BMJ Open

Assessing the effect of empathy-enhancing interventions in health education and training: a systematic review of randomised controlled trials

Rachel Winter [ORCID],[1] Eyad Issa,[1] Nia Roberts,[2] Robert I Norman,[1] Jeremy Howick[3]

[1]College of Life Sciences, University of Leicester, Leicester, UK
[2]Bodleian Health Care Libraries, University of Oxford, Oxford, UK
[3]Faculty of Philosophy, University of Oxford, Oxford, UK

Correspondence to
Dr Rachel Winter;
rw205@le.ac.uk

ABSTRACT

Objective To estimate the effect of empathy interventions in health education and training from randomised controlled trials (RCTs).

Methods MEDLINE, PsycINFO, EMBASE, CINAHL and Cochrane databases were searched from inception to June 2019 for RCTs investigating the effect of empathy-enhancing interventions in medical and healthcare students and professionals. Studies measuring any aspect of 'clinical empathy' as a primary or secondary outcome were included. Two reviewers extracted data and assessed the risk of bias of eligible studies using the Cochrane Risk of Bias Tool. Random effects meta-analyses of the impact of empathy training on participants' empathy levels were performed.

Results Twenty-six trials were included, with 22 providing adequate data for meta-analysis. An overall moderate effect on participant empathy postintervention (standardised mean difference 0.52, 95% CI 0.36 to 0.67) was found. Heterogeneity across trial results was substantial (I^2=63%). Data on sustainability of effect was provided by 11 trials and found a moderate effect size for improved empathy up until 12 weeks (0.69, 95% CI 0.23 to 1.15), and a small but statistically significant effect size for sustainability at 12 weeks and beyond (standardised mean difference 0.34, 95% CI 0.11 to 0.57). In total, 15 studies were considered to be either unclear or high risk of bias. The quality of evidence of included studies was low.

Conclusion Findings suggest that empathy-enhancing interventions can be effective at cultivating and sustaining empathy with intervention specifics contributing to effectiveness. This review focuses on an important, growing area of medical education and provides guidance to those looking to develop effective interventions to enhance empathy in the healthcare setting. Further high-quality trials are needed that include patient-led outcome assessments and further evaluate the long-term sustainability of empathy training.

Protocol registration number PROSPERO (CRD42019126843).

Strengths and limitations of this study

► This is an up-to-date review that excludes non-randomised studies, follows a prepublished protocol and measures the longer term effects of empathy training.

► The quality of the review was limited by the reporting quality of some of the included studies.

► The studies in our review were heterogeneous, which we anticipated.

► We found only four studies that followed-up participants for at least 3 months.

central to all professional healthcare education.[7] Empathy in the clinical setting has been defined in various ways[8] and can be considered as a multidimensional construct incorporating affective, cognitive, behavioural and moral components.[9] A widely accepted definition of clinical empathy involves the ability to understand the patient's situation, perspective and feelings, communicate that understanding to them and act on it in a helpful and therapeutic way.[10] There is still, however, little consensus on the precise nature of clinical empathy, not least reflected in the variety of tools and scales available to measure it. No guidance exists on how to select measures for assessing clinical empathy, and choice of tools is likely to be led by the definition of empathy used or specific domain being measured.[11] A recent systematic review[11] on empathy measurement tools for care professionals identifies certain measures as scoring highest for quality, but concedes even these had low scores in some of the criteria they used.

Although contested by some,[12 13] there is evidence that empathy in medical and healthcare students declines during undergraduate education.[14–16] Researchers agree that empathetic skills can be taught[17–20] and cultivating empathy to protect against a possible decline would seem sensible. No standard empathy

INTRODUCTION
Rationale

Clinical empathy has multiple benefits for patient care[1–4] and practitioner health.[5 6] Indeed, person-centred and empathic care are

curriculum for healthcare training currently exists and empathy-based training does not appear routinely in healthcare education.[14] Understanding what type of empathy training is most effective in healthcare at both cultivating and sustaining empathy would be a useful start in preparing one.

Four systematic reviews of empathy-promoting interventions have been conducted.[17 20–22] Kelm *et al*[17] conducted a qualitative synthesis of empathy-cultivating interventions for medical students or physicians. Their findings support the hypothesis that interventions can increase physician and medical student empathy. However, they also identified a lack of rigorous study design in most studies (such as lack of control groups). More recently, Kiosses *et al*[20] published a systematic review of randomised controlled trials (RCTs) of empathy-promoting interventions for health professionals. However, only 2 out of 17 included reported change in empathy as a primary outcome, focusing instead on general communication skills. Hence, the review did not provide robust evidence of empathy-enhancing interventions. In 2019, Patel *et al*[21] reviewed educational interventions aimed at enhancing both empathy and compassion. They included observational as well as randomised studies and looked only at physicians and physicians-in-training. They were not able to pool their results statistically and did not investigate whether potential benefits of empathy were sustained over time. With the most recent review, Fragkos and Crampton[22] conclude that empathy interventions significantly increase empathy, but limit their study population to medical students only. In addition, they do not explore whether any improvement in empathy is sustained over time.

These problems listed above present barriers for medical educators looking to implement empathy training into their curricula. It is unclear how large the effect size of effective empathy training is; whether the effect is sustained over time or how best to train students and continuing learners from various health backgrounds. It is important to measure the effect of empathy training, both postintervention and sustainability of effect over time. Arthur *et al*[23] found no effect of empathy training immediately after the training, but significant improvement 12 weeks after the end of the training. A delayed improvement in empathy could potentially be accounted for by participants only recognising the benefits of training once they are putting any lessons learnt into action.

In this systematic review and meta-analysis, we addressed these gaps, with an up-to-date synthesis of RCTs of interventions aimed at promoting empathy, delivered to both medical and healthcare students and professionals, with results that are generalisable to all healthcare contexts. In addition, we will consider both immediate and longer term impact of interventions on empathy.

Objectives

The overarching objective of this systematic review and meta-analysis is to combine data from all available RCTs of empathy-enhancing educational interventions in health education and training. This was achieved with two subsidiary objectives:

1. To assess the effectiveness of empathy-enhancing interventions aiming to enhance empathy in undergraduate and postgraduate health education and training.
2. To assess any lasting effect of empathy training.

We also had three secondary aims:

1. To identify the intervention type (eg, communication skills training) that is most effective at enhancing empathy.
2. To identify the duration of training that is most effective.
3. To identify the tools used to measure empathy levels in participants to consider differences in self-reported and observer-reported measures.

METHODS

Protocol and registration

In accordance with the Cochrane Handbook for systematic reviews of interventions,[24] we published a protocol for this systematic review,[25] registered with PROSPERO international prospective register of systematic reviews (registration number CRD42019126843). We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[26]

Eligibility criteria

RCTs investigating the effect of empathy-enhancing interventions on medical and other healthcare students' and professionals' empathy levels as a primary or secondary outcome were eligible for inclusion. Trials measuring empathy via self-reported and/or observer-reported measures were included. See online supplemental eMethods for further details.

Information sources and search strategies

The following databases were searched from inception to 6 June 2019: MEDLINE, PsycINFO, EMBASE, CINAHL and Cochrane. Search strategies are detailed within online supplemental eTable 1. Electronic searches were supplemented by hand-searching the references of retrieved papers.

Study selection

All studies retrieved through the search strategy were stored using EndNote with duplicates removed. Two authors (RW and EI) reviewed titles and abstracts to identify those meeting inclusion criteria. Full text manuscripts were retrieved for potentially relevant articles. If the decision to include or exclude was unclear, the study was discussed with a third author (JH) to reach a consensus. Seven papers were discussed with the third author. A PRISMA flow chart recorded the screening and selection process.

Data collection

One reviewer (RW) extracted, summarised and recorded data to assess quality and synthesise evidence from included studies. A second author (JH) independently extracted a random sample (10%) of studies to ensure agreement on the information extracted and summarised. See online supplemental eMethods for details on information extracted. If data was not reported, study authors were contacted.

Risk of bias in individual studies

Risk of bias was assessed using the Cochrane Collaboration's Tool for assessing the risk of bias in clinical trials (see online supplemental eMethods for further details). Using the criteria provided by Higgins and Green,[24] each item was scored as high, low or unclear risk of bias, and evidence from the study was used to justify each score given. Given that evidence increasingly suggests that sequence generation and allocation concealment are of particular importance in determining the overall risk of bias,[24] a study was classed as being at high risk of bias if it scored as high or unclear risk on either of these domains.

Synthesis of results

We calculated the overall effect size of empathy interventions using the standardised mean difference (SMD) and 95% CIs based on the data provided in the studies: postintervention sample size, mean and SD for experimental versus control group (except where only mean difference and SD between preintervention and postintervention for the experimental and control groups were provided). We used a random effects model (REM) to allow for likely different (though related) intervention effects. If a study had more than one intervention arm, we used the results for the most comprehensive training intervention. If a study provided measures of empathy using different tools, the primary tool to measure empathy was used. If it was unclear which was the primary measure, we used the first reported measure of empathy.

Heterogeneity was anticipated between studies and assessed using Cochran's Q Statistic (heterogeneity was declared if p value <0.10) and quantified using the I^2 statistic, with an I^2 value of 50% or more being considered to represent levels of heterogeneity.

Primary analysis included all studies providing the data needed to calculate the mean and SD (or SE) of the postintervention control and intervention groups. Where studies provided more than one point for outcome assessment, the data closest to the end-point of the intervention was used. Studies that provided no numerical data on empathy-related outcomes or data from which it was not possible to calculate mean values and SD were excluded from the meta-analysis.

Additional analyses

We performed a sensitivity analysis excluding studies that were considered to be at high risk of bias (scoring unclear or high risk of bias for either sequence generation or allocation concealment, with evidence suggesting these domains are of particular importance in establishing risk of bias).[24]

We conducted separate meta-analyses to look at: sustainability of the effects of the intervention; the intervention type that is most effective; the duration of intervention that is most effective; the outcome assessment tools (comparing objective and subjective outcome measures) and participant populations (effectiveness of interventions aimed at student populations compared with those aimed at professional populations). See online supplemental eMethods for further details.

Risk of bias across studies

Reporting bias was assessed qualitatively based on inspection of the characteristics of the studies included. A funnel plot was produced to investigate small study effects, which may indicate the presence of publication bias. The GRADE system was used to evaluate the overall quality of evidence for the primary outcome.[27]

Patient and public involvement

This research was done without patient involvement. Patients were not invited to comment on the study design and were not consulted to develop patient-relevant outcomes or interpret the results. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy.

RESULTS
Study selection

The literature search resulted in 4904 citations with duplicates removed. Figure 1 provides an overview of the selection process (seeonline supplemental eResultsfor further
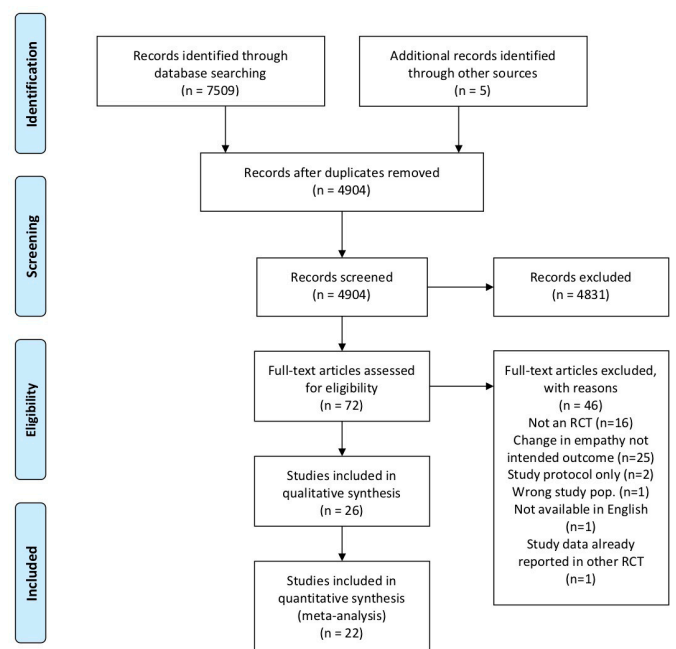

Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram. RCT, randomised controlled trial.

details). Seventy-two articles were retrieved for full-text review. Forty-six studies were excluded (see online supplemental eTable 2). Twenty-six trials were included[23 28–52] (n=2900). Table 1 provides a summary of characteristics (online supplemental eTable 3 gives further details).

Study characteristics

Study publication dates ranged from 1987 to 2019, with 15 out of 26 trials published in the last 5 years.[23 28 30 32 34–38 40 42 47 49 51 52] Thirteen were carried out in the USA and Canada,[29 32–34 40–43 45 46 48 50] seven in Europe,[23 30 36–38 47 51] three in Iran[35 39 49] and one each in Australia,[31] Ghana[28] and China.[52] Fourteen studies provided a definition of empathy.[30 32 34–37 40 43–47 51 52]

Study design

Sample size ranged from 12 to 352 participants (median 90.5; IQR 49.25–154). Thirteen studies had 100 or more participants.[23 28 29 36–38 47 49–51] Seven had fewer than 50 participants.[31 32 40 41 43 46 48] Fifteen studies evaluated empathy interventions for student populations,[28 30 32–35 39 41–43 45 49–52] including seven that looked at medical students,[30 34 35 42 45 50 51] five with nursing students,[33 39 40 50 52] two with student pharmacists,[32 41] one with physiotherapy students[43] and one with a mixed nursing and midwifery student population.[28] Ten trials used professional/qualified populations,[23 29 30 36–38 44 46–48] with four of these focusing on physicians,[31 44 46 48] one on nurses[36] and five with qualified care staff, including healthcare assistants.[23 29 37 38 47] One study used a mixed student and professional population (nursing students and nurse practitioners).[40]

Five trials used multiple sites,[23 30 36 37 40] and five were cluster RCTs.[23 36 37 49 52] Ten studies defined both inclusion and exclusion criteria for the study.[23 28 29 35 37 39 41 49 52] Thirteen defined inclusion criteria only[30–33 36 38 40 42 43 45–47 50] and in three studies inclusion/exclusion criteria were either not given or not clear.[34 48 51]

Study interventions

While the aims of eligible trials in this review were to enhance empathy through an educational intervention, a range of intervention types were employed. The most commonly used approach was a communication skill-based training intervention, with eight studies[28 31 33 34 39 42 48 50] using this. Four studies used perspective-taking training,[23 40 41 49] two had a psychotherapy focus,[30 46] three used empathy skill-based training sessions,[35 44 51] two used an arts and humanities approach,[32 52] one used mindfulness-based training[45] and one a serious gaming intervention.[47] Five studies could not be classified and were described as 'mixed' interventions, using various elements of theoretical knowledge teaching and experiential learning sessions.[29 36–38 43] Seventeen were specifically designed to foster empathy[23 32 34–37 39–44 46–48 50 52] and the remainder used interventions not specifically designed to improve empathy but with the hypothesis that they would. For example, Buffel du Vaure et al[30] explored the impact of

a psychotherapy-focused 'Balint Group' intervention on medical student empathy.

The most frequently used mode of delivery was face-to-face, with 18 interventions using this.[23 28 30 31 33 35–37 40–42 44–46 49–52] Six interventions were delivered online,[29 34 38 39 42 47] one employed a self-directed mode of delivery[32] and one a CD-ROM to deliver the intervention.[48]

Studies ranged in duration of intervention (total time spent participating in the intervention) from 20 min to 42 hours. The mean duration was 10.2 hours (SD 8.8). Five studies did not explicitly state duration.[34 36 38 46 48] Training packages in six studies were considered to be 'short duration', lasting 3 hours or less[32 37 39 42 47 49]; 10 were considered 'medium duration', lasting between 4 and 12 hours[23 28–30 35 39 43 44 50 51]; and 5 were considered 'long duration', lasting more than 12 hours.[31 33 41 45 52] Timespan of the interventions ranged from 1 to 120 days, with a mean length of 38.5 days (SD 40.2).

Outcome measures

Studies used either self-report or other-(objective) report measures to assess a change in participants' empathy. Objective measures included those completed by patients or experts (eg, faculty staff or trained actors playing simulated patients). Most studies (18) used only self-report measures.[23 28 29 32 33 35–39 41–43 45 47–50 52] Four used objective measures[31 34 46 48] and four used a combination of self-report and objective-report tools to measure empathy.[30 40 44 51]

The Jefferson Scale of Empathy (JSE)[53] was the most frequently used self-reported outcome measurement tool, with 13 studies employing it.[23 28 29 32 35 36 39 41–44 49 51 53] Other self-report tools used included the Balanced Emotional Empathy Scale (BEES),[54] the Ekman Facial Decoding test,[55] and the Toronto Empathy Questionnaire (TEQ).[56] The Consultation and Relational Empathy Scale (CARE)[57] was the most frequently used objective measure of empathy, with three studies employing it.[30 40 44] Other objective outcome measures included the Carkhuff Empathy Rating Scale.[58] In addition, some studies developed their own measures of empathy, for example, Tulsky et al[48] used a Likert scale with 10 items to assess perceived oncologist empathy. Butow et al[31] created a manual to code transcripts of videoed patient interactions to assess empathic behaviour, in addition to using the CARE scale.[57] All studies except three[29 31 48] employed a validated tool to measure empathy.

Outcome assessment strategy

Timeframes for measuring outcomes varied. Fifteen studies did not specify a timeframe for postintervention measurements or were unclear.[31–33 35–38 40 41 43–52] For example, Hastings et al[37] reported measuring empathy 6 weeks post-randomisation but were not clear how long after the intervention had ended that this measurement was taken. For studies that were explicit, postintervention measures varied between 2 days and 6 months, with the

Table 1 Summary of characteristics of included studies

Study	Year	County	No participants	Participant type	Intervention type	Duration of intervention (hours)	Outcome assessor	Outcome measure	Effect of intervention
Alhassan[28]	2019	Ghana	210	Nursing and midwifery students	Communication skills training	12	Self	JSE	No significant effect found
Arthur et al[23]	2017	UK	112	Healthcare assistants	Perspective-taking training	12	Self	JSE	No significant effect found
Blair Irvine et al[29]	2012	USA	172	Healthcare professionals	Mixed	4	Self	VST	Significant effect found
Buffel du Vaure et al[30]	2017	France	352	Medical students	Balint group	10.5	Self / Observer	JSE / CARE	Mixed / No significant effect for JSE, significant effect for CARE
Butow et al[31]	2007	Australia	30	Physicians	Communication skills training	15	Observer	CRP	No significant effect found
Collins et al[33]	2017	USA	25	Student pharmacists	Literature intervention	2	Self	JSE	No significant effect found
Daniels et al[33]	1998	Canada	53	Nursing students	Communication skills training	18	Self / Self	ECRS / CIC	Significant effect found
Foster et al[34]	2016	USA	70	Medical students	Communication skills training	NE	Observer	ECCS	Significant effect found
Gholamzadeh et al[35]	2018	Iran	63	Nursing students	Empathy skills training	8	Self	JSE	Significant effect found
Gould et al[36]	2017	UK	249	Nursing staff and healthcare assistants	Mixed	NE	Self	JSE	No significant effect found
Hastings et al[37]	2018	UK	236	Qualified care staff	Mixed	3	Self	SECBQ	No significant effect found
Hattink et al[38]	2015	The Netherlands and UK	142	Qualified care staff	Mixed	NE	Self	IRI	Significant effect found
Larti et al[39]	2014	Iran	82	Nursing students	Communication skills training	12	Self	JSE	Significant effect found
Lobchuk et al[40]	2018	Canada	44	Nursing staff and students	Perspective-taking training	2.66	Observer / Self	CARE / CARE (modified)	Mixed / No significant effect found for CARE. Significant effect found on modified CARE
Lor et al[41]	2014	USA	40	Student pharmacists	Perspective-taking training	18	Self	JSE	Significant effect found
LoSasso et al[42]	2017	USA	70	Medical students	Communication skills training	1	Self	JSE	No significant effect found

Continued

Table 1 Continued

Study	Year	County	No participants	Participant type	Intervention type	Duration of intervention (hours)	Outcome assessor	Outcome measure	Effect of intervention
Mueller et al[43]	2001	USA	37	Physical therapy students	Mixed	11	Self	JSE	Significant effect found
Riess et al[44]	2012	USA	99	Physicians	Empathy skills training	4	Observer	CARE	Mixed
							Self	JSE	No significant effect found for CARE, JSE, BEES.
								BEES	Significant effect for EFDT
								EFDT	
Shapiro et al[45]	1998	USA	78	Medical students	Mindfulness training	17.5	Self	ECRS	Significant effect found
Sripada et al[46]	2010	USA	12	Physicians	Psychotherapy intervention	NE	Observer	BLRI	Significant effect found
Sterkenburg and Vacaru[47]	2018	The Netherlands	224	Qualified care staff	Serious game	0.33	Self	EQ	Significant effect found
Tulsky et al[48]	2011	USA	48	Physicians	Communication skills training	NE	Observer	ES	Significant effect found
								EO	
								PE	
Vaghee et al[49]	2018	Iran	127	Nursing students	Perspective-taking training	3	Self	JSE	Significant effect found
Wolf et al[50]	1987	Canada	134	Medical students	Communication skills training	12	Self	HRI	Significant effect found
								MCI	
Wundrich et al[51]	2017	Germany	158	Medical students	Empathy skills training	6	Self	JSE	Mixed.
							Observer	OSCE	No significant effect found for JSE. Significant effect found on OSCE scores
Yang et al[52]	2018	China	177	Nursing students	Narrative medicine intervention	42	Self	JSE	Significant effect found

BEES, Balanced Emotional Empathy Scale; BLRI, Barrett-Lennard Relationship Inventory; CARE, Consultation and Relational Empathy Scale; CIC, Carkhuff Index of Communication; CRP, Coding Role Play; ECCS, Empathic Communication Coding System; ECRS, Empathy Construct Rating Scale; EFDT, Ekman Facial Decoding test; EO, Empathic Opportunities; EQ, Empathy Quotient; ES, Empathic Statements; HRI, Helping Relationship Inventory; IRI, Interpersonal Reactivity Inventory; JSE, Jefferson Scale of Empathy; MCI, Medical Communication Index; OSCE, Objective Structured Clinical Examination; PE, Perceived Empathy; SECBQ, Staff Empathy for People with Challenging Behaviour; VST, Video Situational Testing.

majority of measures taken within 2 weeks of the intervention.[23 28–30 32 41 48] Eleven studies measured the effects at one or more follow-up points (in addition to the postintervention measurement),[23 28 29 31 33 35 37 39 41 49 52] which varied between 4 weeks and 18 months.

Risk of bias within studies

In total, 11 studies[23 28 31 36–39 43 45 47 48] were considered to be at low risk of bias overall (low risk of bias for sequence generation and allocation concealment).[24] Thirteen were considered to be low risk for random sequence generation[23 28 31 43 47 48] and 11 were low risk for allocation concealment.[6–23 23–28 28–31 31–39 43 44 47 48] Blinding was not possible in the majority of studies due to the nature of the interventions (often described to participants as empathy-promoting) and the method of outcome assessment (eg, self-report questionnaires, making explicit what is being measured, such as the JSE). Full details of the risk of bias assessment are reported in the online supplemental eResults with online supplemental eFigure 1 illustrating the overall findings.

Results of individual studies

The majority of studies (19/26) found that the tested intervention significantly improved empathy on at least one outcome measure.[29 30 33–35 38–41 43–52] Seven studies did not find any significant increase in empathy.[23 28 31 32 36 37 42] Of the studies that reported a significant improvement in empathy on at least one outcome measure, 11 were aimed at student populations (representing approximately 73% of student population studies)[30 33–35 41 43 45 49–52] and 7 were aimed at professionals (representing 70% of professional population studies).[29 38 39 44 46–48] Fifteen studies reported a significant improvement in empathy using a self-rated outcome measure (this represents 68% of studies (15/22) using a self-report outcome

tool).[29 30 33 35 38–41 43 45–47 49 50 52] Four studies reported an increase in empathy when using an objective measure (representing 50% (4/8) of studies using an objective outcome measure).[34 44 48 51] Seventeen studies employed an educational intervention that had been specifically designed to foster empathy.[23 32 34–37 39–44 46–48 50 52] Of these, 12 (70%) were successful.[5 34 39–44 46–48 51 52] Four out of five studies that were classed as 'long duration' (lasting >12 hours) reported a significant improvement in empathy postintervention[33 41 45 52]; 50% of 'medium duration' studies (between 3 and 12 hours) reported a significant increase in empathy[29 35 39 50 51] and 33% of 'short duration' studies (<3 hours) reported a significant improvement.[47 49]

Synthesis of results

Of the 26 studies included in this review, 4 were excluded from meta-analysis as they did not provide adequate data from which to calculate the SMD and SD.[31 36 46 51] For the studies that were excluded from the primary analysis, Butow et al[31] reported a positive but not statistically significant effect and Gould et al[36] found no significant difference between control and intervention groups. Wündrich et al[51] reported no significant influence of the intervention as measured by the JSE (student version) but did report a positive and statistically significant effect on the observer-assessed outcome. Sripada et al[46] also reported a statistically significant positive effect. Of the 22 studies that had adequate data for pooling, all but one (Arthur et al[23]) showed a benefit of intervention. The primary analysis identified that the overall effect of empathy interventions in terms of improving participant empathy was statistically significant (SMD 0.52, 95% CI 0.36 to 0.67) (figure 2). The Q value indicated
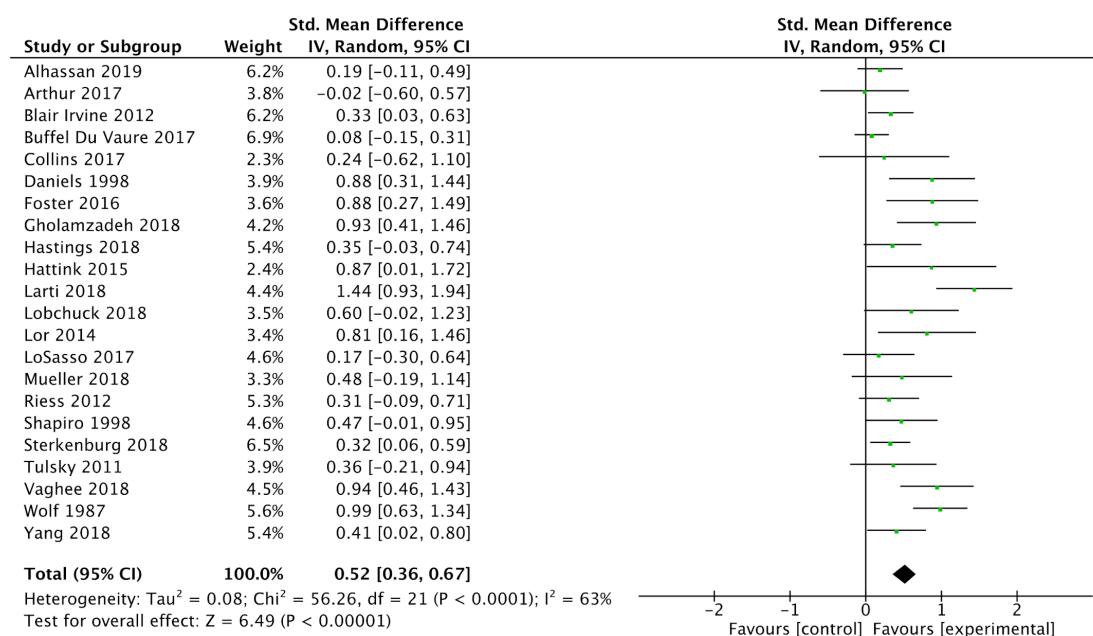
Study or Subgroup	Weight	Std. Mean Difference IV, Random, 95% CI
Alhassan 2019	6.2%	0.19 [−0.11, 0.49]
Arthur 2017	3.8%	−0.02 [−0.60, 0.57]
Blair Irvine 2012	6.2%	0.33 [0.03, 0.63]
Buffel Du Vaure 2017	6.9%	0.08 [−0.15, 0.31]
Collins 2017	2.3%	0.24 [−0.62, 1.10]
Daniels 1998	3.9%	0.88 [0.31, 1.44]
Foster 2016	3.6%	0.88 [0.27, 1.49]
Gholamzadeh 2018	4.2%	0.93 [0.41, 1.46]
Hastings 2018	5.4%	0.35 [−0.03, 0.74]
Hattink 2015	2.4%	0.87 [0.01, 1.72]
Larti 2018	4.4%	1.44 [0.93, 1.94]
Lobchuck 2018	3.5%	0.60 [−0.02, 1.23]
Lor 2014	3.4%	0.81 [0.16, 1.46]
LoSasso 2017	4.6%	0.17 [−0.30, 0.64]
Mueller 2018	3.3%	0.48 [−0.19, 1.14]
Riess 2012	5.3%	0.31 [−0.09, 0.71]
Shapiro 1998	4.6%	0.47 [−0.01, 0.95]
Sterkenburg 2018	6.5%	0.32 [0.06, 0.59]
Tulsky 2011	3.9%	0.36 [−0.21, 0.94]
Vaghee 2018	4.5%	0.94 [0.46, 1.43]
Wolf 1987	5.6%	0.99 [0.63, 1.34]
Yang 2018	5.4%	0.41 [0.02, 0.80]
Total (95% CI)	**100.0%**	**0.52 [0.36, 0.67]**

Heterogeneity: Tau^2 = 0.08; Chi^2 = 56.26, df = 21 (P < 0.0001); I^2 = 63%
Test for overall effect: Z = 6.49 (P < 0.00001)

Favours [control] Favours [experimental]

Figure 2 Meta-analysis of eligible studies providing adequate data to calculate standardised mean difference with 95% CI.

Table 2 Summary of effect sizes for studies included in meta-analyses

	Standardised mean difference (95% CI)	Heterogeneity (I^2)	References
Overall effect of empathy interventions	0.52 (0.36 to 0.67)	63%	8–23 23–30 32–35 37–45 47–49 52
Effect of intervention with least risk of bias	0.44 (0.19 to 0.69)	63%	23 28 37–39 43 44 47 48
Sustainability of effect			
Follow-up measurement before 12 weeks	0.69 (0.23 to 1.15)	84%	28 29 35 37 39 49
Follow-up measurement at 12 weeks or later	0.34 (0.11 to 0.57)	0%	23 37 41 52
Effect by type of intervention			
Communication skills training	0.69 (0.32 to 1.06)	78%	28 33 34 39 42 48 50
Perspective-taking training	0.60 (0.17 to 1.04)	55%	23 40 41 49
Mixed educational programmes	0.39 (0.18 to 0.61)	0%	29 37 38 43
Empathy skills training	0.60 (−0.02 to 1.21)	71%	35 44
Arts/humanities interventions	0.38 (0.03 to 0.73)	0%	32 52
Effect by duration of intervention			
Short (3 hours or less)	0.42 (0.21 to 0.63)	24%	32 37 40 42 47 49
Medium (4–12 hours)	0.51 (0.21 to 0.80)	82%	23 28–30 35 39 43 44 50
Long (more than 12 hours)	0.57 (0.32 to 0.82)	0%	33 41 45 52
Effect by participant population			
Student population	0.62 (0.38 to 0.85)	74%	28 30 32–35 39–43 45 49 50 52
Professional/qualified population	0.33 (0.18 to 0.47)	0%	23 29 37 38 44 47 48
Effect by outcome assessor			
Self-assessment	0.52 (0.37 to 0.68)	58%	23 28–30 32 33 35 37–45 47 49 50 52
Observer-assessment	0.28 (−0.18 to 0.75)	81%	30 34 40 44 48

significant heterogeneity, with p equal to 0.0001 and I^2 equal to 63%. A summary of findings is presented in table 2.
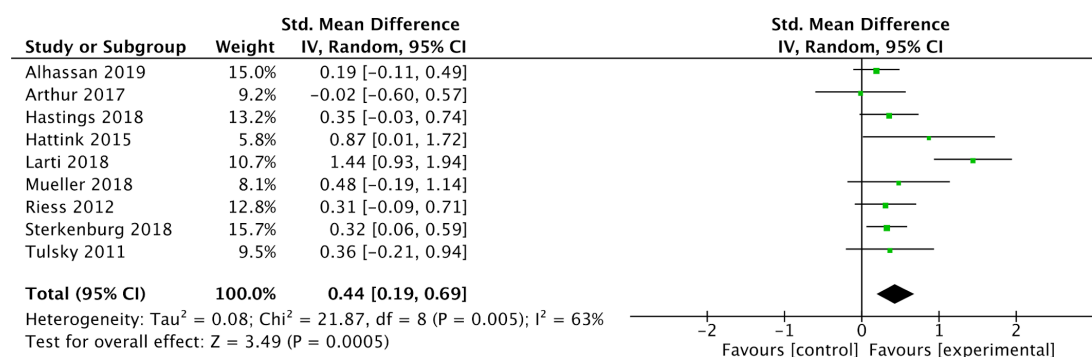
Additional analyses

Sensitivity analysis

For the sensitivity analysis of the least biased studies (table 2), 11 were judged to have low risk of bias for random allocation or allocation concealment[23 28 31 36–39 43 44 47 48] and 9 of these provided sufficient data to be included in a meta-analysis (figure 3).[23 28 37–39 43 44 47 48]

Sustainability of improved empathy analysis

Eleven studies provided follow-up data assessing sustainability of changes to empathy, in addition to postintervention measurement.[23 28 29 31 33 35 37 39 41 49 52] Eight were eligible for inclusion in a subgroup analysis[23 29 35 37 39 41 49 52] (see online supplemental eResults 1 for further details), which found a moderate effect size for sustainability up to 12 weeks and a smaller, but still significant effect size for sustainability of impact of training at 12 weeks or later (figure 4 and table 2).

Type of intervention analysis

A meta-analysis comparing subgroups of different types of intervention (see online supplemental eFigure 2 in the

Study or Subgroup	Weight	Std. Mean Difference IV, Random, 95% CI
Alhassan 2019	15.0%	0.19 [−0.11, 0.49]
Arthur 2017	9.2%	−0.02 [−0.60, 0.57]
Hastings 2018	13.2%	0.35 [−0.03, 0.74]
Hattink 2015	5.8%	0.87 [0.01, 1.72]
Larti 2018	10.7%	1.44 [0.93, 1.94]
Mueller 2018	8.1%	0.48 [−0.19, 1.14]
Riess 2012	12.8%	0.31 [−0.09, 0.71]
Sterkenburg 2018	15.7%	0.32 [0.06, 0.59]
Tulsky 2011	9.5%	0.36 [−0.21, 0.94]
Total (95% CI)	**100.0%**	**0.44 [0.19, 0.69]**

Heterogeneity: Tau² = 0.08; Chi² = 21.87, df = 8 (P = 0.005); I^2 = 63%
Test for overall effect: Z = 3.49 (P = 0.0005)

Figure 3 Meta-analysis of eligible studies, excluding those considered to be at high risk of bias.
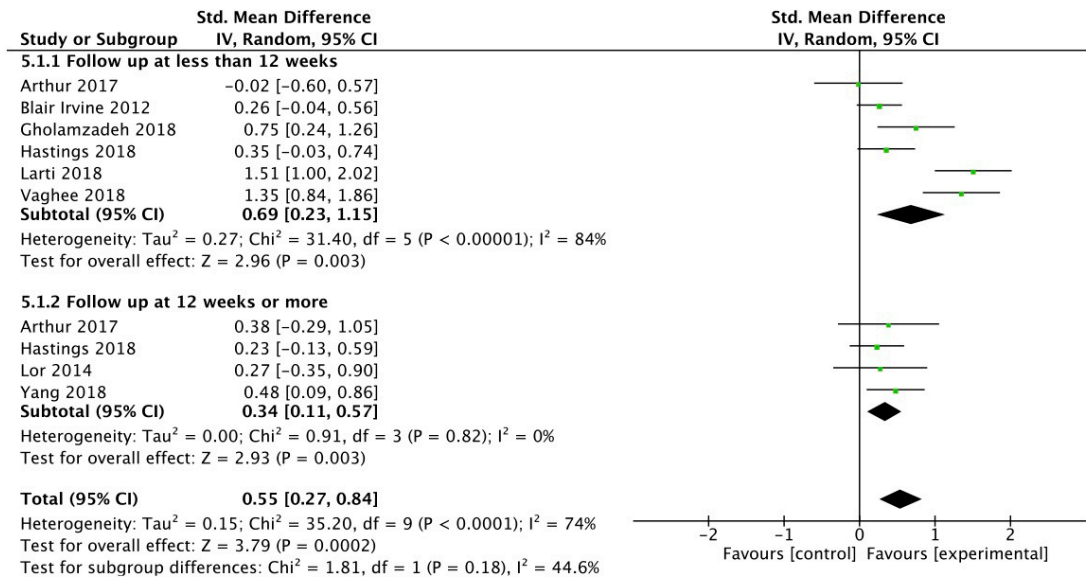
Figure 4 Meta-analysis of studies that provided follow-up observation points to determine long-term effectiveness of intervention.

online supplemental eResults for further details) found the greatest effect was with empathy training that was communication skill-based (table 2). The smallest effect reported was for interventions that were described as 'mixed educational programmes' and ones based in the arts and humanities (table 2). It is worth noting however that only two studies used arts and humanities interventions (compared with seven in the communication skill group) and this may well impact on the effect size.

Duration of intervention analysis
Interventions of medium and longer duration (online supplemental eFigure 3) were most effective. Interventions of short duration had the smallest effect size (table 2).

Participant population analysis
Studies using healthcare student participant populations appeared to have a larger effect size than those directed at professional/qualified populations (online supplemental eFigure 4). Studies included in a subanalysis of interventions for students showed a moderate effect size of training compared with a smaller but still significant effect size for training directed at professional/qualified populations (table 2).

Outcome assessor analysis
Studies using a self-assessment outcome scale showed a moderate and significant benefit to empathy for the intervention tested (online supplemental eFigure 5) compared with a small and statistically not significant effect size for observer-assessed outcome studies (table 2).

Risk of bias across studies
A funnel plot used in the primary meta-analysis (22 studies) did not reveal evidence of publication bias (figure 5). An evaluation of evidence using

GRADE software found the quality of evidence was low (online supplemental eTable 4). This was due to a high or uncertain risk of bias based on random sequence generation and/or allocation concealment in a number of studies and a high degree of heterogeneity across studies.

DISCUSSION
Summary of evidence
Training healthcare practitioners and trainees improved their empathy by a modest amount. The effect of training seemed to diminish, but lasts to beyond 12 weeks.

Comparison with other evidence
Our review supports the evidence of previous similar reviews, finding benefits of empathy training[17 20–22]
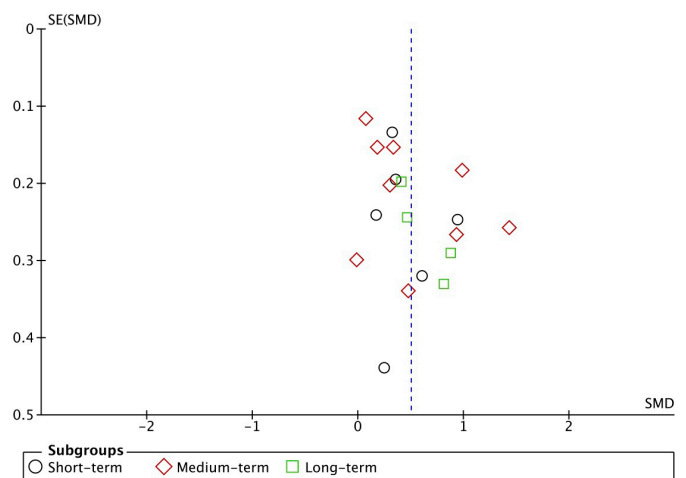

Figure 5 Funnel plot of effect sizes and SEs. SMD, standardised mean difference.

and that practitioner empathy training makes a difference to patients.[59] Our study adds to this evidence by providing an estimate of empathy training from higher quality (randomised) trials and by showing that the effect lasts well beyond the intervention.

Strengths and limitations

This review, to the best of our knowledge, is the first systematic review and meta-analysis limited to RCTs of clinical empathy training for all healthcare students and professionals. This is an up-to-date review that excludes non-randomised studies, follows a prepublished protocol and assesses both the immediate and longer term effects of empathy training. Our broad study population with both healthcare students and professionals means findings are generalisable to all areas of healthcare education and training.

We chose to include only the results of the primary measure of empathy reported by each study. Where it was unclear which was the primary measure, we used the measure that was reported first. We recognise that this might have been biased, as authors may have chosen to report the most positive outcomes first. However, we found that this was not necessarily the case. For example, the first measure of empathy reported by Buffel du Vaure *et al*[30] (who did not specific which measure was primary) had a smaller effect than the second.

We recognise the heterogeneity of the studies in our review and anticipated this. This means that further research is required to identify the most effective empathy training methodology. Also, the strength of findings in this review may be limited by the reporting quality of some of the included studies. A sensitivity analysis of studies of highest quality found a slightly smaller but still significant effect size. Another limitation in reviewing the evidence in this field is the multiple tools used by investigators to measure clinical empathy. With the lack of a definitive definition of clinical empathy and a range of tools measuring different aspects of empathy, the impact of an intervention may vary depending on the measurement tool used. This is demonstrated by Riess *et al*[44] who found a statistically significant improvement in empathy when measured using the CARE scale but no significant changes using the JSE. In contrast, Buffel du Vaure *et al*[30] reported the opposite. Perhaps because of the larger sample size or other factors, our review found a benefit of training independently of how it was measured. A further limitation with this review is that we only identified four studies that followed participants up for at least 3 months. The trials identified however found a positive effect. Lastly, we did not measure the qualitative experiences of participants in this review.

Implications for research and practice

Interventions for cultivating student and trainee empathy should be further developed and implemented. Optimising implementation will require additional qualitative research on the experiences of empathy teachers and learners. Also, the longer term effects (>12 weeks) of empathy training have not been studied adequately and future research should address this. With competition for time and space in both undergraduate and postgraduate healthcare curriculums, future research in this area needs to be robust. Designers of future trials of empathy training in healthcare can use the results of this review as a guide to their intervention development.

CONCLUSION

Teaching students and other learners how to enhance empathy is moderately effective over a sustained period of time and is likely to benefit present and future patients. Future research should focus on empathy interventions with patient-led outcome assessment and on assessing effectiveness of training over more sustained periods of time. Medical educators and curriculum designers can use this research to think of ways to integrate empathy training into busy curricula.

ORCID iD
Rachel Winter http://orcid.org/0000-0002-8775-2461

REFERENCES

1 Kim SS, Kaplowitz S, Johnston MV. The effects of physician empathy on patient satisfaction and compliance. *Eval Health Prof* 2004;27:237–51.

2 Vermeire E, Hearnshaw H, Van Royen P, *et al*. Patient adherence to treatment: three decades of research. A comprehensive review. *J Clin Pharm Ther* 2001;26:331–42.

3 Di Blasi Z, Harkness E, Ernst E, *et al*. Influence of context effects on health outcomes: a systematic review. *Lancet* 2001;357:757–62.

4 Kelley JM, Kraft-Todd G, Schapira L, *et al*. The influence of the patient-clinician relationship on healthcare outcomes: a systematic review and meta-analysis of randomized controlled trials. *PLoS One* 2014;9:e94207.

5 Thomas MR, Dyrbye LN, Huntington JL, *et al*. How do distress and well-being relate to medical student empathy? A multicenter study. *J Gen Intern Med* 2007;22:177–83.

6 DiLalla LF, Hull SK, Dorsey JK, *et al*. Effect of gender, age, and relevant course work on attitudes toward empathy, patient spirituality, and physician wellness. *Teach Learn Med* 2004;16:165–70.

7 The NHS Constitution for England. Health education England. Department of health, 2015. Available: https://assets.publishing. service.gov.uk/government/uploads/system/uploads/attachment_ data/file/480482/NHS_Constituti on_WEB.pdf [Accessed 5 Feb 2019].

8 Cummins J, Bennett V. *Compassion in practice: nursing, midwifery and care staff, our vision and strategy*. Leeds: NHS commissioning board, 2012. https://www.england.nhs.uk/wp-content/uploads/2012/ 12/compassion-in-practice.pdf

9 Francis R. *Report of the mid Staffordshire NHS Foundation trust public inquiry: Executive summary*. London: The Stationery office, 2013. https://assets.publishing.service.gov.uk/government/uploads/ system/uploads/attachment_data/file/279124/0947.pdf

10 Mercer SW, Reynolds WJ. Empathy and quality of care. *Br J Gen Pract* 2002;52:S9–12.

11 Hong H, Han A. A systematic review on empathy measurement tools for care professionals. *Educ Gerontol* 2020;46:72–83.

12 Colliver JA, Conlee MJ, Verhulst SJ, *et al*. Reports of the decline of empathy during medical education are greatly exaggerated: a reexamination of the research. *Acad Med* 2010;85:588–93.

13 Smith KE, Norman GJ, Decety J. The complexity of empathy during medical school training: evidence for positive changes. *Med Educ* 2017;51:1146–59.

14 Neumann M, Edelhäuser F, Tauschel D, *et al*. Empathy decline and its reasons: a systematic review of studies with medical students and residents. *Acad Med* 2011;86:996–1009.

15 Hojat M, Vergare MJ, Maxwell K, *et al*. The devil is in the third year: a longitudinal study of erosion of empathy in medical school. *Acad Med* 2009;84:1182–91.

16 Nunes P, Williams S, Sa B, *et al*. A study of empathy decline in students from five health disciplines during their first year of training. *Int J Med Educ* 2011;2:12–17.

17 Kelm Z, Womer J, Walter JK, *et al*. Interventions to cultivate physician empathy: a systematic review. *BMC Med Educ* 2014;14:219.

18 Chen D, Lew R, Hershman W, *et al*. A cross-sectional measurement of medical student empathy. *J Gen Intern Med* 2007;22:1434–8.

19 Batt-Rawden SA, Chisolm MS, Anton B, *et al*. Teaching empathy to medical students: an updated, systematic review. *Acad Med* 2013;88:1171–7.

20 Kiosses VN, Karathanos VT, Tatsioni A. Empathy promoting interventions for health professionals: a systematic review of RCTs. *J Compassionate Health Care* 2016;3:7.

21 Patel S, Pelletier-Bui A, Smith S, *et al*. Curricula for empathy and compassion training in medical education: a systematic review. *PLoS One* 2019;14:e0221412.

22 Fragkos KC, Crampton PES. The effectiveness of teaching clinical empathy to medical students: a systematic review and meta-analysis of randomized controlled trials. *Acad Med* 2020;95:947–57.

23 Arthur A, Maben J, Wharrad H, *et al*. Can healthcare assistant training (ChAT) improve the relational care of older people? study protocol for a pilot cluster randomised controlled trial. *Trials* 2015;16:559.

24 Higgins JJ, Green S. *The cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]*. Chichester: The Cochrane Collaboration, 2011.

25 Winter R, Roberts N, Howick J. *Assessing the effect of empathy-enhancing interventions in health education and training: a systematic review of randomised controlled trials*. PROSPERO: International prospective register of systematic reviews, 2019.

26 Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.

27 Schunemann H, Brozek J, Oxman A. *Grade Handbook for grading quality of evidence and strength of recommendation*. Hamilton: Group TGW, 2008.

28 Alhassan M. Effect of a 2-day communication skills training on nursing and midwifery students' empathy: a randomised controlled trial. *BMJ Open* 2019;9:e023666.

29 Blair Irvine A, Billow MB, Eberhage MG, *et al*. Mental illness training for licensed staff in long-term care. *Issues Ment Health Nurs* 2012;33:181–94.

30 Buffel du Vaure C, Lemogne C, Bunge L, *et al*. Promoting empathy among medical students: a two-site randomized controlled study. *J Psychosom Res* 2017;103:102–7.

31 Butow P, Cockburn J, Girgis A, *et al*. Increasing oncologists' skills in eliciting and responding to emotional cues: evaluation of a communication skills training program. *Psychooncology* 2008;17:209–18.

32 Collins KL, Zweber A, Irwin AN, *et al*. Impact of a fictional reading intervention on empathy development in student pharmacists. *Curr Pharm Teach Learn* 2017;9:498–503.

33 Daniels TG, Denny A, Andrews D. Using microcounseling to teach RN nursing students skills of therapeutic communication. *J Nurs Educ* 1988;27:246–52.

34 Foster A, Chaudhary N, Kim T, *et al*. Using virtual patients to teach empathy: a randomized controlled study to enhance medical students' Empathic communication. *Simul Healthc* 2016;11:181–9.

35 Gholamzadeh S, Khastavaneh M, Khademian Z, *et al*. The effects of empathy skills training on nursing students' empathy and attitudes toward elderly people. *BMC Med Educ* 2018;18:198.

36 Gould LJ, Griffiths P, Barker HR, *et al*. Compassionate care intervention for hospital nursing teams caring for older people: a pilot cluster randomised controlled trial. *BMJ Open* 2018;8:e018563.

37 Hastings RP, Gillespie D, Flynn S, *et al*. Who's challenging who training for staff empathy towards adults with challenging behaviour: cluster randomised controlled trial. *J Intellect Disabil Res* 2018;62:798–813.

38 Hattink B, Meiland F, van der Roest H, *et al*. Web-Based StAR e-learning course increases empathy and understanding in dementia caregivers: results from a randomized controlled trial in the Netherlands and the United Kingdom. *J Med Internet Res* 2015;17:e241.

39 Larti N, Ashouri E, Aarabi A. The effect of an empathy role-play program for operating room nursing students. *J Educ Eval Health Prof* 2018;15:29.

40 Lobchuk M, Hoplock L, Halas G, *et al*. Heart health whispering: a randomized, controlled pilot study to promote nursing student perspective-taking on carers' health risk behaviors. *BMC Nurs* 2018;17:21.

41 Lor KB, Truong JT, Ip EJ, *et al*. A randomized prospective study on outcomes of an empathy intervention among second-year student pharmacists. *Am J Pharm Educ* 2015;79:18s.

42 LoSasso AA, Lamberton CE, Sammon M, *et al*. Enhancing student Empathetic engagement, history-taking, and communication skills during electronic medical record use in patient care. *Acad Med* 2017;92:1022–7.

43 Mueller K, Prins R, de Heer HD. An online intervention increases empathy, resilience, and work engagement among physical therapy students. *J Allied Health* 2018;47:196–203.

44 Riess H, Kelley JM, Bailey RW, *et al*. Empathy training for resident physicians: a randomized controlled trial of a neuroscience-informed curriculum. *J Gen Intern Med* 2012;27:1280–6.

45 Shapiro SL, Schwartz GE, Bonner G. Effects of mindfulness-based stress reduction on medical and premedical students. *J Behav Med* 1998;21:581–99.

46 Sripada BN, Henry DB, Jobe TH, *et al*. A randomized controlled trial of a feedback method for improving empathic accuracy in psychotherapy. *Psychol Psychother* 2011;84:113–27.

47 Sterkenburg PS, Vacaru VS. The effectiveness of a serious game to enhance empathy for care workers for people with disabilities: a parallel randomized controlled trial. *Disabil Health J* 2018;11:576–82.

48 Tulsky JA, Arnold RM, Alexander SC, *et al*. Enhancing communication between oncologists and patients with a computer-based training program: a randomized trial. *Ann Intern Med* 2011;155:593–601.

49 Vaghee S, Kashani Lotfabadi M, Salarhaji A, *et al*. Comparing the effects of contact-based education and acceptance and Commitment-Based training on empathy toward mental illnesses among nursing students. *Iran J Psychiatry* 2018;13:119–27.

50 Wolf FM, Woolliscroft JO, Calhoun JG, *et al*. A controlled experiment in teaching students to respond to patients' emotional concerns. *J Med Educ* 1987;62:25–3.

51 Wündrich M, Schwartz C, Feige B, *et al*. Empathy training in medical students - a randomized controlled trial. *Med Teach* 2017;39:1096–8.

52 Yang N, Xiao H, Cao Y, *et al*. Does narrative medicine education improve nursing students' empathic abilities and academic achievement? a randomised controlled trial. *J Int Med Res* 2018;46:3306–17.

53 Hojat M, Gonnella JS, Nasca TJ, *et al*. Physician empathy: definition, components, measurement, and relationship to gender and specialty. *Am J Psychiatry* 2002;159:1563–9.

54 Mehrabian A, Epstein N. A measure of emotional empathy. *J Pers* 1972;40:525–43.

55 Ekman P. *Emotions revealed: recognizing faces and feelings to improve communication and emotional life*. New York: Holt, 2004: 328. 0405184.

56 Spreng RN, McKinnon MC, Mar RA, *et al*. The Toronto empathy questionnaire: scale development and initial validation of a factor-analytic solution to multiple empathy measures. *J Pers Assess* 2009;91:62–71.

57 Mercer SW, Maxwell M, Heaney D, *et al*. The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation process measure. *Fam Pract* 2004;21:699–705.

58 Carkhuff RR. *Helping and human relations: a primer for lay and professional helpers*. Montreal-Toronto, Holt: Rinehart and Winston Inc, 1969.

59 Howick J, Moscrop A, Mebius A, *et al*. Effects of empathic and positive communication in healthcare consultations: a systematic review and meta-analysis. *J R Soc Med* 2018;111:240–52.