Mentoring as an intervention to promote gender equality in academic medicine: a systematic review

Allan House,1 Naila Dracup,2 Paula Burkinshaw,3 Vicky Ward,4 Louise D Bryant5

ABSTRACT

Background Mentoring is frequently suggested as an intervention to address gender inequalities in the workplace.

Objectives To systematically review evidence published since 2006 on the effectiveness of mentoring interventions aimed at achieving gender equality in academic medicine.

Design Systematic Review, using the Template for Intervention Description and Replication as a template for data extraction and synthesis.

Sample Studies were included if they described a specific mentoring intervention in a medical school or analogous academic healthcare organisation and included results from an evaluation of the intervention.

Eligibility criteria Mentoring was defined as (1) a formally organised intervention entailing a supportive relationship between a mentor, defined as a more senior/experienced person and a mentee defined as a more junior/inexperienced person; (2) mentoring intervention involved academic career support (3) the mentoring relationship was outside line management or supervision of performance and was defined by contact over an extended period of time.

Outcomes The impact of mentoring was usually reported at the level of individual participants, for example, satisfaction and well-being or self-reported career progression. We sought evidence of impact on gender equality via reports of organisation-level effectiveness, of promotion or retention, pay and academic performance of female staff.

Results We identified 32 publications: 8 review articles, 20 primary observational studies and 4 randomised controlled trials. A further 19 discussed mentoring in relation to gender but did not meet our eligibility criteria. The terminology used, and the structures and processes reported as constituting mentoring, varied greatly. We identified that mentoring is popular with many who receive it; however, we found no robust evidence of effectiveness in reducing gender inequalities. Primary research used weak evaluation designs.

Conclusions Mentoring is a complex intervention. Future evaluations should adopt standardised approaches used in applied health research to the design and evaluation of effectiveness and cost-effectiveness.

INTRODUCTION

The total number of clinical academic staff in UK medical schools has reduced in recent years with a 2.5% decline since 2010.1 The most significant declines are in numbers of clinical research academics at reader/senior clinical lecturer grades and at the stage of career progression from PhD to independent postdoctoral research. A recent survey of clinical research fellows identified a decline in the number of awards at these key career transition points.2 Around 25% of the clinical academic population is aged over 55: lack of capacity further back along the pipeline is a concern for the future of clinician-led research and teaching in the UK. Against this background there are specific considerations around the representation of women. While the number of female clinical academics has risen in comparison to men, the proportion of women decreases by seniority. In medicine, around 19% of clinical professors and 37% of senior clinical lecturers are women with proportions increasing at a rate of around 1%–2% a year.13

There is already a substantial literature related to concerns about recruitment and retention in clinical academic careers and the gender imbalance in senior roles. A number of intersecting factors have been identified in this literature: family responsibilities (caring, mobility), competition for research grants,
‘work–life balance’, competing work priorities, pressures within the National Health Service, limited availability of senior academic roles and financial considerations. However, the evidence is inconclusive on which the most important factors are, for whom, at what career point—and what effective strategies to address them would look like. One of the most significant recent gender balance interventions in the UK—the Athena SWAN charter scheme—as yet has failed to demonstrate strong evidence of changing the representation of women in senior roles.

Mentoring is a popular intervention that is widely endorsed in many walks of life, but particularly in relation to the workplace. It has been recommended as an intervention to help professionals develop their career, to achieve a balanced portfolio of work and non-work activities, and to maintain well-being. Despite being an individual level intervention, it has also been suggested that mentoring is one means to increase retention and involvement of under-represented groups and to address inequalities in the workplace, especially those of gender and ethnicity. For these reasons, support for mentoring schemes has proved popular as a part of organisational responses to the challenge of addressing gender inequalities in academic life. Given that many individuals report having found mentoring helpful, a question arises about what evidence there is for organised mentoring schemes, in terms of their effectiveness in promoting gender equality in academic medicine.

Earlier reviews have concluded that lack of evidence on the effectiveness of mentoring schemes is a result of two deficits in the literature. First is a lack of well-designed evaluations; for example when Sambunjak et al undertook their review just over a decade ago, the primary reason that they were unable to find evidence for the effectiveness of mentoring was that they were unable to identify any randomised trials or more sophisticated quasi-experimental designs such as interrupted time series analysis. Second, a major impediment is substantial variability and lack of specificity in the stated aims and objectives of different approaches to mentoring, in the structures and processes subsumed under the rubric of mentoring and in the outcomes expected or actually measured in mentoring schemes.

Since the exhaustive systematic review of Sambunjak et al found no evidence to support the effectiveness of mentoring interventions in reducing gender inequalities, the present systematic review aimed to identify evidence published since that date, with a particular focus on answering the question—is there new evidence (published since 2006) about whether organised mentoring schemes reduce gender inequalities in academic medicine? In undertaking this review, we did not simply replicate the methods used by Sambunjak et al, but took the opportunity to ask a subsidiary question about whether it is feasible to consider mentoring as analogous to so-called complex interventions in healthcare. Complex interventions are usually described as interventions that contain several interacting components. There are, however, several dimensions of complexity: it may be to do with the range of possible outcomes or their variability in the target population rather than with the number of elements in the intervention package itself. For this reason, we organised the data using the Template for Intervention Description and Replication (TIDieR) framework, which is a widely adopted approach to standardising the description of complex interventions in healthcare, enabling the most effective components of the intervention to be identified.

METHODS

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guidelines have been used. We did not register a protocol for the review prior to commencement. We sought published articles, either single case studies or reviews, which reported mentoring schemes delivered in institutions that employed academic medical staff.

There is no standardised definition of what constitutes a systematic review: here, we use the six criteria suggested by Krnic Martinic et al as below.

Research question

Is there evidence, published since a previous substantial review published in 2006, about whether organised mentoring schemes reduce gender inequalities in academic medicine?

Search strategy

Searches were developed by an Information Specialist for the concepts: mentoring, coaching or sponsorship and medical schools or university departments. Subject headings and free-text words were identified for use in the search concepts by the text analysis tools PubReminer and Yale MeSH Analyser, the Information Specialist and project team members. Further terms were identified and tested from known relevant papers. The search was peer-reviewed by an additional Information Specialist.

In April 2017 we searched:

- Cochrane Database of Systematic Reviews: Issue 4 of 12 April 2017.
- Cochrane Central Register of Controlled Trials: Issue 3 of 12 March 2017.
- Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations.
- Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to April 2017.
- PsycINFO 1806 to April week 1 2017 BASE (Bielefeld Academy Search Engine).
- Equality Challenge Unit.
European Commission-Research and Innovation SWAFS (Science With and For Society), European Mentoring and Coaching Council.
National Research Mentoring Network.
NICE Evidence Search.
Birkbeck University Transforming Institutions by Gendering Contents and Gaining Equality in Research.

See online supplemental appendix I for full (Medline) search strategies.

The results of the database searches were stored and deduplicated in an EndNote library. Further relevant studies were sought by citation searching (forwards and backwards) of the included studies and reviews. The search was repeated in December 2019, and citations were also sought to studies identified and included from the first search.

Eligibility criteria
- Published English language studies that provided detail of a specific mentoring scheme, where mentoring was defined as (1) an intervention that entailed promoting and supporting a relationship between a mentor (also sometimes described as coach or sponsor) defined as a more senior/experienced person and a mentee (sometimes called a protégé) defined as a more junior/inexperienced colleague who was a member of academic staff and not a student; (2) the content of the intervention involved a mixture of practical guidance and career development rather than technical skills training and (3) the mentoring relationship was outside formal line management or supervision of performance and was defined by scheduled contact over an extended period of time.
- The mentoring scheme was organised by a medical school or analogous academic healthcare institution.
- Publication date was 2006 or later, that is too recently for inclusion in the review of Sambunjak et al.

Exclusion criteria
- Other interventions designed to reduce gender inequalities or support women in academic life, such as financial interventions to sustain research over a period of maternity leave.
- Schemes that used the term mentoring to describe a relationship other than defined above, such as peer-mentoring or upward mentoring where the mentors are more junior than the mentees.
- Schemes designed to support career development through more formal means such as accredited training or supervision programmes sometimes described as preceptorships, academic review and feedback, one-off or occasional meetings.
- Testimony and personal accounts of experience; reports involving small numbers (25 or fewer participants); reports of schemes where no formal evaluation was undertaken.
- Schemes designed to support career development rather than technical skills training and we have not listed here the few that do. Most schemes used the name of the organisation linked to the term mentoring as the identifier for their programme.

Survey of academics’ self-reported evaluation of their own experience, unless related to a specific and described scheme.

Study selection
Studies for inclusion were agreed after review of titles and abstracts, and full texts when necessary. To ensure consistency two authors (AH and ND) reviewed 20% of titles and abstracts and additional titles where they were in doubt, to discuss decisions about application of eligibility criteria.

Study quality
Because of the wide range of study designs in the studies we identified, we were not able to apply a single standard assessment of quality. We, therefore, noted study design, as did Sambunjak et al, as the single most important indication of likely biases in evaluation research. We did not exclude any studies based solely on research design, if they met the eligibility criteria noted above.

Data extraction and synthesis
Data were extracted to a standard form designed using the TIDieR framework as a template (see table 1 for a summary of the framework). We then used this framework as the basis for a narrative synthesis of the results from identified studies.

Patient and public involvement
No patient involvement.

RESULTS
The database searches identified 5855 records and from other sources we identified a further 182 records. Once duplicates were removed there were 4240 records. Citation searches identified 220 further records. After screening titles and abstracts, we reviewed 146 papers. A final 24 primary publications were identified for the review as well as eight review articles: we found 20 primary observational studies which were either cross-sectional or before-and-after studies with no study using interrupted time series or similar analysis, and we identified four randomised controlled trials (RCTs). We also identified 19 publications discussing mentoring in relation to gender but not meeting our eligibility criteria.

Characteristics of mentoring schemes
Naming (TIDiER item 1)
With occasional exceptions, it is not usual to develop a new name for a mentoring scheme and we have not listed here the few that do. Most schemes used the name of the organisation linked to the term mentoring as the identifier for their programme.

Rationale, theory, aims and objectives of mentoring schemes (TIDiER item 2)
Few studies referred explicitly to a theory that lay behind their choice of mentoring scheme. Those that did named...
a general theory such as self-determination theory\textsuperscript{49} or a
general theory of mentoring.\textsuperscript{54} Implicit in many schemes,
although not stated as a theory, were ideas related to self-
determination; that mentoring might help people meet
their aspirations by supporting them to meet needs in
areas such as competence, autonomy or relatedness.\textsuperscript{55}
We found one report that provided a logic model for the
programme.\textsuperscript{47}

Mentoring specifically for female academics in medi-
cine is frequently regarded (either explicitly or implicitly)
as an intervention aimed at reducing gender inequalities
in career development, but we found no publications that
linked mentoring to theories about the origins of such
inequality. For the purposes of this review, we; therefore,
treated mentoring specifically aimed at women as an
example of tailoring of a more generic intervention and
it is covered later in this article. Box 1 summarises the
aims and objectives of the schemes reviewed.

Four main purposes of mentoring were identified:
1. Increasing knowledge: including provision of advice
about how promotions processes worked in the organ-
isation, or developing and presenting a curriculum
vitae (CV). For some, mentoring was seen as a way to
increase the knowledge of mentees in less immediately
applied ways, such as helping women to understand
how the gendered nature of academic life works, shar-
ing practical knowledge about University systems, and
allowing what has been called ‘a redistribution of fem-
inist knowledge and social capital’.\textsuperscript{56}
2. Improving skills: generic academic skills included
grant writing or lecturing for example. Sometimes
women were seen to need additional skills, for exam-
ple, mentoring was linked to help with negotiating
salary rises, self-presentation at interviews or meetings,
and balancing career and parenthood. Some skills
were so broadly formulated that it was difficult to know
what was imparted in mentoring: examples included
skills appropriate for advancement\textsuperscript{42} skills to enhance
capability and productivity\textsuperscript{57 58} and improving academ-
ic performance.\textsuperscript{59}
3. Developing self-confidence: female academics are
widely represented in the mentoring literature as lack-
ing assertiveness, self-confidence or self-efficacy, neg-
atively impacting on career development. Mentoring

Table 1  Items included in the TIDieR checklist: information to include when describing an intervention

| Brief name | 1 Provide the name or a phrase that describes the intervention. |
| Why | 2 Describe any rationale, theory or goal of the elements essential to the intervention. |
| What | 3 Materials: Describe any physical or informational materials used in the intervention, including those provided to participants or used in intervention delivery or in training of intervention providers. Provide information on where the materials can be accessed (such as online supplemental appendix, URL). |
| | 4 Procedures: Describe each of the procedures, activities and/or processes used in the intervention, including any enabling or support activities. |
| Who provided | 5 For each category of intervention provider (such as psychologist, nursing assistant), describe their expertise, background and any specific training given. |
| How | 6 Describe the modes of delivery (such as face to face or by some other mechanism, such as internet or telephone) of the intervention and whether it was provided individually or in a group. |
| Where | 7 Describe the type(s) of location(s) where the intervention occurred, including any necessary infrastructure or relevant features. |
| When and how much | 8 Describe the number of times the intervention was delivered and over what period of time including the number of sessions, their schedule and their duration, intensity or dose. |
| Tailoring | 9 If the intervention was planned to be personalised, titrated or adapted, then describe what, why, when and how. |
| Modifications | 10 If the intervention was modified during the course of the study, describe the changes (what, why, when and how). |
| How well | 11 Planned: If intervention adherence or fidelity was assessed, describe how and by whom, and if any strategies were used to maintain or improve fidelity, describe them. |
| | 12 Actual: If intervention adherence or fidelity was assessed, describe the extent to which the intervention was delivered as planned. |

TIDieR, Template for Intervention Description and Replication.
Box 1 Aims and objectives of mentoring schemes

Mentoring is designed to change

Knowledge
► Knowledge of promotion criteria and requirements for promotion.33 53 65
► Awareness of academic responsibilities for example, teaching, research, student supervision and practice.45
► Of self-determination theory.38

Skills
► Leadership.33 57 66
► Decision making.53
► Goal setting: clinical, research and teaching.33 66
► Skills at applying for promotions.53
► Negotiating.66
► Effective communication.50
► Professional networking.15 34 66
► Expectation management.50
► Teaching skills.71
► Entrepreneurial skills for promoting research.72
► Overseeing departments for Chairs.67

Attitudes, psychological characteristics
► Job-related well-being, self-esteem, confidence and work–life balance.32
► Self-esteem.63
► Self-efficacy.63
► Job-related well-being and satisfaction.63
► Job interference with family/work–life balance.48 63
► Confidence.50 66
► Establishing trust.43
► Psychological need satisfaction.49
► Fostering independence.90

Connectedness
► Increasing connectedness between junior and senior faculty.34 42
► Encourage networking across institutions via cross-institutional mentoring scheme.53
► Foster links with under-represented minority communities.30
► Improve organisational transparency and communication.33

Retention of identified staff groups:
► Retention of all staff.15 42 68
► Retention of department chairs.67
► Under-represented minority faculty.13 30 52 53 58 60 66
► Staff from non-traditional disciplinary backgrounds.52 57

Materials and procedures (TIDIER items 3+4)
Materials to support mentors usually took the form of briefing documents, for example, a paper outlining the goals and structure of the scheme.33 Other schemes sign-posted mentors to guidance books and manuals,15 online resources about mentoring,34 or created a mentoring resource section in the University library with relevant books and articles.60

Written agreements were also used; here mentors confirmed their commitment to programme coordination or faculty development60 or were required to sign an agreement with their mentee in their first or second meeting, setting out frequency and duration of meetings and aims and expectations.52 Such written partnership agreements, sometimes coupled with a requirement for progress reports, were designed to hold both the mentor and mentee accountable.61

Materials to support or brief mentees took the same two forms. Written materials included information about the scheme and guidance on how to use mentoring as well as more generally available literature on mentoring.23 60 Written agreements included signed formal programmes outlining key concepts of the programme such as confidentiality, active listing and a commitment to meet every 3 months,65 agreements to set boundaries and encourage accountability15 and a mentoring agreement certificate outlining areas of mentoring focus and frequency of meetings.66. In one scheme, each pair was expected to write short minutes of each meeting and submit to the programme coordinators, encouraging commitment to the scheme.52

Who provided the intervention (TIDIER item 5)
Definition of mentors
In most schemes, mentors were defined by seniority in the organisation, as senior faculty, professors, experienced researchers or department heads33 52 53 57 62 66 67 69 72 with several reports noting that this led to a preponderance of male mentors. Sometimes an additional expectation was that they already had some mentoring experience50 or had what were deemed to be qualities of a good mentor.61 64 For schemes with a particular focus, mentors might be defined according to another characteristic, for example, being female60 or expressing support for faculty members from minority ethnic groups.53

Selecting mentors
We identified three main ways in which mentors were selected:

1. Prospective mentors were invited to propose themselves or they could be nominated by peers or potential mentees. No other selection criteria were applied.25 52 61 63 65
2. Staff were invited to apply and then a decision was made by the scheme organiser, on grounds not usually specified beyond suitability for the scheme.16 34 49 66
3. Mentors were selected by committee or scheme coordinators, based on track record in mentoring or in research.38 50 52 53 57 59 62 66 68

and role modelling were considered to improve this situation but the function was not always specified beyond general descriptions such as reference to personal change to aid professional and personal development.15 48 50

4. Building relationships and networks: for some, the power of mentoring comes when it is linked to career sponsorship, that is, help in making contacts, building strategic alliances, networking and obtaining national and international recognition. Since there is a lack of women in senior leadership positions, this sponsorship is seen as being done mainly by and for men, which compounds existing gender inequalities in academic career development.
**Support and training for mentors**

In some schemes, training was provided in mentoring skills, on the aims of the specific mentoring scheme being introduced or into topics likely to come up during mentoring.\(^{16, 34, 38, 53, 60, 61, 65, 66, 69}\) Such training was mainly done just with mentors, but occasionally with joint workshops so that mentees and mentors get to know one another.\(^{16, 61}\) Additional support was offered by arranging peer support meetings\(^{33, 66, 68}\) or offering supervision to mentors by a member of team organising scheme\(^{38, 52, 63, 68}\).

Some schemes offered incentives to mentors. Financial incentives included compensation through a ‘small’ salary offset,\(^{42, 68}\) an incentive payment for each mentor who completed training,\(^{49}\) funding to gain additional time and release from teaching duties\(^{62}\) and payment of expenses.\(^{52}\) Non-financial incentives included accrediting service points which counted towards annual review,\(^{65}\) including mentoring in performance reviews, and considering mentoring when reviewing applications for promotion.\(^{69}\) Peer-reviewed nominations and awards at department and institutional level were used to encourage participation and good practice.\(^{32, 69}\)

**How mentoring was delivered (TIDIER item 6)**

**Selecting mentees and matching to mentors**

Definition of mentees varied. Some schemes were targeted specifically at certain healthcare disciplines or those with particular sources of funding.\(^{38, 57}\) Others defined mentees according to career stage, for example, assistant professors,\(^{33}\) staff who were full time salaried for 1–5 years in Assistant Professor or teaching positions,\(^{30}\) Academic Clinical Lecturers and Clinician Scientist Fellows.\(^{34}\) Schemes that took this approach were sometimes willing to offer mentoring to new recruits at any career stage.\(^{65, 67, 69}\) Some schemes made no specification and were open to all staff, men and women, but predominantly catered for early career researchers, teaching staff and lecturers.\(^{16, 62}\)

Certain specialised mentoring schemes were open only to mentees from defined groups—for example under-represented minorities\(^{49, 53}\) or women (see later).

**Selection of mentees**

In this review, we included only those reports where it was clear that the staff referred to (using whatever terminology the authors employed) were in academic rather than managerial or administrative positions. Most schemes accepted all eligible staff, and all were invited to apply. One scheme enrolled eligible staff automatically.\(^{66}\) In other cases, staff were sent an invitation to apply stating their reasons for wanting mentoring, followed by a selection procedure.\(^{16, 52, 53}\) Occasionally mentees were selected by heads of faculty or scheme coordinators as likely to benefit from mentoring.\(^{59, 63}\) As with mentors, training about the nature of mentoring was provided by some Institutions\(^{38, 49, 53, 59, 64, 65}\) and support was offered through supervision of the progress of the mentoring relationship\(^{52}\) and financial or other incentives.\(^{49, 60, 62}\)

**Matching mentors and mentees**

Although multiple mentors per mentee have been recommended so that different needs can be met by different mentors,\(^{32, 61, 64}\) (network mentoring), in reality most schemes offer a 1:1 pairing of mentor and mentee. While informal mentoring is usually arranged by a mentee approaching a potential mentor personally, in organised schemes some degree of senior involvement is expected. Occasionally the process of matching was managed by randomisation\(^{48, 49}\) but usually there was purposeful matching. Many schemes tried to allow an element of personal choice by the mentee, with the degree of interactivity in this process varying considerably:

- Mentees matched with their nominated mentor from a selection of ‘pen portraits’ of available mentors; or were matched in a ‘clinic’ held by the scheme organiser.\(^{63}\)
- Mentees were encouraged to select a potential mentor based on a match of clinical, research or administrative interests or experience.\(^{33}\)
- Mentees were encouraged to select several potential mentors and to meet potential mentors before making a decision to establish ‘good fit’.\(^{69}\)
- Junior faculty were asked to fill out a short survey assessing areas of expertise and interests and to name a potential mentor—matching was then made by scheme organisers.\(^{60}\)
- All prospective mentees completed a needs assessment and a discussion held with department chair who selected a mentor based on availability and suitability.\(^{65}\)
- A booklet outlined details of the mentors including research interests and what they were willing to contribute. Mentees and mentors were introduced and the programme coordinator then matched the pairs ‘in a way that was acceptable’ to the pairs.\(^{62}\)
- Matching mentor and mentee involved a questionnaire and interviews. Mentees discuss with the scheme coordinator what they want and the coordinator helps them to select.\(^{52}\)
- Mentees had an interview with the mentoring coordinator to determine their needs.\(^{59}\)
- Mentees nominated three ‘candidate’ mentors from faculty. They were provisionally matched with one chosen mentor followed by an initial conversation to see if they feel able to work together.\(^{57}\)

Where matching was based on something other than personal compatibility and preference, it was usually based on specialty, academic department or faculty,\(^{15}\) or research interests.\(^{38, 64}\) Two schemes noted the difficulty of matching by ethnicity.\(^{60, 61}\)

**Form and content of sessions**

Once a mentoring relationship was established, its form and content was not usually specified, although occasionally it was highly organised.\(^{43}\) For example, some mentees were encouraged to take the lead in identifying goals and aspirations,\(^{57}\) alternatively, task-specific...
mentoring was discouraged and mentors were advised to act as a ‘life coach’, providing broad career support and direction.68 Mentees and mentors could decide the frequency, length and content of their meetings without a prescribed duration period.38 59 In one scheme, the meetings took place predominantly via digital media and email.70

We identified four themes in the content of mentoring sessions.

1. Task-specific mentoring.
2. Career planning, CV development and promotion planning.
3. Developing personal and interpersonal skills.
4. Emotional support and wider advice.

Füger52 recommends that mentoring sessions are task specific and for many, mentoring appears to have that focus.48 62 In some cases suggested topics for meetings were very detailed and the mentee would attend meetings prepared to present detailed accounts of teaching,71 be helped to analyse data and prepare manuscripts and presentations, or take grant applications for review.42 57 61 Most task-specific sessions focused on teaching and research68 including writing for grant applications and for publication. These may be could be supplemented by workshops30 48 and further observation and feedback.65 71 In these situations, mentoring blurs into supervision.

Advice on career planning30 57 might again be supplemented by workshops34 65 with a more-or-less formal curriculum covering topics like balancing research, teaching or clinical practice,15 goal setting30 68 developing a research portfolio38 building a sustainable research career,57 preparing an academic CV66 or understanding the organisation and committee participation.66

Developing personal and interpersonal skills49 included advice on negotiation skills and conflict resolution,68 leadership styles, performance evaluation and presentation skills,60 decision making, problem solving, providing feedback and role modelling.66 Understanding the importance of, knowing how to maximise and facilitate networking opportunities were widely supported34 52 60 66 68

Emotional support and wider advice feature in some accounts of mentoring, usually in schemes without a strong task-orientation. Giving moral support to help a mentee cope with the stresses of academic life, and helping build motivation48 61 were seen as important especially when those stresses may be related to discrimination in the workplace (see later section on gender). Assisting with difficulties such as student complaints65 balancing professional and personal demands65 and confidence building.66

Where, when and how much (item 7+8) We found no studies that specified a priori how mentoring meetings should be arranged and delivered. However, several reports (see the evaluation section) recorded what had happened in practice.

Tailoring (item 9) Tackling gender inequality was a rationale mentioned in some reports, although few were explicit about how mentoring was to be modified or tailored to meet that aim.16 48 52 59 60 63 72 73 The most common approach in practice was to predefine the gender of either mentee or mentor as one of the requirements for participation. Several schemes reported that mentoring was open to women only as mentees.52 59 60 63 The focus included assisting female faculty with networking and making connections and offering targeted support in grant writing and related academic activities.

We found no study that examined the impact of mentoring in those populations defined by what is usually called intersectionality—that is, schemes where the focus might be, for example, on female staff from minority ethnic backgrounds, or living with a disability.

There was no consensus on gender-matching, even in mentoring schemes for female mentees only. Although schemes reported that they tried to match by sex67 and that such matching was desirable52 others noted that matching should be based on mentees’ needs16 38 or that for practical reasons mentors could be male and female to ensure that they had adequate numbers of mentoring staff at senior levels.63 One scheme explicitly did not match on gender after an internal survey revealed that 100% of males and 87% of females either disagreed with or were neutral about the suggestion that same sex dyads were ‘easier’.68

Two schemes offered training on ‘gender in academic life’ as part of mentoring.52 68 The content of such sessions was not well specified, for example ‘interpersonal competencies (gender and generational issues, negotiation skills, conflict resolution, individual preferences using the Myers-Briggs Type Indicator and effective coaching)’ and seemed based on helping mentees to recognise common barriers for women rather than acquisition of specific skills in dealing with them.

Formal evaluation of schemes (items 11+12) Adherence/fidelity/quality Mentoring pairs met typically 2–5 times within a year, but we found a wide range from once annually, to weekly. Most mentoring conversations lasted between 1 and 2 hours.15 33 38 53 57 63 66 68

Implementation problems were identified in a few schemes, including lack of organisational recognition of the addition to workload of mentors, a lack of mentor interest, incompatibility in mentoring pairs25 34 a lack of mentor experience64 perceived lack of commitment from the institution52 and practical problems such as difficulty scheduling meetings67 and a lack of protected time for meetings.15 Such problems were typically not monitored systematically.

Effectiveness We found no study that compared outcomes of mentoring in men and women.
The most common approach to evaluation was to undertake questionnaire surveys of mentors or mentees. One scheme interviewed staff as they left the scheme and one analysed transcripts of feedback meetings.

Two groups of outcomes were based on the personal experiences of participants. Most commonly, we found evaluations of participant experience of academic life and the degree to which they thought it was changed positively by mentoring. Outcomes included self-reported experience of job-related well-being, overall satisfaction, job satisfaction, job-related anxiety versus contentment, increased confidence and self-efficacy, development of networking skills, better time management and better work–life balance. Second, a range of self-reported outcomes in broad terms were evaluated, for example, improved clinical leadership skills, enhanced national or international profile, feeling welcomed as a member of the scientific or clinical community, following one’s chosen career path, and remaining in academic life.

Many schemes reported improvement in these two areas of experience.

Negative outcomes are by comparison were far less widely reported. The most common was that the mentoring relationship ‘never got off the ground’ or that the mentor was incompetent or distant and uninterested. More serious negative outcomes include an experience of being harassed or bullied, manipulated by a mentor who plagiarised or otherwise used a mentee’s work, or some other breach of trust or of a confidence.

Two problems make it difficult to interpret the effectiveness findings. Lost to follow-up will have led to substantial response bias, and lack of a comparison group prevents attribution to the mentoring experience.

More objective outcomes that indicated academic productivity were publications and grant income. These outcomes could be recorded individually or in aggregate. Because mentoring is typically offered to people at a career stage where increasing academic activity is to be expected, the lack of a comparison group in these studies prevents attribution of productivity to the intervention.

Organisational benefits as outcomes included improved retention and promotion of academic staff and some evidence of improved retention and promotion for women in before-and-after studies. Mentoring may orient staff towards policies supporting diversity and gender equality, and the few studies that sought evidence of organisational benefits recognised the importance of helping mentees gain an understanding of organisational structures and policies.

We identified four RCTs of aspects of mentoring schemes; all other studies used either cross-sectional or simple before-and-after designs. Three RCTs tested the effect of training mentors to participate in a scheme compared with no training. One found increased mentoring competence after training but did not measure mentee outcomes. Two found improvements in experience of mentee’s career but no difference between groups and no gender effect.

One RCT examined a broad-based intervention for female mentees that included an element of coaching and mentoring in a cluster randomised trial comparing outcomes in departments where the scheme was introduced to outcomes in departments with no scheme (table 2). Results were generally modest: participants with a PhD but not those with an MD qualification increased their publications more if they were in an intervention department despite reporting that they had reduced their hours of work; both intervention and control departments showed increases in publication rates and perceived self-efficacy. The authors concluded that the scheme might have helped staff to ‘work smarter’, an effect that did not benefit clinical staff to the same extent. The effectiveness of the intervention might have

<table>
<thead>
<tr>
<th>Brief name</th>
<th>NIH-TAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why Rationale, theory, goals</td>
<td>1. To improve key indicators of academic success for mentees. 2. To drive broader changes in culture.</td>
</tr>
<tr>
<td>What Materials, training,</td>
<td>‘Coaching’ and ‘additional mentoring’ were embedded in a wider programme of professional development seminars and strategically planned interventions in departmental work practices.</td>
</tr>
<tr>
<td>Who Definition and selection of mentors</td>
<td>Mentors/coaches not defined or described.</td>
</tr>
<tr>
<td>Definition and selection of mentees</td>
<td>Target population was female assistant professors in intervention departments. 134/178 (75%) participated.</td>
</tr>
<tr>
<td>Matching</td>
<td></td>
</tr>
<tr>
<td>How Practicalities of mentoring meetings</td>
<td>Not described</td>
</tr>
<tr>
<td>Where when and how much</td>
<td>3 years programme. Formal seminars in department.</td>
</tr>
<tr>
<td>Tailoring</td>
<td>Only female staff targeted, no gender-matching of mentors/coaches</td>
</tr>
<tr>
<td>Modifications</td>
<td>None described</td>
</tr>
<tr>
<td>How well Adherence and fidelity</td>
<td>Participation in the formal programmes (Manuscript Writing, Total Leadership) was recorded, but not experience of additional coaching or mentoring.</td>
</tr>
</tbody>
</table>

NIH-TAC, National Institutes for Health-Transforming Academic Culture Trial; TIDieR, Template for Intervention Description and Replication.
been masked by substantial organisation-wide changes taking place at the same time as the trial. The coaching/mentoring component of the intervention was not well specified, especially in relation to any gender-specific component.

Two schemes estimated costs attributed to the mentoring scheme and costs saved by organisational benefits of the scheme. They concluded that mentoring pays for itself by increasing academic staff retention and thereby reducing recruitment costs and by increasing external research funding.

In summary, we identified a substantial literature on mentoring in academic medicine, and numerous references to the problem of inequalities in academic careers. We were able to characterise mentoring as a complex intervention using the TIDieR framework but were unable to synthesise findings from the included studies because of lack of a standardised approach to defining mentoring, the structures and processes involved in its delivery or the outcomes. The literature on mentoring schemes is still dominated by descriptions of evaluations that use weak research designs, with few randomised trials or alternative rigorous quantitative methods.

DISCUSSION
This study provides an up-to-date systematic review, according to PRISMA guidelines, of published evidence of the effectiveness of mentoring schemes in achieving equality for female academics in medicine. The extraction and synthesis of data using a widely adopted framework for reporting complex interventions in healthcare (TIDieR) enabled us to ‘unpack’ the broad construct of mentoring into its component parts, in the same way the framework has been used in other settings.

There is no reason to doubt that, for early career academics, a supportive advisory relationship with an experienced and interested colleague can be helpful to career development and to morale. Indeed, it would be odd if it were not so, and no doubt this accounts for some of the promotion of mentoring schemes despite lack of evidence of effectiveness. However, we found few randomised controlled trials evaluating organised mentoring schemes. Alternative designs can be used in applied health research when randomised trials are not feasible, for example, interrupted time series or modelling based on routine data, but we identified no such studies. As a result, we found no study that provided definitive evidence to support the introduction of mentoring schemes organised at an institutional level, rather than informally between individuals, as a way of tackling gender inequalities in academic medicine.

It is a limitation of our review that we included only English language publications. This limits what we can say about the impact of mentoring in countries which may have very different cultures in relation to gender equality, as well as different approaches to medical training and career support. To take one example, we know that the problem of gender inequality in medicine is substantial in Japan, where fewer than 10% of medical school professors are women. There, efforts at change appear to have had limited impact but a full evaluation is not available in an English language academic publication. Nonetheless we believe that our findings can be applied to many countries in Europe, to North America and to other countries with similar educational and cultural traditions.

Drawing conclusions about the effectiveness of mentoring interventions remains difficult; however, there are serious weaknesses in the evaluation literature, and this is a limitation of the review. The structures and processes of mentoring were often not specified in enough detail that they could be properly evaluated. Even in longitudinal studies, the outcomes were rarely specified a priori or linked to structures and processes in a logic model or programme theory. When rationale was specified and standardised outcomes used, most of the evidence for the popularity and individual effectiveness of mentoring in advancing academic career came from cross-sectional studies. Where mentoring was aimed at supporting female academics there was little description of what constituted gender-specific mentoring, reported outcomes were not gender-specific and there was no between-group (and specifically between-gender) comparisons—limiting an answer to our research question. Terminology used to describe mentoring was inconsistent, further limiting the review process.

The lack of randomised trials or other higher-quality designs in the literature means that indication bias is a major possibility as an explanation for reported benefits; that is, it is the most energetic and ambitious academics or those most enthusiastic about having personal support in their careers who opt for mentoring schemes and then report good career and personal outcomes from the process. A first step would be to use a descriptive approach that allows mentoring schemes to be developed and evaluated in consistent and replicable ways, with clear outcomes related logically to the implementation of the scheme. We present the TIDieR framework here and show that it can be used to characterise mentoring schemes in just this way. If mentoring is to be used to tackle gender inequalities then the framework used to plan and evaluate the scheme must pay more careful attention to the form and content of the programme in relation to desired and prespecified gender-specific outcomes. Understanding which components of this complex intervention are most effective may help target resources more effectively.

Not everybody agrees that such an enterprise is worthwhile. The suggestion that mentoring can be useful for tackling inequality issues by providing women with an additional resource to work on their continuing professional development is widely endorsed. However, mentoring schemes must be careful not to further reinforce gender power relations and exclusionary structures and practices, for example by trying to ‘fix the women’ in terms of their confidence, for example. We need to remain open to the possibility of questioning the validity...
of mentoring as an individualised response to problems that may have more collective or structural solutions. 80-82

CONCLUSIONS
Establishing the place of mentoring schemes in medical schools is not a trivial matter. If it is effective in reducing gender inequalities (or indeed improving career development) then it needs to be universally implemented. If it is ineffective then we need to know because there are significant costs in academic time, opportunity costs in diverting energies away from potentially more useful interventions, and unintended negative consequences from poorly planned and implemented schemes. It is also recognised that much of the gender equality work, for example, in setting up schemes or mentoring female clinical academics, will disproportionately fall to female colleagues. Gender inequality in academic life is a major wrong in terms of the social injustice it represents, and its negative effects on academic productivity and sustainability. We argue that there is a pressing need for better quality implementation and effectiveness research using appropriate research methods to answer questions about mentoring interventions which purport to address this inequality.

Correction notice This article has been corrected since it was published. ‘the women’ has been corrected to ‘fix the women’ under Discussion.

Acknowledgements The authors would like to thank the School of Medicine at the University of Leeds for funding this review to support the work of our gold Athena SWAN award https://medicinеhealth.leeds.ac.uk/medicinе/doc/equality-inclusion-4. Contributors AH led the study and was primarily responsible for oversight of the project design, data synthesis and analysis. AH led on the writing of this original article. ND was the information specialist who designed and ran the literature searches. PB was a Research Fellow in an earlier study that aimed to identify interventions to support gender equality in academic medicine including mentoring. PB contributed to the drafting of this article. WW was coinvestigator on the earlier study and was involved in synthesising interventions to support gender equality in academic medicine including mentoring. WW contributed to the drafting of this article. LDB was principal investigator on the earlier study which identified and synthesised interventions to support gender equality in academic medicine including mentoring. LDB made a significant contribution to the drafting of this article.

Funding This research was funded by the Faculty of Medicine and Health, University of Leeds award/grant number is not applicable.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study. The nature of the evidence does not support a data pooling meta-analysis and so there is no secondary dataset.

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
Allan House http://orcid.org/0000-0001-8721-8026
Louise D Bryant http://orcid.org/0000-0002-1972-7395

REFERENCES