64, 1(56) = 1.99, p = 0.044</td>
<td>n/a</td>
</tr>
<tr>
<td>Patel, 2015 [50]</td>
<td>Single US hospital</td>
<td>30</td>
<td>Residents (PGY2-4)</td>
<td>2 years</td>
<td>2 year healthcare leadership in quality residency track</td>
<td>Enhanced intellectual engagement, three weeks prior to course</td>
<td>Leadership, management, and core feedback skills, interpretation of emotional intelligence inventory, interpersonal and communication skills and professionalism</td>
<td>1. The core curriculum was rated very highly valuable and necessary (mean 4.65 and 6.865 respectively). 2. Increased understanding of quality improvement knowledge assessment was at 5 points for cohort 1 (50.3), 4 points for cohort 2 (50.1). Improvements were marked by the track directors. No significance reported</td>
<td>2. Participants reported improved skills and knowledge. 3a. Participants reported use of new knowledge and skills in both current and new leadership roles. 3b. All graduates have remained in health leadership positions (grades 86% (53/62)) have opted to take on new responsibilities. 4a. Individual projects completed on the programme had a range of positive outcomes. These amounted primarily to official recommendations and initiation of new pathways, but several of the projects are reported to have improved patient care.</td>
<td>n/a</td>
</tr>
<tr>
<td>Nakajako, 2015 [49]</td>
<td>Four African and 4 US universities</td>
<td>55</td>
<td>Unclear</td>
<td>1 year</td>
<td>1 year fellowship</td>
<td>Enhanced intellectual engagement, three weeks prior to course</td>
<td>Leadership, communication, monitoring and evaluation, health information, research methodology, grant writing, implementation science, and responsible conduct of research</td>
<td>2. Participants felt empowered to start service improvement projects. 2b. Statistically significant increase in self-reported understanding of several domains</td>
<td>2. Significant improvement in participants’ knowledge for all modules (p&lt;0.001). 3b. Participants have received both local and national awards recognizing their leadership and commitment to the community. Nine graduates are in positions of leadership such as medical directorships. 4b. All graduates have received grants to support their projects. Multiple projects have achieved sustainable funding and impact across advocacy, health programme development and policies.</td>
<td>n/a</td>
</tr>
<tr>
<td>Kuo, 2010 [41]</td>
<td>Single US residency program</td>
<td>16</td>
<td>Residents (Pediatrics)</td>
<td>3 years</td>
<td>Baquep residency programme</td>
<td>Leadership, communication, monitoring and evaluation, health information, research methodology, grant writing, implementation science, and responsible conduct of research</td>
<td>1. Participants evaluated five (five point scale) the content and speakers (access from 4 to 6.5); participants showed significant improvement in their leadership and management test scores (61/67 to 98% in two different cohorts)</td>
<td>1. Participants evaluated five (five point scale) the content and speakers (access from 4 to 6.5); participants showed significant improvement in their leadership and management test scores (61/67 to 98% in two different cohorts)</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Brandeis, 2013 [42]</td>
<td>Single US residency program (radiology)</td>
<td>44</td>
<td>Residents</td>
<td>1 year</td>
<td>7x 60min modules</td>
<td>Leadership, communication, monitoring and evaluation, health information, research methodology, grant writing, implementation science, and responsible conduct of research</td>
<td>2. Participants reported improved skills and knowledge. 3a. Participants reported use of new knowledge and skills in both current and new leadership roles. 3b. All graduates have remained in health leadership positions in Uganda (grades 86% (32/37)) have opted to take on new responsibilities.</td>
<td>2. Participants reported improved skills and knowledge. 3a. Participants reported use of new knowledge and skills in both current and new leadership roles. 3b. All graduates have remained in health leadership positions in Uganda (grades 86% (32/37)) have opted to take on new responsibilities.</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Green, 2008 [43]</td>
<td>US network of community-owned health care providers and physicians</td>
<td>26</td>
<td>Teams from eight organizational units</td>
<td>2 years</td>
<td>Coaching and leadership initiative</td>
<td>Coaching and leadership initiative</td>
<td>Leadership and management basics, managing change and interpersonal skills, professional issues, quality, information, finance, and a capstone service</td>
<td>2a. Participants felt empowered to start service improvement projects. 2b. Statistically significant increase in self-reported understanding of several domains</td>
<td>2b. Participants showed significant improvement in their leadership and management test scores (61/67 to 98% in two different cohorts)</td>
<td>n/a</td>
</tr>
<tr>
<td>Hemmer, 2007 [47]</td>
<td>Single US residency/ fellowship program (pathology)</td>
<td>16</td>
<td>Residents and fellows</td>
<td>1 year</td>
<td>6 x 2-3 day workshops (average 20hrs per workshop)</td>
<td>Coaching and leadership initiative</td>
<td>Leadership and management basics, managing change and interpersonal skills, professional issues, quality, information, finance, and a capstone service</td>
<td>1. Participants evaluated five (five point scale) the content and speakers (access from 4 to 6.5); participants showed significant improvement in their leadership and management test scores (61/67 to 88% in two different cohorts)</td>
<td>1. Participants evaluated five (five point scale) the content and speakers (access from 4 to 6.5); participants showed significant improvement in their leadership and management test scores (61/67 to 88% in two different cohorts)</td>
<td>n/a</td>
</tr>
<tr>
<td>McCurdy, 2008 [48]</td>
<td>Single US academic medical centre</td>
<td>22</td>
<td>Faculty members who went, at the time of the course, in a leadership position or likely to move into a leadership position soon</td>
<td>12 months</td>
<td>Multiple-year programme (6 months per cohort)</td>
<td>Coaching and leadership initiative</td>
<td>Leadership and management basics, managing change and interpersonal skills, professional issues, quality, information, finance, and a capstone service</td>
<td>2a. Significant improvement in participants’ knowledge for all modules (p&lt;0.001). 3b. Participants have received both local and national awards recognizing their leadership and commitment to the community. Nine graduates are in positions of leadership such as medical directorships. 4b. All graduates have received grants to support their projects. Multiple projects have achieved sustainable funding and impact across advocacy, health programme development and policies.</td>
<td>2a. Significant improvement in participants’ knowledge for all modules (p&lt;0.001). 3b. Participants have received both local and national awards recognizing their leadership and commitment to the community. Nine graduates are in positions of leadership such as medical directorships. 4b. All graduates have received grants to support their projects. Multiple projects have achieved sustainable funding and impact across advocacy, health programme development and policies.</td>
<td>n/a</td>
</tr>
<tr>
<td>Hadley, 2014 [46]</td>
<td>Single US training agency</td>
<td>30</td>
<td>Residents (PGY2 doctors paired with a management training)</td>
<td>6 months</td>
<td>Project work, mentoring, action learning sets</td>
<td>Coaching and leadership initiative</td>
<td>Leadership and management basics, managing change and interpersonal skills, professional issues, quality, information, finance, and a capstone service</td>
<td>1. Participants evaluated five (five point scale) the content and speakers (access from 4 to 6.5); participants showed significant improvement in their leadership and management test scores (61/67 to 88% in two different cohorts)</td>
<td>1. Participants evaluated five (five point scale) the content and speakers (access from 4 to 6.5); participants showed significant improvement in their leadership and management test scores (61/67 to 88% in two different cohorts)</td>
<td>n/a</td>
</tr>
<tr>
<td>Rewen, 2015 [47]</td>
<td>Single US hospital network</td>
<td>50</td>
<td>Senior physicians “chairs and mentors”</td>
<td>6 months</td>
<td>Leadership, discussions, projects</td>
<td>Coaching and leadership initiative</td>
<td>Leadership and management basics, managing change and interpersonal skills, professional issues, quality, information, finance, and a capstone service</td>
<td>3. Participants reported being more engaged in their conversations, and adapting learning day-to-day</td>
<td>3. Participants reported being more engaged in their conversations, and adapting learning day-to-day</td>
<td>n/a</td>
</tr>
<tr>
<td>Osborn, 2004 [49]</td>
<td>Two US Pediatric Association locations</td>
<td>22 (total 2 columns)</td>
<td>Pediatricians from AMCs</td>
<td>2-3 years</td>
<td>Workshops only</td>
<td>Coaching and leadership initiative</td>
<td>Leadership and management basics, managing change and interpersonal skills, professional issues, quality, information, finance, and a capstone service</td>
<td>2b. Reported increased knowledge and skills in all areas related to giving workshops</td>
<td>2b. Reported increased knowledge and skills in all areas related to giving workshops</td>
<td>n/a</td>
</tr>
</tbody>
</table>

BMJ Leader


Supplemental material BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

<table>
<thead>
<tr>
<th>Source (First Author, Year)</th>
<th>Setting</th>
<th>Learner Number</th>
<th>Learner Type</th>
<th>Intervention Length</th>
<th>Intervention Description</th>
<th>Teaching Methods</th>
<th>Educational Context</th>
<th>Main Findings</th>
<th>JBI Score</th>
<th>BMJ Leader Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welch, 2009 (48)</td>
<td>Single US hospital department</td>
<td>6</td>
<td>Residents (Psychiatry PGY1)</td>
<td>8 months</td>
<td>Weekly seminars (8 modules x 4 sessions each), with projects and mentoring</td>
<td>Lectures, discussions, projects, mentoring</td>
<td>Financial management, human resources management, Planning and marketing, Information management, Risk management, Governance and organizational dynamics, Business and clinical operations, Professional responsibility</td>
<td>4b. Project resulted in decreased non-attendance by new patients by 50% across 12 months</td>
<td>9/5</td>
<td>n/a</td>
</tr>
<tr>
<td>Monaghan, 2018 (50)</td>
<td>Single UK hospital trust</td>
<td>12 (6 doctors)</td>
<td>Doctors (non-clinical training and clinical fellows, managers of various departments)</td>
<td>6 months</td>
<td>4-month paired learning doctors/managers</td>
<td>6-month paired learning doctors/managers</td>
<td>2a. Physician participants reported feeling more prepared for a range of leadership requirements, including understanding decisions and working in teams and with managers</td>
<td>4</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Voigt, 2016 (51)</td>
<td>Six Dutch teaching hospitals</td>
<td>68 (90)</td>
<td>Residents (Radiology 2012)</td>
<td>1 year</td>
<td>Four 1hr meetings to discuss and work on projects</td>
<td>Facilitated discussions, project work</td>
<td>Quality improvement and leadership, not otherwise specified</td>
<td>2a. Interviewees reported feeling empowered. 2b. Interviewees reported increased awareness of organizational aspects of healthcare delivery</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Henkova, 2017 (52)</td>
<td>Single US hospital department</td>
<td>98</td>
<td>Residents (radiology)</td>
<td>NR (variable, journal club)</td>
<td>One hour long journal club meeting every fortnight</td>
<td>Journal club, projects, mentoring, leadership role placement</td>
<td>Leadership (topics chosen by the group as an ad hoc basis)</td>
<td>0</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Pearson, 2018 (53)</td>
<td>NHS leadership fellowship</td>
<td>12</td>
<td>Residents (senior, medicine, general practice, surgery, obstetrics and gynaecology, paediatric surgery and psychiatry)</td>
<td>1 year</td>
<td>1-year on-site programme: fellowship in a host organization</td>
<td>Symposium and conferences, one-to-one coaching sessions, action learning sets, shadowing opportunities and reflective practice including completion of a portfolio, Project work for host organizations</td>
<td>Minimal benefits. Communication, working styles and leadership framework mentioned</td>
<td>0</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Crinic, 2004 (54)</td>
<td>Single UK residency program (internal medicine and paediatrics)</td>
<td>13</td>
<td>Residents (PGY1-4)</td>
<td>1 year</td>
<td>Monthly seminar series</td>
<td>Interactive lectures</td>
<td>N/A</td>
<td>N/A</td>
<td>9/5</td>
<td>n/a</td>
</tr>
<tr>
<td>Dickie, 2014 (55)</td>
<td>Single US hospital</td>
<td>Unclear</td>
<td>Residents (Psychiatry PGY1-4)</td>
<td>4 years</td>
<td>Modular leadership programme over 4 years with variable components</td>
<td>Seminars, with voluntary simulations, action teams work, electives, mentoring</td>
<td>Philosophy of leadership, healthcare delivery systems, quality assurance, risk management, qualities of exceptional leaders</td>
<td>3b. 35% had assumed leadership roles within three years. Publication of 27 articles in core radiology journals, completion of multiple projects. 4b. New external collaboration for residents. 1. Mean participant reported that all components of the course were very or slightly useful. 70% did not access multi-source feedback component. 2a. Most participants reported improved attitudes towards leadership and their ability to make decisions in their organisation. Reports of increased confidence. 2b. Reports of increased awareness of new working styles and characteristics of good leadership.</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Foster, 2008 (56)</td>
<td>Single US medical centre</td>
<td>12</td>
<td>Residents</td>
<td>2 years</td>
<td>2-week intensive orientation, 11-week MPhil degree and leadership course, weekly U2-day didactic sessions, monthly journal club, monthly open evening sessions, 2-year leadership programme</td>
<td>MPhil degree, leadership course, change project, mentoring</td>
<td>Leadership of small systems in health care</td>
<td>2a. Participants reported increased confidence across a range of key curriculum areas. 3b. Participants reported the programme had helped them gain new leadership roles, and that they had applied learning from the programme in their new leadership roles</td>
<td>4b. Projects completed</td>
<td>9/5</td>
</tr>
<tr>
<td>Freeman, 2018 (57)</td>
<td>Single UK training programme</td>
<td>30</td>
<td>Cardiology fellows-in-training and early career professionals</td>
<td>2 years</td>
<td>2-year cardiology leadership course</td>
<td>Mentoring, not otherwise clear</td>
<td>Leadership of small systems in health care</td>
<td>2a. Participants reported increased confidence across a range of key curriculum areas. 2b. Participants reported increased confidence in their ability to make decisions in their organisation. 3b. Participants reported the programme had helped them gain new leadership roles, and that they had applied learning from the programme in their new leadership roles. 4a. Of those that acquired a new leadership opportunity, 90% of respondents reported participating in the leadership academy program had an impact on their success within the newly acquired leadership role. 4b. Participants rated the course effectiveness 4.1/5 (based on video recorded simulations, no control group)</td>
<td>9/5</td>
<td>n/a</td>
</tr>
<tr>
<td>Saroro, 2017 (58)</td>
<td>Single German university hospital</td>
<td>50</td>
<td>Residents (PGY1-4 across specialties)</td>
<td>4 weeks</td>
<td>Weekly 2-3hr sessions after clinical duties</td>
<td>Didactic modules, standardised situational simulations, one-on-one feedback on recorded simulations, &quot;practicing communication techniques&quot;</td>
<td>&quot;Full Range Leadership Model&quot; (Bass), transformational and transactional leadership, simulation of critical incidents, communication techniques</td>
<td>2a. No change in knowledge or skills after the leadership courses compared with control (both had small increases in mean scores)</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Schub, 2015 (59)</td>
<td>US surgical department specialty</td>
<td>9</td>
<td>Residents (DIT)</td>
<td>6 months</td>
<td>Virtual strength assessments (VTSA); mentorship meetings; &quot;thought of the day&quot; internal and external faculty training and development; leadership basic training course</td>
<td>Virtual leadership, curriculum not otherwise specified</td>
<td>Leadership of small systems in health care</td>
<td>2b. No significant change in attitude towards leadership</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Steaker, 2007 (60)</td>
<td>Single UK hospital network</td>
<td>Unclear (beyond 400)</td>
<td>Emerging physician leaders</td>
<td>9 months</td>
<td>Eight sessions offered roughly once monthly on Fridays at an off-site centre</td>
<td>Seminars, business case project</td>
<td>Market was healthcare, Healthcare finance, Writing a business plan, Emotional intelligence, situational leadership, conflict resolution and negotiation, Medico-legal issues</td>
<td>3b. All business plans were submitted over 13 courses. 4b. Of all business plans that have been implemented</td>
<td>9/5</td>
<td>n/a</td>
</tr>
<tr>
<td>Source</td>
<td>First Author, Year</td>
<td>Setting</td>
<td>Learner Number</td>
<td>Learner Type</td>
<td>Intervention Length</td>
<td>Intervention Description</td>
<td>Teaching Methods</td>
<td>Educational Content</td>
<td>Main Findings by Kirkpatrick Level</td>
<td>GB Score</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------</td>
<td>---------</td>
<td>----------------</td>
<td>--------------</td>
<td>---------------------</td>
<td>--------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>-------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Lyons, O. et al.</td>
<td>BMJ Leader 2020:0:1–8. doi: 10.1136/leader-2020-000360</td>
<td>BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)</td>
<td><strong>Table 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2b. Increased knowledge of desirable leadership characteristics in the organisation 4.3
3b. Projects completed as part of the programmes.
3. Participants dissatisfied with the programmes, partly due to high expectations not being met. Roll-off in attendance
4a. Participants reported increased self-awareness
1. High ratings for the course 4.63/5 for comparison to other leadership and education programmes experienced.
1. "universal praise" from participants for the mandatory first 2 years, 100% opted into the optional years 3-4.
1. “universal praise” from participants for the mandatory first 2 years, 100% opted into the optional years 3-4. 2b. If participants have worked on projects as part of the programme, with one having submitted academic manuscript resulting from her project.
<table>
<thead>
<tr>
<th>Source (First Author, Year)</th>
<th>Setting</th>
<th>Learner Number</th>
<th>Learner Type</th>
<th>Intervention Length</th>
<th>Intervention Description</th>
<th>Teaching Methods</th>
<th>Educational Context</th>
<th>Main Findings by Kirkpatrick Level</th>
<th>JBI Score</th>
<th>MERSQI Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>program in Australia</td>
<td>residents and 14 resident medical officers</td>
<td>24</td>
<td>10 months</td>
<td>56 sessions 60-120 minutes</td>
<td>Large group lectures, discussion and reflection on action learning (planning or patient safety)</td>
<td>trainee presentations at workshops</td>
<td>Leadership competency framework; team building, feedback, and action planning</td>
<td>All workshops rated between 3.5-4</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Danglay, 2018 [94]</td>
<td>Single UK hospital trust</td>
<td>NR (n=50)</td>
<td>Specialty trainees ST4-5TH</td>
<td>3 hrs</td>
<td>Case scenarios, role playing</td>
<td>Interactive discussions and self-reflection</td>
<td>Leadership competency, strategic planning and vision, financial management, business planning, communication skills, change management, quality improvement, negotiation and problem solving, leadership development</td>
<td>2a. Improved confidence, intention to apply&lt;br&gt;2b. Improved leadership understanding in a range of domains&lt;br&gt;3a. Completion of 5 case studies</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Frigel, 2010 [92]</td>
<td>Single UK hospital department</td>
<td>29</td>
<td>Residents (Pediatric Oncology)</td>
<td>NR</td>
<td>Twice monthly seminar series</td>
<td>Guided reflection on challenging leadership experiences</td>
<td>Leadership training; educational content</td>
<td>&lt;p&gt;1. 70% of fellows continue to participate in later years despite no longer being mandatory.&lt;br&gt;2a. Significant increase in confidence in all items on the Hospital Questionnaire. Examples include working effectively as a team, effective communication, self-awareness.&lt;br&gt;2b. Participants agreed that the information was relevant to their future careers&lt;/p&gt;</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>Hunt, 2017 [94]</td>
<td>Single UK training camp</td>
<td>29</td>
<td>Residents (trainees at the start of F1 year surgery)</td>
<td>NR</td>
<td>Surgical skills boot camp</td>
<td>Individual, small group, and workshop</td>
<td>Leadership training; educational content</td>
<td>&lt;p&gt;1. Significant increase in self-awareness (only 53% agree or strongly agree that understanding of own behaviours and motivations improved immediately post, only 46% agree or strongly agree at 4 months)&lt;br&gt;2c. Increase in self-assessed competency in all the 26 categories in each of the program’s five cohorts (significant change not required)&lt;br&gt;3c. Commentaries and assessments revealed an increasing level of empowerment in their leadership role and increased desire for selection to leadership roles&lt;/p&gt;</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>MacDonald, 2011 [97]</td>
<td>Programme across 3 UK states</td>
<td>&lt;p&gt;n=30&lt;br&gt;not specific to five cohorts of 30 or more&lt;/p&gt;</td>
<td>Community practice physicians/five cohorts</td>
<td>Weekly three-hour evening sessions</td>
<td>Weekly three-hour evening sessions</td>
<td>The business of medicine, quality improvement, transformational leadership</td>
<td>2b. Respondents reported increased personal leadership ability, sustained at 6 months. (5.2/6 - 3a).&lt;br&gt;3b. Completion of some clinical projects</td>
<td>7.5</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>Tasker, 2018 [98]</td>
<td>Single UK hospital trust</td>
<td>25</td>
<td>Residents (PSYCH 5 Psychiatry residents)</td>
<td>1 year</td>
<td>Six full day workshops: mentoring programme</td>
<td>Medical leadership competency framework; didactic teaching, simulation, role playing, small group work, case discussion</td>
<td>1. All workshops rated between 3-4&lt;br&gt;2a. Participants felt the programme helped them in achieving leadership competencies&lt;br&gt;2b. Community-based teaching was reinforced with leadership champions and then applied to their projects&lt;br&gt;3a. All participants completed and passed a formalised assessment at the end of the programme&lt;br&gt;3b. Two participants reported increased leadership skills, knowledge and self-awareness. They were less successful at: delegating; saying ‘no’; adopting different leadership styles; and evaluating meetings.&lt;br&gt;4b. One participant was noted to have achieved a change implementation through their project (not otherwise specified).&lt;br&gt;5. Mean rating of 4.75 (post) and 4.65 (retrospectively)&lt;br&gt;6. Respondents reported increased self-awareness, which was sustained at 6-month follow-up. (4.5/6)&lt;br&gt;7. Respondents reported increased proactivity in management</td>
<td>4</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Patterson, 2013 [98]</td>
<td>Single UK GP deanery (Yorkshire &amp; Humber)</td>
<td>5</td>
<td>Residents (GP trainees, F1)</td>
<td>8 months</td>
<td>Facilitated leadership projects</td>
<td>Leadership, change management, and teamwork skills</td>
<td>Leadership training; educational content</td>
<td>1. 100% of fellows would recommend the programme to peers&lt;br&gt;2. Participants reported increased personal leadership ability, sustained at 6 months. (5.2/6 - 3a)&lt;br&gt;3. Participants described how the conceptual issues learned in the workshops were reinforced with leadership champions and then applied to their projects&lt;br&gt;4a. Participants reported increased personal leadership ability. (4.5/6 - 3a)&lt;br&gt;4b. Increased leadership skills, knowledge and self-awareness. They were less successful at: delegating; saying ‘no’; adopting different leadership styles; and evaluating meetings.</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Wise, 2016 [90]</td>
<td>Single Israel health provider</td>
<td>256</td>
<td>Physician-managers</td>
<td>8 weeks</td>
<td>3.5 full days over 2 weeks with one overnight; 3 full-day meetings at 3 and 5 weeks</td>
<td>Theoretical knowledge, experiential learning, practical tools, deep personal exercises and simulations, individual, dyadic, and group learning</td>
<td>Models of self-awareness, outcome thinking, determining a personal and organisational vision, and creating a personal approach to leadership</td>
<td>3a. Participants felt the programme helped them in achieving leadership competencies&lt;br&gt;3b. Residents described how the conceptual issues learned in the workshops were reinforced with leadership champions and then applied to their projects&lt;br&gt;3b. Residents described how the conceptual issues learned in the workshops were reinforced with leadership champions and then applied to their projects&lt;br&gt;3b. Residents described how the conceptual issues learned in the workshops were reinforced with leadership champions and then applied to their projects&lt;br&gt;3b. Residents described how the conceptual issues learned in the workshops were reinforced with leadership champions and then applied to their projects&lt;br&gt;3b. Residents described how the conceptual issues learned in the workshops were reinforced with leadership champions and then applied to their projects</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Steinert, 2003 [96]</td>
<td>Single Canadian department</td>
<td>16</td>
<td>Faculty (family medicine)</td>
<td>2 days</td>
<td>Two-day workshop</td>
<td>Time management, goals and priorities, leadership styles and skills, and conducting effective meetings</td>
<td>1. All participants rated the workshop as “very useful”&lt;br&gt;2. Several the participants reported that they would change their behaviour after the workshop, regarding time management, goals and meetings&lt;br&gt;3d. Most respondents had successfully attempted determining short-term goals; handling paper more effectively; determining their ‘prime time’; protecting time for specific tasks; and setting meeting agendas. They were less successful at: delegating; saying ‘no’; adopting different leadership styles; and evaluating meetings.</td>
<td>3</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Sataran, 2014 [92]</td>
<td>Single UK hospital department</td>
<td>24</td>
<td>Surgeons who wish to assume administrative or leadership roles</td>
<td>18 months</td>
<td>46th annual VA house staff retreat</td>
<td>Seminars, project work (teams)</td>
<td>Leadership competency, strategic planning and vision, financial management, business planning, communication skills, change management, quality improvement, negotiation and problem solving, leadership development and coaching, didactic content and skill development</td>
<td>1. Participants believed the course should be offered again and had a positive effect on their professional lives. 2. All participants rated workshop as “very useful”&lt;br&gt;2. Several the participants reported that they would change their behaviour after the workshop, regarding time management, goals and meetings&lt;br&gt;3d. Most respondents had successfully attempted determining short-term goals; handling paper more effectively; determining their ‘prime time’; protecting time for specific tasks; and setting meeting agendas. They were less successful at: delegating; saying ‘no’; adopting different leadership styles; and evaluating meetings.</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Lee, 2004 [93]</td>
<td>Single UK residency programme (paediatrics, Hawaii)</td>
<td>10</td>
<td>Residents (PGY2)</td>
<td>3hrs</td>
<td>3-hour interactive workshop during resident retreat</td>
<td>Case scenario, role playing, interactive discussions and self-reflection</td>
<td>Leadership competencies, strategic planning and vision, financial management, business planning, communication skills, change management, quality improvement, negotiation and problem solving, leadership development and coaching, didactic content and skill development</td>
<td>1. Increased confidence, intention to apply&lt;br&gt;2. Respondents reported increased leadership skills, knowledge and self-awareness. They were less successful at: delegating; saying ‘no’; adopting different leadership styles; and evaluating meetings.</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Satterfield, 2004 [94]</td>
<td>Single UK department of psychiatry</td>
<td>13</td>
<td>Residents and post-doctoral fellows</td>
<td>5 months</td>
<td>8 &amp; 15 seminars</td>
<td>Seminars with guest speakers and case discussions</td>
<td>Exploration of leadership roles in research, clinical practice, teaching, and administrative organizational dynamics and gender, negotiation skills and conflict resolution strategies, role of consultation, communication, peer support, and mentoring by peers and women in the development of leadership skills; &quot;keeping the balance of work, relationships, and personal health.&quot;</td>
<td>1. Participants believed the course should be offered again and had a positive effect on their professional lives.&lt;br&gt;2. One participant said it encouraged them to seek out mentors</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Awad, 2004 [92]</td>
<td>Single UK residency program (surgery)</td>
<td>NR</td>
<td>Residents (Surgery)</td>
<td>6 months</td>
<td>N/A</td>
<td>Education and learning strategies, conflict resolution, serving as teacher, time management, delegation, leadership styles, managing stress, safety and quality, team building, feedback and action planning</td>
<td>2b. Statistically significant increase in score on a 3–6 item internal stress scorecard alignment ≥30%, communication ≥30%, and integrity ≥30%</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

---

BMJ Leader

<table>
<thead>
<tr>
<th>Source (First Author, Year)</th>
<th>Setting</th>
<th>Learner Number</th>
<th>Learner Type</th>
<th>Intervention</th>
<th>Description</th>
<th>Teaching Methods</th>
<th>Educational Context</th>
<th>Main Finding by Kirkpatrick level</th>
<th>AB Score</th>
<th>NROD Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayrl, 2008 [86]</td>
<td>Single US residency program (family medicine)</td>
<td>NR</td>
<td>Residents (Family Medicine PGY-2)</td>
<td>2 years</td>
<td>Nine half-day sessions (first year); monthly one-half sessions (second year)</td>
<td>Interactive lectures and group assignments</td>
<td>Teaches through assignment, confrontation, and integrity. A practice management curriculum: determining/balancing personal and professional goals, practice opportunities, facilities, organization, operation and management, staff policies, legal issues, marketing, resources and hospital issues.</td>
<td>1. Reported that the course was beneficial, a positive experience. 2a. Increased confidence in their abilities to lead a ward team (p = .0002) and felt better prepared to deal with the challenges of being upper-level residents (mean = 3.65, SD .61). The participants also believed that they would use what they learned at this interval as upper-level residents (mean = 3.48, SD .33). 2b. Gained confidence in time management, negotiation and leadership skills. 3a. 60% of participants felt that their leadership skills had improved as a result of the feedback received.</td>
<td>n/a</td>
<td>7</td>
</tr>
<tr>
<td>Hanss, 2012 [97]</td>
<td>Single Canadian hospital</td>
<td>43</td>
<td>Residents (PGY-5 &amp; Surgery)</td>
<td>1 day</td>
<td>One-day conference</td>
<td>Interactive lectures and case-based discussions; role-playing, scenario simulation; small-group problem-solving sessions; and live feedback sessions.</td>
<td>Teaching Methods: mentorship, feedback, and managing personal finances, feeling/making risk, and managing a private practice. Leadership MDS</td>
<td>1. 70% felt that management was well-addressed or very well-addressed of 5% before the course. 2a. Participants reported improved skills in giving feedback, delegating tasks, coping with stress, effective learning while on duty, teaching bedside and in the OR, and managing conflicts. 2b. Participants reported improved skills in giving feedback, delegating tasks, coping with stress, effective learning while on duty, teaching bedside and in the OR, and managing conflicts. 3a. 60% of participants felt that their leadership skills had improved as a result of the feedback received.</td>
<td>n/a</td>
<td>7</td>
</tr>
<tr>
<td>Patel, 2019 [98]</td>
<td>Single US hospital residency</td>
<td>17</td>
<td>Residents (PGY-2)</td>
<td>2 years</td>
<td>Quarterly mentoring meetings for 2 years as part of a healthcare leadership program: Weekly didactic sessions.</td>
<td>Mentoring</td>
<td>Teaching Methods: monthly sessions of one-hour sessions.</td>
<td>1. Reported that the course was beneficial, a positive experience. 2a. Increased confidence in their abilities to lead a ward team (p = .0002) and felt better prepared to deal with the challenges of being upper-level residents (mean = 3.65, SD .61). The participants also believed that they would use what they learned at this interval as upper-level residents (mean = 3.48, SD .33). 2b. Gained confidence in time management, negotiation and leadership skills. 3a. 60% of participants felt that their leadership skills had improved as a result of the feedback received.</td>
<td>n/a</td>
<td>7</td>
</tr>
<tr>
<td>Stefc, 2011 [99]</td>
<td>Single US hospital</td>
<td>24</td>
<td>Residents (Senior)</td>
<td>4 weeks</td>
<td>Monthly didactic sessions.</td>
<td>Didactic teaching, assignment.</td>
<td>Teaching Methods: case-based learning, reflections, and discussions, scenarios and role play.</td>
<td>1. Participants reported high satisfaction with the course, and most would recommend it to colleagues. 2a. Participants felt more interested and prepared for leadership responsibilities. 3a. 70% of participants reported having changed their approaches to projects or problems as a result of the course.</td>
<td>n/a</td>
<td>7</td>
</tr>
<tr>
<td>Gagliano, 2010 [100]</td>
<td>Single US hospital</td>
<td>50</td>
<td>Physicians with leadership responsibilities</td>
<td>2 years</td>
<td>Monthly didactic sessions.</td>
<td>Lectures and case-based discussion.</td>
<td>Teaching Methods: interactive lectures and group sessions.</td>
<td>1. Participants enjoyed the course and found it interesting. 2a. Self-reported new knowledge. Self-reported reassessment of own personal skills. 2b. Gained confidence in time management, negotiation and leadership skills. 3a. 60% of participants felt that their leadership skills had improved as a result of the feedback received.</td>
<td>0</td>
<td>6.5</td>
</tr>
<tr>
<td>Gammer, 2014 [101]</td>
<td>Single US hospital</td>
<td>8</td>
<td>Residents</td>
<td>5 months</td>
<td>Weekly one-hour sessions</td>
<td>Didactic teaching, assignment.</td>
<td>Teaching Methods: 1 per month 1-hour sessions.</td>
<td>1. Participants enjoyed the course and found it interesting. 2a. Self-reported new knowledge. Self-reported reassessment of own personal skills. 2b. Gained confidence in time management, negotiation and leadership skills. 3a. 60% of participants felt that their leadership skills had improved as a result of the feedback received.</td>
<td>0</td>
<td>6.5</td>
</tr>
<tr>
<td>Whest, 2015 [102]</td>
<td>Single US institution</td>
<td>20</td>
<td>Residents (Internal medicine, surgery, emergency, Med)</td>
<td>1 week</td>
<td>1 week programme</td>
<td>Case-based learning interactive task; small group sessions; session/evaluation and feedback.</td>
<td>Teaching Methods: interactive lectures and group sessions.</td>
<td>1. All topics rated &gt;8/10</td>
<td>n/a</td>
<td>6.5</td>
</tr>
<tr>
<td>Hadley, 2015 [103]</td>
<td>Single UK Emergency</td>
<td>NR (40 full-time staff)</td>
<td>Residents (FY2 doctors)</td>
<td>Single brief intervention</td>
<td>Leadership assessment and feedback</td>
<td>Leadership assessment and feedback</td>
<td>Teaching Methods: interactive lectures and group sessions.</td>
<td>1. Participants enjoyed the course and found it interesting. 2a. Self-reported new knowledge. Self-reported reassessment of own personal skills. 2b. Gained confidence in time management, negotiation and leadership skills. 3a. 60% of participants felt that their leadership skills had improved as a result of the feedback received.</td>
<td>n/a</td>
<td>6.5</td>
</tr>
<tr>
<td>Khasy, 2001 [104]</td>
<td>Single US residency program (internal medicine)</td>
<td>NR</td>
<td>Residents (PGY1)</td>
<td>1 day</td>
<td>8 hour retreat</td>
<td>Lectures and small group tasks and discussions, scenarios and role play.</td>
<td>Setting personal vision, leadership vs. management, building a team, practical negotiation and roles and responsibilities as upper-level residents (mean = 3.65, SD .61). The participants also believed that they would use what they learned at this interval as upper-level residents (mean = 3.48, SD .33). 2b. Gained confidence in time management, negotiation and leadership skills. 3a. 60% of participants felt that their leadership skills had improved as a result of the feedback received.</td>
<td>n/a</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>Nies, 2018 [105]</td>
<td>Single US residency programme</td>
<td>NR</td>
<td>Residents (Anesthesiology)</td>
<td>1 year</td>
<td>Monthly 1hr didactic sessions</td>
<td>Didactic teaching, assignment</td>
<td>Personal branding, curriculum vitae, marketing, networking, evaluating and evaluating different types of medical practice, medical staff structure, governance, healthcare reform, future trends in medicine.</td>
<td>1. Increased confidence in their abilities to lead a ward team (p = .0002) and felt better prepared to deal with the challenges of being upper-level residents (mean = 3.65, SD .61). The participants also believed that they would use what they learned at this interval as upper-level residents (mean = 3.48, SD .33). 2b. Gained confidence in time management, negotiation and leadership skills. 3a. 60% of participants felt that their leadership skills had improved as a result of the feedback received.</td>
<td>n/a</td>
<td>6.5</td>
</tr>
<tr>
<td>Schwart, 2014 [106]</td>
<td>International (US and Canada) Psychiatry leadership conference</td>
<td>S41</td>
<td>Residents (all US and Canadian residency programs)</td>
<td>3 days</td>
<td>3-day immersion course</td>
<td>Large and small group sessions, group tasks, peer and teacher feedback.</td>
<td>Psychological challenges in leadership situations, personal conflicts, self-reflection and self-awareness, group process, conflict resolution, navigation of challenging leadership roles.</td>
<td>1. The attendance rate was 86% and the graduation rate was 86%.</td>
<td>n/a</td>
<td>6.5</td>
</tr>
<tr>
<td>Enya Cole, 2018 [107]</td>
<td>Single US hospital</td>
<td>10</td>
<td>Physicians NOS</td>
<td>6 months</td>
<td>2hrs once a fortnight</td>
<td>Multi-source feedbacks,</td>
<td>Teaching Methods: program (internal medicine).</td>
<td>1. Reported that the course was beneficial, a positive experience. 2a. Increased confidence in their abilities to lead a ward team (p = .0002) and felt better prepared to deal with the challenges of being upper-level residents (mean = 3.65, SD .61). The participants also believed that they would use what they learned at this interval as upper-level residents (mean = 3.48, SD .33). 2b. Gained confidence in time management, negotiation and leadership skills. 3a. 60% of participants felt that their leadership skills had improved as a result of the feedback received.</td>
<td>n/a</td>
<td>6.5</td>
</tr>
</tbody>
</table>

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s).
<table>
<thead>
<tr>
<th>Source (First Author, Year)</th>
<th>Setting</th>
<th>Learner Number</th>
<th>Learner Type</th>
<th>Intervention Length</th>
<th>Intervention Description</th>
<th>Teaching Methods</th>
<th>Educational Content</th>
<th>Main Findings by Kirkpatrick level</th>
<th>JBI Score</th>
<th>MERSQI Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lyons, et al. (2011)</td>
<td>Single US hospital</td>
<td>NR</td>
<td>Residents (PGY1-3)</td>
<td>5.5 hours</td>
<td>2.5 hour workshop delivered by instructors</td>
<td>Role-play, discussion, games</td>
<td>Principles of leadership, innovative, creative problem solving, and communication techniques</td>
<td>1. “nearly 100% numerical rating of 5”&lt;br&gt;2a. Participants reported increased ability to act in crisis situations&lt;br&gt;3a. Participants reported increased ability to be confident and vulnerable as a leader&lt;br&gt;4. Nearly 100% of faculty attended basic programme. Overall reaction post-session was positive to extremely positive.</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Thoresen, et al. (2010)</td>
<td>Single US academic department</td>
<td>NR</td>
<td>Residents (PGY1-3)</td>
<td>3 months</td>
<td>2.5 hour workshop delivered by instructors</td>
<td>Role-play, discussion, games</td>
<td>Principles of leadership, innovative, creative problem solving, and communication techniques</td>
<td>1. “nearly 100% numerical rating of 5”&lt;br&gt;2a. Participants reported increased ability to act in crisis situations&lt;br&gt;3a. Participants reported increased ability to be confident and vulnerable as a leader&lt;br&gt;4. Nearly 100% of faculty attended basic programme. Overall reaction post-session was positive to extremely positive.</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Vey, et al. (2013)</td>
<td>Single-Canadian hospital</td>
<td>29</td>
<td>Residents</td>
<td>3 months</td>
<td>2.5 hour meeting every 3 months</td>
<td>Role-play, discussion, games</td>
<td>Principles of leadership, innovative, creative problem solving, and communication techniques</td>
<td>1. “nearly 100% numerical rating of 5”&lt;br&gt;2a. Participants reported increased ability to act in crisis situations&lt;br&gt;3a. Participants reported increased ability to be confident and vulnerable as a leader&lt;br&gt;4. Nearly 100% of faculty attended basic programme. Overall reaction post-session was positive to extremely positive.</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Bircher, et al. (2013)</td>
<td>Single US residency program (extension of year training)</td>
<td>NR</td>
<td>Residents (PGY trainees)</td>
<td>2 years</td>
<td>Unspecified number of participants</td>
<td>Didactic teaching, online learning environment, supervision, project work as individuals</td>
<td>Alignments of competencies, a systems and collaborative approach, effective learning strategies</td>
<td>1. Increased confidence in having difficult conversations&lt;br&gt;2a. Comments cited increased awareness and understanding of leadership principles&lt;br&gt;3a. Increased confidence in having difficult conversations&lt;br&gt;3a. Participants reported increased interest in leadership and management Assessment of skills learnt</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Kocher, et al. (2010)</td>
<td>Single US academic medical centre</td>
<td>30</td>
<td>Faculty members</td>
<td>2 months</td>
<td>2.5 hour meeting every 3 months</td>
<td>Didactic teaching, online learning environment, supervision, project work as individuals</td>
<td>Alignments of competencies, a systems and collaborative approach, effective learning strategies</td>
<td>1. Increased confidence in having difficult conversations&lt;br&gt;2a. Comments cited increased awareness and understanding of leadership principles&lt;br&gt;3a. Increased confidence in having difficult conversations&lt;br&gt;3a. Participants reported increased interest in leadership and management Assessment of skills learnt</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Rees, et al. (2011)</td>
<td>Single US hospital department</td>
<td>NR</td>
<td>Residents (Senior Residents, Emergency Medicine)</td>
<td>3 months</td>
<td>Programme outlining the leadership roles of residents</td>
<td>Didactic teaching, online learning environment, supervision, project work as individuals</td>
<td>Alignments of competencies, a systems and collaborative approach, effective learning strategies</td>
<td>1. Increased confidence in having difficult conversations&lt;br&gt;2a. Comments cited increased awareness and understanding of leadership principles&lt;br&gt;3a. Increased confidence in having difficult conversations&lt;br&gt;3a. Participants reported increased interest in leadership and management Assessment of skills learnt</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Cherry, et al. (2011)</td>
<td>Single US university</td>
<td>10</td>
<td>Junior Faculty</td>
<td>6 months</td>
<td>Didactic teaching, online learning environment, supervision, project work as individuals</td>
<td>Alignments of competencies, a systems and collaborative approach, effective learning strategies</td>
<td>1. Increased confidence in having difficult conversations&lt;br&gt;2a. Comments cited increased awareness and understanding of leadership principles&lt;br&gt;3a. Increased confidence in having difficult conversations&lt;br&gt;3a. Participants reported increased interest in leadership and management Assessment of skills learnt</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Johnson, et al. (2011)</td>
<td>Single US hospital department</td>
<td>NR</td>
<td>Residents (Senior medical residents)</td>
<td>3 months</td>
<td>2.5 hour meeting every 3 months</td>
<td>Didactic teaching, online learning environment, supervision, project work as individuals</td>
<td>Alignments of competencies, a systems and collaborative approach, effective learning strategies</td>
<td>1. Increased confidence in having difficult conversations&lt;br&gt;2a. Comments cited increased awareness and understanding of leadership principles&lt;br&gt;3a. Increased confidence in having difficult conversations&lt;br&gt;3a. Participants reported increased interest in leadership and management Assessment of skills learnt</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>O’Donnell, et al. (2011)</td>
<td>Single US hospital (residency program)</td>
<td>NR</td>
<td>Residents (PGY1)</td>
<td>4 weeks</td>
<td>2.5 hour meeting every 3 months</td>
<td>Didactic teaching, online learning environment, supervision, project work as individuals</td>
<td>Alignments of competencies, a systems and collaborative approach, effective learning strategies</td>
<td>1. Increased confidence in having difficult conversations&lt;br&gt;2a. Comments cited increased awareness and understanding of leadership principles&lt;br&gt;3a. Increased confidence in having difficult conversations&lt;br&gt;3a. Participants reported increased interest in leadership and management Assessment of skills learnt</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Wragg, et al. (2011)</td>
<td>Single US residency programme</td>
<td>NR</td>
<td>Residents</td>
<td>6 months</td>
<td>2.5 hour meeting every 3 months</td>
<td>Didactic teaching, online learning environment, supervision, project work as individuals</td>
<td>Alignments of competencies, a systems and collaborative approach, effective learning strategies</td>
<td>1. Increased confidence in having difficult conversations&lt;br&gt;2a. Comments cited increased awareness and understanding of leadership principles&lt;br&gt;3a. Increased confidence in having difficult conversations&lt;br&gt;3a. Participants reported increased interest in leadership and management Assessment of skills learnt</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

**Supplementary Table 1: Summary of included studies. NR= Not Reported; NOS=Not Otherwise Specified; n/a=not applicable. MERSQI=Medical Education Research Study Quality Instrument; JBI=Joanna Briggs Institute Critical Appraisal Checklist for Qualitative Research.**
References:


41 Kuo AK, Thyne SM, Chen HC, et al. An innovative residency program designed to develop leaders to improve the health of children. *Acad Med* 2010;85:653. doi:10.1097/ACM.0b013e3181e60f6


58 Saravo B, Netzel J, Kiesewetter J. The need for strong clinical leaders - Transformational and transactional leadership as a framework for resident leadership training. *PLoS One* 2017;12:e0183019. doi:https://dx.doi.org/10.1371/journal.pone.0183019


105 Ninan D, Patel D. Career and Leadership Education in Anesthesia Residency Training. *Cureus* 2018;10:e2546. doi:https://dx.doi.org/10.7759/cureus.2546


114 Cherry RA, Davis DC, Thordnyke L. Transforming culture through physician leadership development. *Physician Exec* 2010;36:38–44.


<table>
<thead>
<tr>
<th>MERSQI Component</th>
<th>Classification</th>
<th>All studies (117)</th>
<th>MERSQI&gt;12 (16)</th>
<th>JBI&gt;6 (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study Design</strong></td>
<td>Single Group Cross-Sectional or Post-programme only</td>
<td>54 (46%)</td>
<td>0 (0%)</td>
<td>8 (57%)</td>
</tr>
<tr>
<td></td>
<td>Single Group Pre and Post Programme</td>
<td>54 (46%)</td>
<td>9 (56%)</td>
<td>5 (36%)</td>
</tr>
<tr>
<td></td>
<td>Non-Randomised Two Group</td>
<td>8 (7%)</td>
<td>6 (38%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td></td>
<td>Randomised Controlled Trial</td>
<td>1 (1%)</td>
<td>1 (6%)</td>
<td>1 (7%)</td>
</tr>
<tr>
<td><strong>Institution #</strong></td>
<td>Single</td>
<td>81 (69%)</td>
<td>10 (63%)</td>
<td>9 (64%)</td>
</tr>
<tr>
<td></td>
<td>Double</td>
<td>1 (1%)</td>
<td>1 (6%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td></td>
<td>Multi</td>
<td>35 (30%)</td>
<td>5 (31%)</td>
<td>5 (36%)</td>
</tr>
<tr>
<td><strong>Response Rate</strong></td>
<td>&lt;50% or NR</td>
<td>44 (38%)</td>
<td>3 (19%)</td>
<td>1 (7%)</td>
</tr>
<tr>
<td></td>
<td>50-75%</td>
<td>23 (20%)</td>
<td>3 (19%)</td>
<td>6 (43%)</td>
</tr>
<tr>
<td></td>
<td>&gt;75%</td>
<td>48 (41%)</td>
<td>9 (56%)</td>
<td>7 (50%)</td>
</tr>
<tr>
<td><strong>Type of Data</strong></td>
<td>Self-reported</td>
<td>70 (60%)</td>
<td>2 (13%)</td>
<td>9 (64%)</td>
</tr>
<tr>
<td></td>
<td>Observed</td>
<td>47 (40%)</td>
<td>14 (88%)</td>
<td>5 (36%)</td>
</tr>
<tr>
<td><strong>Questionnaire Construct Validity</strong></td>
<td>Reported</td>
<td>9 (8%)</td>
<td>7 (44%)</td>
<td>2 (14%)</td>
</tr>
<tr>
<td></td>
<td>Not Reported</td>
<td>108 (92%)</td>
<td>9 (56%)</td>
<td>12 (86%)</td>
</tr>
<tr>
<td><strong>Questionnaire Content Validity</strong></td>
<td>Reported</td>
<td>45 (38%)</td>
<td>14 (88%)</td>
<td>5 (36%)</td>
</tr>
<tr>
<td></td>
<td>Not Reported</td>
<td>72 (62%)</td>
<td>2 (13%)</td>
<td>9 (64%)</td>
</tr>
<tr>
<td><strong>Relationships to Other Variables</strong></td>
<td>Reported</td>
<td>8 (7%)</td>
<td>5 (31%)</td>
<td>3 (21%)</td>
</tr>
<tr>
<td></td>
<td>Not Reported</td>
<td>108 (92%)</td>
<td>11 (69%)</td>
<td>11 (79%)</td>
</tr>
<tr>
<td><strong>Data Analysis Comprehensiveness</strong></td>
<td>Comprehensive</td>
<td>23 (20%)</td>
<td>14 (88%)</td>
<td>6 (43%)</td>
</tr>
<tr>
<td></td>
<td>Less Comprehensive</td>
<td>94 (80%)</td>
<td>2 (13%)</td>
<td>8 (57%)</td>
</tr>
<tr>
<td></td>
<td>Descriptive only</td>
<td>102 (87%)</td>
<td>7 (44%)</td>
<td>11 (79%)</td>
</tr>
<tr>
<td></td>
<td>Beyond Descriptive</td>
<td>15 (13%)</td>
<td>9 (56%)</td>
<td>3 (21%)</td>
</tr>
<tr>
<td><strong>Outcomes (Kirkpatrick Level)</strong></td>
<td>Level 1</td>
<td>80 (68%)</td>
<td>8 (50%)</td>
<td>14 (100%)</td>
</tr>
<tr>
<td></td>
<td>Level 2a</td>
<td>70 (60%)</td>
<td>7 (44%)</td>
<td>13 (93%)</td>
</tr>
<tr>
<td></td>
<td>Level 2b</td>
<td>79 (68%)</td>
<td>11 (69%)</td>
<td>11 (79%)</td>
</tr>
<tr>
<td></td>
<td>Level 3a</td>
<td>51 (44%)</td>
<td>7 (44%)</td>
<td>10 (71%)</td>
</tr>
<tr>
<td></td>
<td>Level 3b</td>
<td>54 (46%)</td>
<td>14 (88%)</td>
<td>7 (50%)</td>
</tr>
<tr>
<td></td>
<td>Level 4a</td>
<td>9 (8%)</td>
<td>1 (6%)</td>
<td>2 (14%)</td>
</tr>
<tr>
<td></td>
<td>Level 4b</td>
<td>26 (22%)</td>
<td>7 (44%)</td>
<td>4 (29%)</td>
</tr>
</tbody>
</table>

Supplementary Table 2: Study characteristics organised by MERSQI heading. Brackets in headings refer to original MERSQI items where headings have been adapted for clarity.
### Supplementary Table 3: Proportion of studies which met Joanna Briggs Institute (JBI) Critical Appraisal Items for Qualitative Studies. Descriptions are adapted from the JBI tool. Higher Reliability Studies scored 6 or more on the JBI tool.

<table>
<thead>
<tr>
<th>JBI Component</th>
<th>Description of component</th>
<th>High-reliability studies (n=14)</th>
<th>Mixed-Methods Studies (n=53)</th>
<th>Qualitative Studies (n=10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Philosophical Perspective</td>
<td>Congruity between the stated philosophical perspective and the research methodology</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Objectives</td>
<td>Congruity between the research methodology and the research question or objectives</td>
<td>39 (63%)</td>
<td>33 (62%)</td>
<td>6 (60%)</td>
</tr>
<tr>
<td>Methods</td>
<td>Congruity between the research methodology and the methods used to collect data</td>
<td>38 (61%)</td>
<td>32 (60%)</td>
<td>6 (60%)</td>
</tr>
<tr>
<td>Analysis</td>
<td>Congruity between the research methodology and the representation and analysis of data</td>
<td>18 (29%)</td>
<td>15 (28%)</td>
<td>3 (30%)</td>
</tr>
<tr>
<td>Interpretation</td>
<td>Congruity between the research methodology and the interpretation of results</td>
<td>17 (27%)</td>
<td>13 (25%)</td>
<td>4 (40%)</td>
</tr>
<tr>
<td>Researcher Location</td>
<td>Statement locating the researcher culturally or theoretically</td>
<td>10 (16%)</td>
<td>8 (15%)</td>
<td>2 (20%)</td>
</tr>
<tr>
<td>Researcher Influence</td>
<td>Influence of the researcher on the research addressed</td>
<td>13 (21%)</td>
<td>12 (23%)</td>
<td>1 (10%)</td>
</tr>
<tr>
<td>Participant Representation</td>
<td>Participants and their voices adequately represented</td>
<td>21 (34%)</td>
<td>16 (30%)</td>
<td>5 (50%)</td>
</tr>
<tr>
<td>Ethics</td>
<td>Evidence of ethical approval by an appropriate body</td>
<td>26 (42%)</td>
<td>23 (43%)</td>
<td>3 (30%)</td>
</tr>
<tr>
<td>Conclusions Supported</td>
<td>Conclusions drawn in the research report flow from the analysis or interpretation of the data</td>
<td>20 (32%)</td>
<td>16 (30%)</td>
<td>4 (40%)</td>
</tr>
<tr>
<td>No.</td>
<td>Searches</td>
<td>Results</td>
<td>Type</td>
<td>Actions</td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------------------------------------------------------</td>
<td>---------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>1</td>
<td>PHYSICIAN*2/</td>
<td>87338</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>2</td>
<td>PHYSICIAN EXECUTIVES/</td>
<td>4211</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>3</td>
<td>CONSULTANTS/</td>
<td>6648</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>4</td>
<td>INTERNSHIP/</td>
<td>49084</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>5</td>
<td>RESIDENCY/</td>
<td>49084</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>6</td>
<td>MEDICAL STAFF/</td>
<td>3359</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>7</td>
<td>(physician* or surgeon* or doctor* or intern* or intern* or resident* or registrar* or consultant* or &quot;house officer&quot;* or &quot;medical staff&quot;) lid.</td>
<td>230256</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>8</td>
<td>1 or 2 or 3 or 4 or 5 or 6 or 7</td>
<td>304027</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>9</td>
<td>LEADERSHIP</td>
<td>40117</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>10</td>
<td>PRACTICE MANAGEMENT/</td>
<td>1344</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>11</td>
<td>&quot;leader&quot;* lid.</td>
<td>21361</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>12</td>
<td>(practice and manager*) lid.</td>
<td>9878</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>13</td>
<td>9 or 10 or 11 or 12</td>
<td>59933</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>14</td>
<td>PROGRAM EVALUATION/</td>
<td>61451</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>15</td>
<td>PROGRAM DEVELOPMENT/</td>
<td>28428</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>16</td>
<td>CURRICULUM</td>
<td>73527</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>17</td>
<td>EDUCATION, MEDICAL, CONTINUING/</td>
<td>24544</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>18</td>
<td>EDUCATION, MEDICAL, GRADUATE/</td>
<td>28493</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>19</td>
<td>(teach* or train* or educator* or course* or program* or pathway* or curricu*l) lid.</td>
<td>793506</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>20</td>
<td>14 or 15 or 16 or 17 or 18 or 19</td>
<td>873227</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>21</td>
<td>8 and 13 and 20</td>
<td>1068</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>22</td>
<td>limit 21 to yr=&quot;2013 -Current&quot;</td>
<td>461</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>23</td>
<td>GENERAL PRACTITIONERS/</td>
<td>7485</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>24</td>
<td>FAMILY PHYSICIANS/</td>
<td>16222</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>25</td>
<td>MEDICAL STAFF.HOSPITALI</td>
<td>22748</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>26</td>
<td>(GP or GPs or &quot;general practitioner&quot;* or &quot;family practitioner&quot;)* lid.</td>
<td>22053</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>27</td>
<td>23 or 24 or 25 or 26</td>
<td>65255</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>28</td>
<td>13 and 20 and 27</td>
<td>101</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>29</td>
<td>limit 28 to yr=&quot;2013 -Current&quot;</td>
<td>63</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
<tr>
<td>30</td>
<td>29 not 22</td>
<td>40</td>
<td>Advanced</td>
<td>Display Results</td>
</tr>
</tbody>
</table>

**Supplementary Figure 1: Medline (OVID) Search Strategy, January 2020**