Kiosses et al. Journal of Compassionate Health Care (2016) 3:7
DOI 10.1186/s40639-016-0024-9

Journal of Compassionate Health Care

REVIEW Open Access

Empathy promoting interventions for health professionals: a systematic review of RCTs

Vassilios N. Kiosses[1], Vassilios T. Karathanos[1] and Athina Tatsioni[1,2*]

Abstract

Purpose: Authors assessed systematically the effectiveness of interventions aiming at health professionals' increase of empathic responses.

Method: Authors searched Pubmed, Cochrane Database of Clinical Trials, Scopus, and PsycInfo for randomized controlled trials (RCTs) (latest search on November 2012). They included trials in English that evaluated interventions, which may promote empathy in health professionals. Studies were categorized according to the type of the outcome on empathy (attitude or opinion, knowledge or skills, and behavior). Authors considered change in empathy as the main outcome. Standardized mean differences (SMD) with 95 % confidence interval (95 % CI) were calculated for the studies that provided adequate data. Primary analysis included all the studies that provided adequate data per outcome category. In addition, authors proceeded in subgroup analyses for the following groups (a) type of intervention (experiential vs. non-experiential; (b) training specifically for promoting empathy vs. other; (c) type of assessor (external observer, health professional participants, and patients); and (d) type of process used for empathy evaluation, (simulated interview, actual interview, and questionnaire completion without interview).

Results: Out of 722 items, 17 articles were eligible. Trials were highly heterogeneous in terms of participants, interventions, and outcome measures. Interventions usually covered a broad training on communication skills. Thirteen studies used experiential while four non-experiential learning approaches. There were only two studies that evaluated interventions specifically aiming at promoting empathy; one of these trials reported significant increase in residents' empathy related knowledge, skills, and behavior. Based on 13 trials with adequate data, health professionals in intervention group improved empathic behavior when compared to control group (SMD 0.8, 95 % CI 0.4, 1.2; *P* value <0.001). None of the trials assessed patients' health care outcomes.

Conclusion: There are interventions, which may contribute to a significant improvement in empathic behavior among health professionals. However, the type of intervention that would be effective needs to be supported by future studies. Whether empathic behavior may last, or whether it may affect patients' outcomes is yet to be defined.

Keywords: Empathy, Health care, Randomized controlled trial, Systematic review, Meta-analysis

* Correspondence: atatsioni@gmail.com
[1]University of Ioannina School of Health Sciences, Faculty of Medicine, Ioannina, Greece
[2]Tufts University School of Medicine, Boston, MA, USA

Background

Empathy is described sometimes as a cognitive attribute featuring understanding of experiences of others; at other times, as an emotional state of the mind featuring sharing of feelings; and at still other times as a concept involving both cognition and emotion [1]. People, who demonstrate it, identify with another's feelings. The ability to empathize is directly dependent on a person's ability to feel his/her own feelings and identify them [2, 3]. Health professional—patient communication is the means to deliver care that is adapted to an individual's emotional, cognitive, and biological needs. Clinical empathy includes understanding the patient's situation, perspective and feelings as well as their attached meanings; communicating understanding and checking its accuracy; and acting on that understanding with the patient in a therapeutic way [4]. Medical and other health care professional schools have included educational interventions to maintain and enhance empathy in undergraduate students [5]. The reason for which clinical empathy has been introduced to health care curricula is related to empathy's expected positive attributes, including dutifulness, prosaic behavior, moral reasoning, reduced malpractice litigation, improved history taking and physical examination, patient satisfaction, physician satisfaction, improved therapeutic relationships, and overall improved clinical outcomes [6–14]. However, to our knowledge there is no systematic approach to indicate whether interventions aiming at the improvement of health professionals' empathy may contribute to any of the above outcomes. Limited evidence mainly from clinical studies with non-experimental design has supported the correlation of empathy with patient outcomes [5, 15]. This may have led some researchers to question whether enhancing empathy would have any incremental beneficial effect on medical care [16]. Moreover, reduced empathy may sometimes allow physicians to complete clinical tasks more accurately [17, 18]. In certain clinical contexts, such as surgery [14] or oncology, is argued that keeping an emotional distance from patients maintains clinical neutrality [19] while being empathic has a psychological cost for health care professionals [20, 21], which can lead to 'compassion fatigue' [22].

The ambiguity associated with the definition of empathy obstructs investigators to clearly see what they intend to study, and hinders their ability of how to measure it in the context of patient care [1]. Social relationships may require both mutual understanding and feeling of emotions. However, for patient—physician relationships, health professionals need to be aware of patient's concerns. Empathy in patient care has been introduced as a multidimensional concept involving at least three factors: "perspective taking," "compassionate care," and "standing in the patient's shoe" [1]. The ability to capture all three dimensions in studies evaluating empathy has important implications not only for the conceptualization and measurement of empathy in patient care but also for the assessment of patient outcomes [1]. In addition, research findings on empathy can be subject to serious challenges if the conceptualization, definition, and measurement issues remain unsettled [1].

Previous work [1] presented studies with randomized and non-randomized designs that included a wide variety of interventions aimed at enhancing empathy either by evaluating an empathy-focused training, or by evaluating empathy training as part of a communication skills training program for health professionals. This extended work offered substantial insight on the professional groups that may receive the intervention, the type of interventions, and the type of measures for empathy. However, several issues remained unsolved including what type of interventions are effective, and which outcomes may actually be improved. Therefore, the authors of the present paper tried to systematically assess the extent to which interventions aiming at the improvement of health professionals' empathy were evaluated in randomized controlled trials (RCTs). Authors focused on RCTs in order to capture the best quality information and to ensure the highest robustness of the results [23]. Developing an intervention for improving empathy may need additional methodological approaches, such as the use of theory of planned behavior. However, in order to support effectiveness and ensure reproducibility of the results among health professionals, the assessment of the developed intervention needs to be supported by well-designed RCTs [23]. In an effort to clarify what intervention might be promising for improving empathy and whether the evaluation mode for the intervention correlated with the results, authors considered a number of factors in their analyses, including the type of intervention, the type of assessors, and the type of process for assessing empathy. To address the challenges in the type of measures for empathy, they categorized outcomes on attitudes, knowledge, skills, and behavior. To increase generalizability, authors included interventions for all health professionals both at undergraduate and postgraduate level. For interventions specifically aiming at increasing empathy, authors also explored whether studies assessed the impact of these interventions on patients' health care outcomes.

Methods

Search strategy

Authors searched Pubmed, Cochrane Database of Clinical Trials, Scopus, and PsycInfo (from inception to November 2012) using the following search algorithm: ("empathy"[MeSH Terms] OR "empathy"[All Fields]) AND ("Clinical Trials as Topic"[Mesh] OR "randomized

controlled trial"[pt] OR "controlled clinical trial"[pt] OR randomized [tiab] OR placebo[tiab] OR randomly[tiab] OR trial[ti]) AND ("Clinical Trials as Topic"[Mesh] OR "randomized controlled trial"[pt] OR "controlled clinical trial"[pt] OR randomized[tiab] OR placebo[tiab] OR randomly[tiab] OR trial[ti]). Electronic searches were supplemented by perusal of the references of the retrieved papers as well as the references of review articles. Two independent investigators (VNK, VTK) screened abstracts and papers in full text. Discrepancies were resolved with consensus and the participation of an arbitrator (AT) where necessary.

This systematic review was performed according to PRISMA guidelines [24].

Eligibility criteria

Authors included only randomized control trials (RCTs) irrespective of the type, i.e., parallel, crossover, cluster, and pragmatic design, which evaluated training interventions and included empathy change in health professionals, or health care students during their encounters with patients as an outcome. Authors included both trials that evaluated training for specifically promoting empathy and studies that assessed interventions aiming at communication or interpersonal skills. They considered as eligible both studies with a clear definition of empathy and articles that did not include any clarification. They did not set any exclusion criteria for the type of measures that investigators employed to assess change in empathy.

In case a trial was reported in multiple papers (duplicated publications), authors considered as eligible the paper including the most complete information. They excluded RCTs that were published at the protocol stage, RCTs that may have measured but did not report results on eligible outcomes, and studies that were not written in English.

Data extraction

Data were extracted in predefined forms. Two independent investigators (VNK, VTK) extracted all data. Discrepancies were resolved with consensus and the participation of a third arbitrator (AT) where necessary.

Extracted items included name of first author, year of publication, country, study design, sample size, description of the recruited population, and the number of centres that participated, the percentage of male, the mean age of the participants, and duration of the study. If a paper described empathy, authors recorded how empathy was defined. They also recorded the description of the intervention in the experimental group, including content and whether it was experiential or not, frequency, and duration as well as the intervention—if any—in the comparator group. During the experiential learning, trainees are involved in the learning process through experience. It is learning by doing and it is distinct from didactic learning.

In addition, authors reported the outcomes as described in each paper as well as the assessors and measures used by the investigators; and whether empathy was the primary endpoint. They recorded all primary and secondary outcomes in articles that assessed interventions for specifically promoting empathy. Based on the outcome categories provided by the MERSQI tool for assessing the quality of medical education studies [25, 26] authors grouped reported outcomes in the four following types: (a) satisfaction /attitude /perceptions /opinions; (b) knowledge /skills; (c) behavior; and (d) patients /health care outcome. For trials that assessed interventions aiming at communication or interpersonal skills, authors only recorded empathy regardless of whether it was included as primary or secondary outcome. If an instrument was used to measure any of the outcomes, they recorded whether the article reported construct or content validity of the scale. Finally, authors captured the number of participants who were analysed for each measure. They also extracted the difference and the reported measure of dispersion, both for within group and between group comparisons, and the corresponding *P*-values. If a study reported multiple follow up points, authors recorded these values for each point separately.

Quality assessment of the studies

To assess the quality of reporting of the eligible RCTs, authors used the CONSORT statement [27] for reporting randomized controlled trials. Specifically, for each trial, they reported whether it described the mode of randomization, allocation concealment, blinding, and if yes, who were blinded, power calculations, the primary outcome, and the percentage of withdrawals. In addition, they recorded whether trials described the trainers—if any—for the interventions, and whether fidelity was evaluated for the intervention in each study. Authors also recorded whether potential adverse events of the intervention were reported.

Analyses /synthesis

To address the potential effectiveness of interventions, authors presented the results of the studies per outcome category. In case a primary study described results on empathy using multiple measures, authors calculated the combined estimate of empathy for the study by the inverse of variance fixed effects model (FEM) [28]. To combine effect estimates across studies, standardized mean differences (SMDs) and 95 % confidence intervals (CIs) were calculated from the changes in means (post—pre-intervention) and their standard deviations

(SD). If the post—pre-intervention changes were not reported, post-intervention means were used for the synthesis of the results. Authors performed random effects model (REM) meta-analysis of standardized mean differences (SMD) [28]. Heterogeneity was evaluated with Cochran's Q statistic (statistically significant for $P < 0.10$) and it was quantified with the I [2] metric (low, moderate, large, very large for values of <25, 25–49, 50–74, >75 %, respectively) [29]. Primary analysis included all the studies that provided adequate data to calculate SMD. Sensitivity analysis included also studies that part of their results had to be imputed to facilitate their inclusion in the meta-analysis, i.e., studies that provided median instead of mean values (in these studies median was assumed to equal the mean), and trials that provided the mean value but did not provide the SD (in these trials the missing SD was imputed by the largest SD that was recorded among the studies of the same outcome category). Studies that did not provide results on effect estimate and dispersion were excluded from the analyses. In addition, authors proceeded in subgroup analyses for the following groups (a) type of intervention (experiential vs. non-experiential; (b) training specifically for promoting empathy vs. other; (c) type of assessor (external observer, health professional participants, and patients); and (d) type of process used for empathy evaluation, (simulated interview, actual interview, and questionnaire completion without interview). Analyses were performed in STATA 10.0 (STATA Corp., College Station, TX, USA). *P* values were two tailed.

Results

Eligible studies

Electronic searches yielded 722 unique items. Pubmed included 465, Cochrane Database of Clinical Trials included 52, Scopus included 1, and PsychInfo included 204. Authors excluded 621 items after screening the titles. Additionally they excluded 50 articles after screening the abstracts either because they were not written in English ($n = 7$), or because the intervention was not relevant ($n = 43$). Authors retrieved 51 publications in full text. They excluded 34 articles (one was duplicated; two papers were dissertations; and 31 were not RCTs). Searches of the reference lists of the retrieved articles and of review papers did not reveal additional eligible papers. Thus, the total number of the eligible papers included in our systematic review was 17 (Fig. 1).

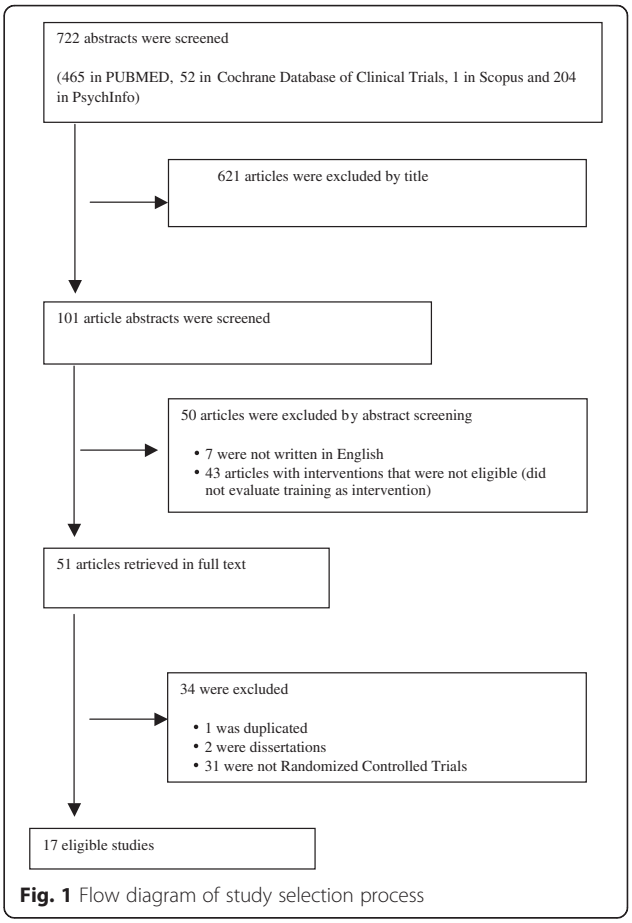

Fig. 1 Flow diagram of study selection process

Characteristics of eligible studies (Table 1)

Eligible studies were published from 1979 to 2012. Thirteen out of the 17 studies were published after 2000 [30–42]. Almost half of them were conducted in USA; [30, 31, 35, 36, 39, 43–45] seven in Europe [32–34, 38, 40–42] and one in Australia [37].. All studies were reported as parallel RCTs except for two [31, 39] with a cross-over design.

Sample size ranged from 16 to 452 participants (median 79; IQR 48–133). Ten studies[30,31,33,34,36,-38,40–42] included health care professionals (physicians, residents, and nurses). Seven studies [32, 35, 39, 43–46] included under-graduate or post-graduate students. There were 7 multicentre studies; [31, 33–35, 39, 40, 42] five studies [32, 43–46] did not provide data on the number of centres. Two studies [34, 42] included only females; three studies [32, 44, 45] gave no information about gender. Male proportion for the rest of the studies ranged between 15 % and 81 % (median 53 %; IQR 21 %-68 %). Mean age of the participants ranged from 21.2 to 49.3 years. Nine studies [30, 32, 34, 40–45] provided no data on age. Duration of the studies varied from one month to 60 months (median 10.5 months; IQR 6–12 months).

Two [31, 36] articles provided definition for empathy. Boncivici [36] referred to empathy as a 'multidimensional concept and skill with cognitive, affective and behavioral components'. Riess [31] defined empathy as 'a process with both cognitive and affective components, which enables individuals to understand and respond to others' emotional states

Table 1 Characteristics of eligible studies

First author, Publication year Country	Sample size	Population/ Number of Centres	Male (%)	Mean age (years)	Study duration (months)
Blaire-Irvine 2012 USA	172	Licensed care staff/ ND	15	ND	1.5
Riess 2012 USA	99	Surgery, Medicine, Anesthesiology, Psychiatry, Ophthalmology, and Orthopedics Residents/ 2	79	30	10
Daeppen 2011 Switzerland	131	2nd year master students/ 1	ND	ND	48
Tulsky 2011 UK	48	Oncologists/ 2	81	49.3	1
Rask 2009 Denmark	24	Nurses in oncology clinic/ 1	0	9 nurses <40 14 nurses >40	7
Shapiro 2009 Canada	79	1st year medical students/ ND	39	23.2	48
Bonvicini 2008 USA	155	Physicians/ ND	63	37.3	11
Butow 2008 Australia	30	Oncologists/ 6	50	Range 41–44	12
Delvaux 2005 Belgium	62	Specialists in medical or surgical oncology, radiotherapy, hematology, and gynecology/ ND	55	43	6
Shapiro 2004 USA	16	Medical students/ 1	45	23.4	12
Fallowfield 2002 UK	160	Oncologists/ 34	72	ND	4
Jenkins 2002 UK	93	Oncologists/ ND	70	ND	60
Razavi 2002 Belgium	133	Nurses/ 88	0	ND	6
Smith 1995 USA	26	1st year residents/ 2	57	ND	ND
Daniels 1988 Canada	53	2nd year registered nurses students/ ND	0	21.2	9
Wolf 1987 USA	134	1st year medical students/ 1	ND	ND	12
Robbins 1979 USA	51	Residents/ 2	ND	ND	12

ND, No data

and contributes to compassionate behavior and moral agency'.

Type of interventions (Table 2)

Fifteen studies described experiential interventions [31, 32, 34–46]. Two [31, 39] out the 15 studies included interventions for specifically promoting empathy while the rest assessed interventions on communication or interpersonal skills. Experiential interventions included role-playing, self-awareness exercises, and feedback as well as group discussions where the participants are responsible for their own learning. Two studies [30, 33]

Table 2 Type of interventions in eligible studies. Interventions are grouped as experiential and non-experiential

First author, publication year	Intervention	Frequency	Duration of intervention	Comparison group
Experiential Interventions				
Riess 2012	Empathy training: Videos of clinical interactions	Three 60-min modules	1 month	Standard post graduate medical education
Daeppen 2011	Motivational interviewing training: Sessions of practical exercises and role playing, Learning objectives, training activities, including exercises, DVD extracts and role-plays	Two 4-h sessions	2 days	No training
Rask 2009	Communication skills training: Program including role playing, theoretical presentations, video demonstrations, group discussions	Two 2-day sessions lasting 33 h	2 months	No training
Saphiro 2009	Communication skills training: Therapeutic communication elective program: including regular meetings on one to one basis with patients under psychiatrist's supervision	Weekly meetings	4 months	No training
Boncivici 2008	Communication skills training: Audiotaped physician—patient interaction, didactic and experiential teaching modalities, coaching, practice sessions	3 times, 6 h each (18 h)	3 months	No training
Butow 2008	Communication skills training: Intensive face-to-face workshop	1.5 intensive and four 1.5 h monthly	6 months	No training
Delvaux 2005	Communication skills training: Theoretical information, role play, how to handle patients' distress	Six 3-h consolidation workshops	3 months	19-h basic training program
Shapiro 2004	Empathy training: Reading and discussion sessions	1 h twice monthly	4 months	No training
Fallowfield 2002	Communication skills training: Learner-centered course incorporating cognitive, behavioral and experiential components, written feedback	ND	3 days	Written feedback
Jenkins 2002	Communication skills training: Courses were learner centered, incorporating cognitive, experiential, and behavioral components	ND	3 days	No training
Razavi 2002	Communication skills training: Theoretical information, case presentations, role play	5 days per week	3 weeks	No training
Smith 1995	Psychosocial Training: Intensive program including interviewing, somatization, patient education, self-awareness	3 times per week	1 month	No training
Daniels 1988	Microtraining skills: Videotapes, microtraining/ microcounseling	6 segments 3 to 5 h length (total 25 h)	ND	No training
Wolf 1987	Training to facilitate response to patients' concerns: Series of videotapes/ video program (interpreting body language in everyday practice)	6-h lectures, 4 weekly 3-h sessions	1 month	6-h lectures
Robbins 1979	Interpersonal skills: Videotapes, discussions and sensitivity exercise	8-day course, 1–2 h every morning (30 h)	8 days	No training
Non-experiential Interventions				
Blaire-Irvine 2012	Mental illness training: Internet Program: "Caring Skills: Working with Mental Illness" including video modeling vignettes, right way and wrong way exemplars, testimonials, narrations for creating empathy	2 visits, 1 week apart - 2 courses each visit (10–30 min length each course)	2 weeks	No training
Tulsky 2011	Training to facilitate response to patients' concerns: 1-h lecture plus CD-ROM training program (with 5 modules: effective communication, empathic opportunities, responding)	ND	1 month	1-h lecture

ND No Data

described non-experiential interventions including items such as audio-taped interactions between physicians and patients, CD-ROMs or Internet courses, where the participants had no active presence during learning process.

Frequency of intervention varied between studies (Table 2). Three papers—two referred to experiential and one to non-experiential interventions—did not specify the intensity of the interventions [33, 40, 41]. Experiential interventions lasted from 8 h to 6 months while non-experiential from 2 weeks to 3 months. In four studies [30–32, 39] – two [31, 39] of which included interventions for specifically promoting empathy - there were experiential interventions with duration less than 2 weeks. Eleven studies [31, 32, 34, 35, 37, 39, 41–43, 45, 46] with experiential interventions, and two [30, 36] with non-experiential interventions reported no training for the comparison group. Three [30, 36, 38] of the remaining four studies provided a brief intervention. In one study [40], investigators gave written feedback to the control group.

Type of outcomes (Table 3)

All studies evaluated change in physicians' empathic behavior as an outcome. Riess [27] was the only study that also evaluated change in attitude towards empathy, knowledge, and skills using self-reported questionnaires.

To evaluate change in health professionals' behavior, studies used interviews with simulated patients [32, 37], or actual interviews [33, 36, 40, 41, 45], or both [38, 42]. All these studies used external assessors; Tulsky [33] also used the patients as assessors for the actual interviews. There were six studies that used questionnaires completed by the health professionals to evaluate change in their behavior [30, 31, 34, 44–46]. Shapiro [35] used an external observer who filled in the questionnaire. In three studies [31, 34, 43], patients completed the questionnaire.

Effectiveness of interventions (Table 4)

Riess [31] did not find significant change in residents' attitude towards empathy; however, this study showed a significant increase in physicians' knowledge of the neurobiology and physiology of empathy ($P < 0.001$) as well as in physicians' skill at decoding subtle facial expressions of emotion ($P < 0.001$).

Out of four studies that used simulated interviews to assess change in empathic behavior, there was only one [42] that reported a significant improvement in the use of emotional "distress" words by nurses ($P < 0.001$ for frequency score and $P = 0.04$ for density score). However, the use of emotional "distress" words was not increased in the same study when investigators evaluated actual interviews. Out of the seven studies that used actual interviews, there were four trials that reported significant improvement in empathic behavior for health professionals. Specifically, Tulsky [33] showed a significant increase in the number of empathic statements per conversation ($P = 0.024$), and in the number of continuer response to empathic opportunities ($P = 0.028$) for oncologists. Bonvicini [36] reported significant increase in physicians' empathic expression using Global Rating Score ($P < 0.01$), and Hierarchical Empathy Communication Coding System ($P < 0.01$). Fallowfield [40] supported improvement in the number of empathic expressions ($P = 0.005$) for oncologists while Razavi [42] concluded that nurses increased the use of emotional "anxiety" words ($P = 0.028$) in actual interviews.

Out of the nine studies that used questionnaires completed by the health professionals to evaluate change in their behavior, three trials reported a significant improvement for the intervention group. Specifically, Blair-Irvine [30] showed a significant increase in the psychosocial construct of empathy Likert scale ($P = 0.04$); Daniels [44] supported a significant increase in Carkhuff index of communication ($P < 0.05$) for nurses; while in Wolf [44], medical students exhibited greater ability to respond to patients' emotional concerns in hypothetical scenarios using Medical Communication Index ($P < 0.001$), and greater preferences for responses that addressed patients' emotions using Helping Relationship Inventory ($P < 0.001$). In addition, external observers who completed Staff-Patient Interaction Rating Scale in one trial [35], found an increase in expressed empathy for the medical students ($P = 0.04$). Finally, out of the two studies that patients filled in the questionnaires, one [31] reported significant increase in residents' empathy using Consultation and Relational Empathy Measure ($P = 0.04$).

Three studies [39, 40, 43] did not provide numerical data on results; and therefore, they were excluded from further analyses. Delvaux [38] reported results for several measures of empathic behavior as relative risk (RR), and therefore this trial was also excluded from standardized mean difference (SMD) meta-analyses. It reported that intervention was not effective in improving empathic behavior (RR 1.5, 95 % CI 1.0, 2.3; P value 0.6). In Table 4, the results of meta-analyses per outcome category as well as per subgroup were presented (additional information on the corresponding forest plots is available in Appendix). Ten studies [30–36, 42, 44, 46] provided adequate results and were combined showing that interventions were effective in improving empathic behavior (SMD 0.7, 95 % CI 0.3, 1.1; P value <0.001) (Table 4; see also Appendix). After inclusion of three [37, 41, 45] additional studies for which part of their results had to be imputed, the interventions remained effective for improving empathy (SMD 0.8, 95 % CI 0.4, 1.2; P value <0.001) (Table 4; see also Appendix). Similar

Table 3 Results of eligible studies. Results are presented per outcome category, i.e., attitude, knowledge / skills, and behavior

First author, publication year	Rater/ Process	Outcome: measure	Participants analyzed (intervention / control	Difference within group		Difference between groups	
				Intervention Estimate[a] (SD or 95 % CI)	Comparison group Estimate[a] (SD or 95 % CI)	Estimate[a] (SD or 95 % CI)	P-value
Type of outcome: Attitude							
Riess 2012[b]	Health professional/ questionnaire only	Attitude towards empathy: Jefferson Scale of Physician Empathy	99 (54 / 45)	1.2 (SD 9.3)	−1.1 (SD 6.7)	Effect size 0.3	NS
Type of outcome: Knowledge / Skills							
Riess 2012[b]	Health professional/ questionnaire only	Physician's knowledge of the neurobiology and physiology of empathy: Neurobiology and Physiology of Empathy Test	99 (54 / 45)	2.3 (SD 2.4)	0.4 (SD 2.3)	Effect size 0.8	<0.001
Riess 2012[b]	Health professional/ questionnaire only	Physician's skill at decoding subtle facial expressions of emotion: Ekman test	99 (54 / 45)	2.1 (SD 2.5)	0.2 (SD 2.2)	Effect size 0.8	<0.001
Type of outcome: Behavior							
Daeppen 2012[c]	External observer/ simulated interview	MITI 3.0 behavioral coding system Global Scores: Empathy	131 (66 / 65)	Before: NA After: 4.0 (SD 0.6)	Before: NA After: 3.4 (SD 0.7)	NA	NA
Butow 2008[d]	External observer/ simulated interview	Key doctor behaviors; category: creating an environment where emotion is likely to be expressed: Basic empathy subscale	30 (15 / 14)	Before: Median 5.0 (IQR 4.0, 5.0) After: Median 5.0 (IQR 4.5, 5.0)	Before: Median 4.0 (IQR 3.8, 5.0) After: Median 4.0 (IQR 3.0, 5.0)	ND	NS
Delvaux 2005[e]	External observer/ simulated interview	Empathy, educated guesses, alerting to reality, confronting, negotiating, summarizing: CRCWEM rated utterances directed to patients	56 (27 / 29)	ND	ND	RR 1.8 (95 % CI 1.0, 3.4)	NS
Delvaux 2005[e]	External observer/ simulated interview	Empathy, educated guesses, alerting to reality, confronting, negotiating, summarizing: CRCWEM rated utterances directed to relatives	56 (27 / 29)	ND	ND	RR 1.1 (95 % CI 0.4, 3.2)	NS
Razavi 2002	External observer/ simulated interview	Emotional depth of utterances: CRCWEM neutral (level 0) utterances	115 (57 / 58)	ND	ND	ND	NS
Razavi 2002[f]	External observer/ simulated interview	Use of emotional "distress" words: Frequency score	115 (57 / 58)	Before: 9.2 (SD 4.9) After: 12.9 (SD 5.6)	Before 10.8 (SD 5.6) After: 9.1 (SD 5.4)	ND	<0.001
Razavi 2002[f]	External observer/ simulated interview	Use of emotional "distress" words: Density score	115 (57 / 58)	Before: 3.6 (SD 2.1) After: 4.4 (SD 2.2)	Before: 4.0 (SD 2.8) After: 3.6 (SD 2.7)	ND	0.04
Tulsky 2011[g]	External observer/ actual interview	Emotion handling skills: Number of empathic statements per conversation	48 (24 / 24)	Before: 0.4 (SD 1.0) After: 0.8 (SD 1.3)	Before: 0.3 (SD 0.7) After: 0.4 (0.8)	RR 1.9 (95 % CI 1.1, 3.3)	0.024
Tulsky	External observer/ actual interview		48 (24 / 24)	ND	ND		0.028

Table 3 Results of eligible studies. Results are presented per outcome category, i.e., attitude, knowledge / skills, and behavior *(Continued)*

2011[g]		Emotion handling skills: Number of continuer response to empathic opportunities				OR 2.1 (95 % CI 1.1, 4.2)	
Bonvicini 2008	External observer/ actual interview	Physicians' empathic expression: GRS, Empathy	155 (79 / 76)	Before: 8.4 (95 % CI (7.8, 9.0) After: 11.6 (95 % CI 11.0 12.2)	Before: 8.1 (95 % CI (7.6, 8.6) After: 7.4 (95 % CI 7.0 7.8)	1.4 (ND)	<0.01
Bonvicini 2008	External observer/ actual interview	Physicians' empathic expression: Hierarchical ECCS	155 (79 / 76)	Before: 2.7 (95 % CI (2.5, 2.8) After: 4.0 (95 % CI 3.9 4.1)	Before: 2.6 (95 % CI 2.4, 2.8) After: 2.5 (95 % CI 2.3 2.7)	3.9 (ND)	<0.01
Delvaux 2005[e]	External observer/ actual interview	Empathy, educated guesses, alerting to reality, confronting, negotiating, summarizing: CRCWEM rated utterances directed to patients	56 (27 / 29)	ND	ND	RR 1.4 (95 % CI 0.7, 2.8)	NS
Fallowfield 2002[h]	External observer/ actual interview	Empathy: Number of empathic expressions	160 (80 / 80)	ND	ND	RR 1.5 (ND)	0.005
Jenkins 2002[i]	External observer/ actual interview	Empathy: Presence of empathic expressions	93 (48 / 45) 186 (97/89)	Before: 57 59 % (ND) After: 56 58 % (ND)	Before: 44 49 % (ND) After: 38 42 % (ND)	ND	ND
Razavi 2002	External observer/ actual interview	Emotional depth of utterances: CRCWEM neutral (level 0) utterances	115 (57 / 58)	ND	ND	ND	NS
Razavi 2002[f]	External observer/ actual interview	Use of emotional "distress" words: Frequency score	115 (57 / 58)	Before: 4.4 (SD 3.8) After: 7.0 (SD 5.8)	Before 4.6 (SD 4.9) After: 4.3 (SD 4.2)	ND	NS
Razavi 2002[f]	External observer/ actual interview	Use of emotional "distress" words: Density score	115 (57 / 58)	Before: 3.7 (SD 3.4) After: 2.7 (SD 4.8)	Before: 3.3 (SD 3.0) After: 3.1 (SD 2.9)	ND	NS
Razavi 2002[j]	External observer/ actual interview	Use of emotional "anxiety" words: Density score	115 (57 / 58)	ND	ND	ND	0.028
Robbins 1979[k]	External observer/ actual interview	Empathy: Carkhuff scale - Empathy level	51 (26 / 25)	Before: 2.3 (ND) After: 2.7 (ND)	Before: 2.3 (ND) After: 2.3 (ND)	ND	ND
Robbins 1979[k]	External observer/ actual interview	Empathy: Carkhuff scale - Number of empathic responses	51 (26 / 25)	Before: 2.0 (ND) After: 4.1 (ND)	Before: 2.3 (ND) After: 2.3 (ND)	ND	ND
Tulsky 2011[l]	Patient/ actual interview	Empathy: Perceived Empathy Scale	48 (24 / 24)	Before: NA After: 0.4 (95 % CI 0.3, 0.5)	Before: NA After: 0.2 (95 % CI 0.1, 0.3)	NA	NA
Shapiro 2009	External observer/ questionnaire only	Expressed empathy: SPIR Scale	79 (38 / 41)	Before: 6.0 (SD 5.7) After: 8.3 (SD 5.0)	Before: 7.3 (SD 6.7) After: 6.6 (SD 5.1)	ND	0.04
Blair-Irvine 2012[m]	Health professional/ questionnaire only	Assessment of psychosocial construct: Empathy (4-item 7-point Likert scale)	172 (84 / 88)	Before: 5.1 (SD 1.0) After: 5.5 (SD 0.9)	Before: 5.7 (SD 1.1) After: 5.2 (SD 1.0)	ND	0.04
Riess 2012[n]	Health professional/ questionnaire only	Empathic responsiveness in personal life: BEES	99 (54 / 45)	0.9 (SD 14.5)	2.7 (SD 14.1)	Effect size 0.12	NS
Rask 2009[o]	Health professional/ questionnaire only	Nurse's perception of patient's experiencing empathy during	23 (12 / 11)	Before: 21.4 (SD 2.4) After: 24.0 (SD 2.9)	Before: 22.1 (SD 3.6) After: 23.2 (SD 3.2)	Effect size 0.42	NS

Table 3 Results of eligible studies. Results are presented per outcome category, i.e., attitude, knowledge / skills, and behavior *(Continued)*

		their communication: NPRI, Empathy subscale					
Shapiro 2004[p]	Health professional/ questionnaire only	Empathy: ECRS	16 (10 / 6)	ND	ND	ND	ND
Shapiro 2004[p]	Health professional/ questionnaire only	Empathy: BEES	16 (10 / 6)	ND	ND	ND	ND
Daniels 1988[q]	Health professional/ questionnaire only	Empathy: Carkhuff index of communication	53 (24 / 29)	Before: 1.7 (SD 0.3) After: 2.4 (SD 0.2)	Before: 1.8 (SD 0.4) After: 1.9 (SD 0.3)	ND	<0.05
Daniels 1988[q]	Health professional/ questionnaire only	Empathy: ECRS	53 (24 / 29)	Before: NA After: 188.8 (42.1)	Before: NA After: 151.9 (40.9)	ND	NA
Wolf 1987	Health professional/ questionnaire only	Ability to respond to patients' emotional concerns in hypothetical scenarios, MCI	134 (65 / 69)	Before: 0.7 (SD 0.7) After: 2.0 (SD 1.0)	Before: 0.9 (SD 0.9) After: 1.1 (SD 0.8)	ND	<0.001
Wolf 1987	Health professional/ questionnaire only	Exhibit greater preferences for responses that addressed patients' emotions: HRI	134 (65 / 69)	Before: 33.0 (SD 7.9) After: 17.9 (SD 6.7)	Before33.1 (SD 9.5) After: 22.3 (SD 8.8)	ND	<0.001
Robbins 1979[k]	Health professional/ questionnaire only	Affective sensitivity - empathy: Affect Sensitivity Scale	51 (26 / 25)	Before: 26.2 (ND) After: 28.8 (ND)	Before: 26.5 (ND) After: 28.0 (ND)	ND	NS
Riess 2012[n]	Patient/ questionnaire only	Physician's empathy: CARE	99 (54 / 45)	0.7 (SD 7.9)	−1.5 (SD 6.0)	Effect size 0.31	0.04
Rask 2009[o]	Patient/ questionnaire only	Patient perception of nurse empathy: PPRI, Empathy subscale	23 (12 / 10)	Before: 22.8 (SD 1.9) After: 22.8 (SD 1.8)	Before: 22.4 (SD 1.7) After: 23.5 (SD 2.9)	Effect size 0.05	NS
Smith 1995[r]	Patient/ questionnaire only	Patient Satisfaction Questionnaire, Patient perception of physician's empathy	26 (14 / 12)	ND	ND	ND	NS

SD Standard Deviation, *CI* Confidence Interval, *ND* No Data, *NS* Non Significant, *IQR* Interquartile Range, *RR* Risk Ratio, *NA* Non Applicable, *MITI* Motivational Interviewing Treatment Integrity, *SPIR* Staff-Patient Interaction Rating Scale, *ECRS* Empathy Construct Rating Scale, *BEES* Balanced Emotional Empathy Scale, *GRS* Global Rating Score, *ECCS* Empathy Communication Coding System, *CARE* Consultation and Relational Empathy Measure, *CRCWEM* Cancer Research Campaign Workshop Evaluation Manual, *NPRI*, Nurse—Patient Relationship Inventory, *PPRI* Physician—Patient Relationship Inventory, *MCI* Medical Communication Index, *HRI* Helping Relationship Inventory

[a] Estimate: mean unless otherwise indicated

[b] Numbers for estimate represent pre-post change

[c] There was not pre-intervention assessment of evaluation; authors reported that higher post-intervention scores on empathy were found in the experimental group (effect size 0.7, 95 % CI 0.6–0.8, $P < 0.001$)

d Results correspond at baseline and 6 month follow up values; difference between groups was also not significant after 12 month follow up

[e] Estimated relative risks were based on a multivariable Poisson regression model was adjusted for physician age, oncology practice, and gender in both types of interviews; for patients' and their relatives' age, gender, and educational level; for patients' Karnofsky performance status; for relatives' ties with the patients; for the number of months since diagnosis; for the presence of a previous and/or current cancer treatment; for the type of information (diagnosis or prognosis-focused and treatment-focused) and of the news given (good, bad, or neutral); and for the fact that it was or not the first interview of the

patient with the physician. There was no convergence of the model for relatives in actual interviews

[f] Frequency score and density score were calculated by PROTocol ANalyzer (PROTAN); comparisons with multivariate analyses of variances (MANOVA) included pre-intervention scores, post-intervention scores, and scores at 3 month follow up (non-significant difference); "distress" words as tagged by the French version of the Harvard Third Psychosociological Dictionary

[g] Analysis for the number of empathic statements was based on total conversations ($n = 264$) while for the number of continuer response to empathic opportunities on conversations with at least one empathic opportunity ($n = 135$) including 275 empathic opportunities (range: 1 to 11 opportunities per conversation). Results for the number of empathic statements were from a mixed-effect Poisson regression model adjusted for study site; physician sex; and mean number of empathic statements, defined as NURSE (name, understand, respect, support, explore) statements used by the physician in the pre-intervention phase. Results for the number of continuer response to empathic opportunities were from a logistic mixed-effect regression model adjusted for study site and physician sex

[h] Number of empathic expressions was assessed after videotapes including consultations were rated by Medical Interaction Process System (MIPS)

[i] Presence or absence of empathic expressions was based on Medical Interaction Process System (MIPS) summary data from the videotaped consultations between the physicians and patients; authors presented the results of comparison between the groups only after intervention stating that the intervention group were more likely to exhibit empathy (56 [58 %] vs. 38 [42 %]; $\chi^2 = 4.823$, df =1, $P = 0.02$)

[j] "Anxiety" words as tagged by the French version of the Regressive Imagery Dictionary (MRID); authors stated
that in actual interviews, trained health care professionals used different words (density scores) tagged by 'anxiety' MRID subcategory compared to untrained health care professionals (group by time effect MANOVA F value: 3.66, $P = 0.028$). Authors do not report the results on frequency or density scores of any other MRID emotion subcategories assessed in the study for health care professionals

[k] Authors presented only comparisons within but not between groups stating that the mean empathy level scores increased significantly in the experimental group ($P < 0.05$ by F tests of group means) while control scores did not. In addition, the number of responses dealing with patient feelings increased in the experimental group ($P < 0.05$ by F tests of group means) but not in the control group

[l] Perceived empathy was not evaluated before intervention. Two hundred two patients (109 in the intervention group and 93 in the control group) were included in the analysis for perceived empathy. Mixed-effect models are adjusted for site and physician sex. Marginal standardized estimates are predicted proportion for binary outcomes, and the 95 % CIs for the standardized estimates and relative risks are from 1000 bootstrap samples. There was no statistically significant difference between the two groups ($P = 0.058$)

[m] Comparison of the posttest scores, adjusted for pretest scores, showed greater gains by the intervention condition compared to the control condition for empathy. Results from the ANCOVA analysis adjusted for pretest scores did not show significant group differences at 8-week follow-up

[n] Numbers represent pre-post change

[o] Group x time interaction effects were investigated by within-between-subjects ANOVAs. For nurses, Group x time interaction effects were also investigated at 3 month follow up without statistically significant between group differences ($P = 0.94$)

[p] Authors presented analyses after combining the two randomized groups to one; results per group and on the comparison between groups were not provided

[q] Significant difference between groups was not confirmed after 9 month follow up; authors reported that after the intervention experimental trainees performed better in ECRS ($P < 0.01$); however, pre-post intervention comparison between groups is not applicable

[r] Mean post-intervention scores were given separately for male and female residents while pre-intervention scores were not provided. Analysis of covariance with groups (trained vs. untrained) and pre-training satisfaction score as a covariate did not yield statistically significant results for patients' perceived physician empathy ($P = 0.65$)

Table 4 Overall results for empathy promoting interventions vs. control among health professionals on empathic attitude, knowledge /skills, and on empathic behavior. Results are also presented separately after the inclusion of trials with imputed values. Separate subgroup analyses are shown for experiential and non-experiential interventions, for training specifically for empathy and for training on interpersonal /communication skills, for the type of assessor (external observer, health professional, and patient), and for the type of evaluation process (simulated interview, actual interview, and questionnaire)

Outcome type (subgroup)	Number of eligible studies	SMD (95 % CI)	*P*-value	Q-statistic I [2] (95 % CI)
Attitude (all)	1 [27]	0.4 (−0.2, 0.1)	0.95	NA
Knowledge /skills (all)	1 [27]	0.8 (0.4, 1.2)	<0.001	NA
Behavior (all)	10 [26–32, 38, 40, 42]	0.7 (0.3, 1.1)	<0.001	76.27 88 % (80 %, 93 %)
Behavior (including studies with imputed values)	13 [26–33, 37, 38, 40–42]	0.8 (0.4, 1.2)	<0.001	132.22 91 % (86 %, 94 %)
Behavior (experiential intervention)	11 [27, 28, 30–33, 37, 38, 40–42]	0.8 (0.3, 1.3)	<0.001	130.44 92 % (88 %, 95 %)
Behavior (non-experiential intervention)	2 [26, 29]	0.7 (0.3, 1,2)	0.001	1.6 NA
Behavior (training specifically for empathy)	1 [27]	0.1 (−0.3, 0.4)	0.15	NA
Behavior (training for communication skills)	12 [26, 28–33, 37, 38, 40–42]	0.9 (0.5, 1.3)	<0.001	111.47 90 % (85 %, 94 %)
Behavior (external observer as assessor)	8 [28, 29, 31–33, 37, 38, 41]	1 (0.4, 1.5)	0.001	84.29 92 % (86 %, 95 %)
Behavior (health professional participant as assessor)	6 [26, 27, 30, 40–42]	0.7 (0.2, 1.3)	0.008	50.15 90 % (81 %, 95 %)
Behavior (patient as assessor)	3 [27, 29, 30]	0.4 (−0.2, 1.2)	0.22	3.27 39 % (0, 81 %)
Behavior (simulated interview as process of evaluation)	3 [28, 33, 38]	0.8 (0.5, 1.1)	<0.001	2.53 21 % (0, 92 %)
Behavior (actual interview as process of evaluation)	5 [29, 32, 37, 38, 41]	1.2 (0.3, 2.1)	0.01	81.41 95 % (91 % 97 %)
Behavior (questionnaire as process of evaluation)	7 [26, 27, 30, 31, 40–42]	0.8 (0.3, 1.3)	0.003	42.31 86 % (73 %, 93 %)

SMD Standardized Mean Difference, *CI* Confidence Interval, *NA*, Non-Applicable

results were supported by subgroup analyses for experiential interventions, non-experiential interventions, and for training health professionals on interpersonal /communication skills programs (Table 4; see also Appendix). However, training on programs that specifically aimed at empathy was assessed by one study and did not yield a significant result. In addition, significant improvement for empathy was noted when an external observer or the health professional that participated was used as assessor but not when the patient assessed the health professional (Table 4; see also Appendix). Finally, interventions were found effective regardless of the evaluation process that was used (Table 4; see also Appendix).

Quality of reporting for eligible studies

Included studies were of fair to moderate quality. Randomization mode and allocation concealment were not reported in any of the studies. Eight of the seventeen studies [31–33, 37, 39, 41, 44, 45] reported single, while three reported double [31, 32, 38] blinding. Three studies [31, 35, 36] used power calculations to determine the sample size and they reached the number of requested participants. However, none of the studies reported that they incorporated potential contamination between the intervention and the control group in their power analysis.

Out of the 15 trials that employed trainers for the intervention, seven [34, 35, 41, 42, 45] described the number and the experience of the trainers. Two trials [30, 33] provided internet-based or computerized interventions without the need for trainers. Trials invariably did not report on whether researchers verified the accurate implementation of interventions. Potential adverse effects of the interventions were not reported in the studies.

All studies clearly stated primary outcomes; eight trials included empathy as primary outcome [30, 31, 33, 36, 39, 42, 43, 46]. Out of the 13 studies that used a scale to assess their outcomes, eleven [30–32,

34–36, 38, 39, 43, 44, 46] reported the psychometric characteristics of the scales.

Nine studies [33–36, 38, 40–43] reported withdrawal rate, which ranged from 0 to 23 %. Jenkins [41] was the only trial that reported a withdrawal rate more than 20 %, i.e., 23 %.

Conclusion

Several trials assessed the change in empathic behavior among health professionals supporting a significant improvement; however, very few were well powered and assessed empathy as primary outcome. Interventions usually covered a broad training on communication skills based on experiential approach. There was a limited number of RCTs that evaluated interventions specifically aimed at promoting empathy; none of them evaluated patients' outcomes.

Clinical trials that assessed changes in empathic behavior were highly heterogeneous in terms of participants, interventions and outcome measures. Interventions varied in terms of content and frequency; their duration was generally brief and did not exceed six months. Follow up did not exceed one year in any of the studies. Studies with significant results at the end of the intervention did not observe the improvement at a later stage when they followed participants for a longer period. Thus, it would be difficult to support even for potential effective interventions that changes in empathic behavior may last. Investigators applied a wide variety of modes, and measures to assess changes in empathic behavior among health professionals. This may reflect the complexity of measuring any change in behavior [47]. In this meta-analysis, patients did not perceive improvement on empathic behavior for health professionals. This may reveal the necessity for a validated, globally accepted process to assess empathy in research taking into consideration the assessment both from health professionals and patients. It may also reveal the lack of agreement on the definition of empathy in clinical practice [48]. In this review, authors found only two publications [31, 36] that provided a definition for empathy. For the purpose of this study, authors considered the definition of empathy within the framework of the Person-Centered Approach as the ability to deeper understand other's frame of reference and involves being able to put oneself in the other's position [49].

A previous systematic review included training in empathy as a part of generalized communication skills training [50], and revealed improvement in outcomes such as trust in physicians, or patient's satisfaction. However, the specific role for empathy as part of communication skills to achieve these results remained unknown. Another systematic review [51] referred to communication skills training for health care professionals working with cancer patients and their families. It identified three trials, one of which showed significant increase in the expression of empathy. In our review, several RCTs showed that training in communication skills might enhance empathic behavior. Effective communication skills training incorporated a number of brief experiential interventions including role-playing, self-awareness exercises, and feedback as well as group discussions; or brief non-experiential interventions including items such as audio-taped interactions between physicians and patients, CD-ROMs or Internet courses. However, most of these studies did not report power calculations. Therefore, their results need to be cautiously interpreted. Moreover, future studies are necessary to corroborate these results. In addition, there was a paucity of RCTs for interventions specifically aiming at empathy training. Whether such interventions may have an impact on patients' health care outcomes as well needs to be investigated in the future. Effective interventions specifically aiming at enhancing empathy may be incorporated in general training programs. However, further research is needed in order to clarify the type of approach, the duration, and the frequency of empathy enhancing interventions within a generic program. There is still debate about whether it is feasible and sound to isolate empathy from general training or empathy has to be taught in the context of a communication skills training. In the present meta-analysis, interventions aimed specifically at empathy were not found effective; however, the number of the included trials was small. Until there is a definitive answer both approaches may complement each other. Thus, medical curricula may provide empathy training within the targeted training in each clinic, such as cardiology, internal medicine or orthopedics. Taking into consideration that empathy is a way of being; there is indeed no need to separate it from the general training. In addition specific communication skills training may provide students or health care professionals with the appropriate process and time in order to further enhance awareness about empathy.

The methodological quality of eligible studies in our review was compromised. They invariably failed to report important items for RCT design and several times did not provide details on interventions. This may increase the difficulty for these interventions for replication in future studies. However, it was encouraging that several recent studies showed

better reporting. The current development of tools for reporting RCTs of behavior change and health education interventions [52] may explain this observation. Almost all studies that included a process with simulated or actual interview used an external rater, which may have attributed a higher objectivity in the measurements [25, 26]. Studies with questionnaires requested health professionals to evaluate their own performance. Two trials used multiple types of assessors, i.e., an external observer and a patient [33], or a health professional and a patient [31], which did not yield concordant results within each study [31, 33].

The present systematic review confirmed that empathy is an attribute that is amenable to change as a result of educational experiences [1]. Counteracting current trends in medical education and practice that are not conducive to empathic engagement in patient care requires a mandate for the development and implementation of targeted educational programs at all levels of training in all academic medical centers [1]. In order to develop an educational program for promoting empathy investigators need to take into consideration that the extent to which the potential for empathy can be actualized or enhanced in a particular person depends on the interaction of several factors, including the person's constitutional makeup, early life experiences, motivation, and a facilitating environment [1]. The content of an educational intervention needs to address the variety of clinical contexts in which empathy may be communicated as well as the variety of verbal and non-verbal ways in which empathy may be communicated. This process requests for a methodological approach beyond the design of an RCT. It requires the contribution of other research fields, such as communication theory, planned behavior theory, and behavioral science. For example, an educational program for promoting empathy may be based on scientifically validated theories of behavior, such as the social cognitive theory, the theory of planned behavior, operant conditioning, implementation intentions, or stage theories. This will facilitate the theoretical understanding of the likely process of change among health professionals. However, once an educational program is developed its rigorous evaluation through RCTs is necessary. These RCTs demand the explicit reporting of additional design characteristics as compared to RCTs for pharmacological interventions. Empathy enhancing interventions may be complex interventions including multiple components that may be tailored to individual participants. Thus, these interventions need a detailed reporting including the procedures for tailoring them to individual participants, the mode through which they were standardized, the process through which the adherence of the providers with the protocol was assessed. In addition they need to report how these interventions were finally implemented and whether any blinding process was also attempted. Acknowledging that there would be factors –several of them unknown—that may not have been taken into consideration during the design of an empathy promoting intervention, randomization would be the way to control for these factors during the evaluation process. In addition, there are several barriers to maintain the results of a potentially effective educational intervention including time restriction, poor reimbursement, and decision-making about access to treatments by another provider. A pragmatic study design may facilitate addressing these issues; or RCTs may evaluate interventions specifically developed to confront these barriers and provide solutions for sustainability of empathic behavior among health professionals.

This review had several limitations. First, authors may have failed to identify additional studies in search strategy since certain investigators may have used different phrases or words to describe empathy. However, there were extensive searches in multiple electronic databases as well as hand searches of the retrieved articles to retrieve all RCTs that included the word "empathy". Included studies used different instruments to measure the same construct, and therefore SMD was used in meta-analysis for combining continuous data. The use of SMD was helpful in generalizing the results; however, interpretability may be limited. Authors also provided a narrative synthesis of the results per outcome category to facilitate interpretation. Finally, authors cannot exclude the possibility of a publication bias.

As a conclusion, limited evidence suggested that certain interventions might effectively enhance empathic behavior among physicians, residents, nurses, or medical students for a brief period of time. However, the exact type of intervention that would be effective needs to be clarified. In addition, whether any improvement in health professionals' empathic behavior may continue to be present after a longer period, or whether it may also affect patients' outcomes is yet to be defined. Future research needs to clarify factors, which may contribute to the enhancement of empathy in patient care, and support a lasting effect of empathy training. In addition, future research is necessary to measure the impact of empathy promoting interventions on patient outcomes.

Appendix

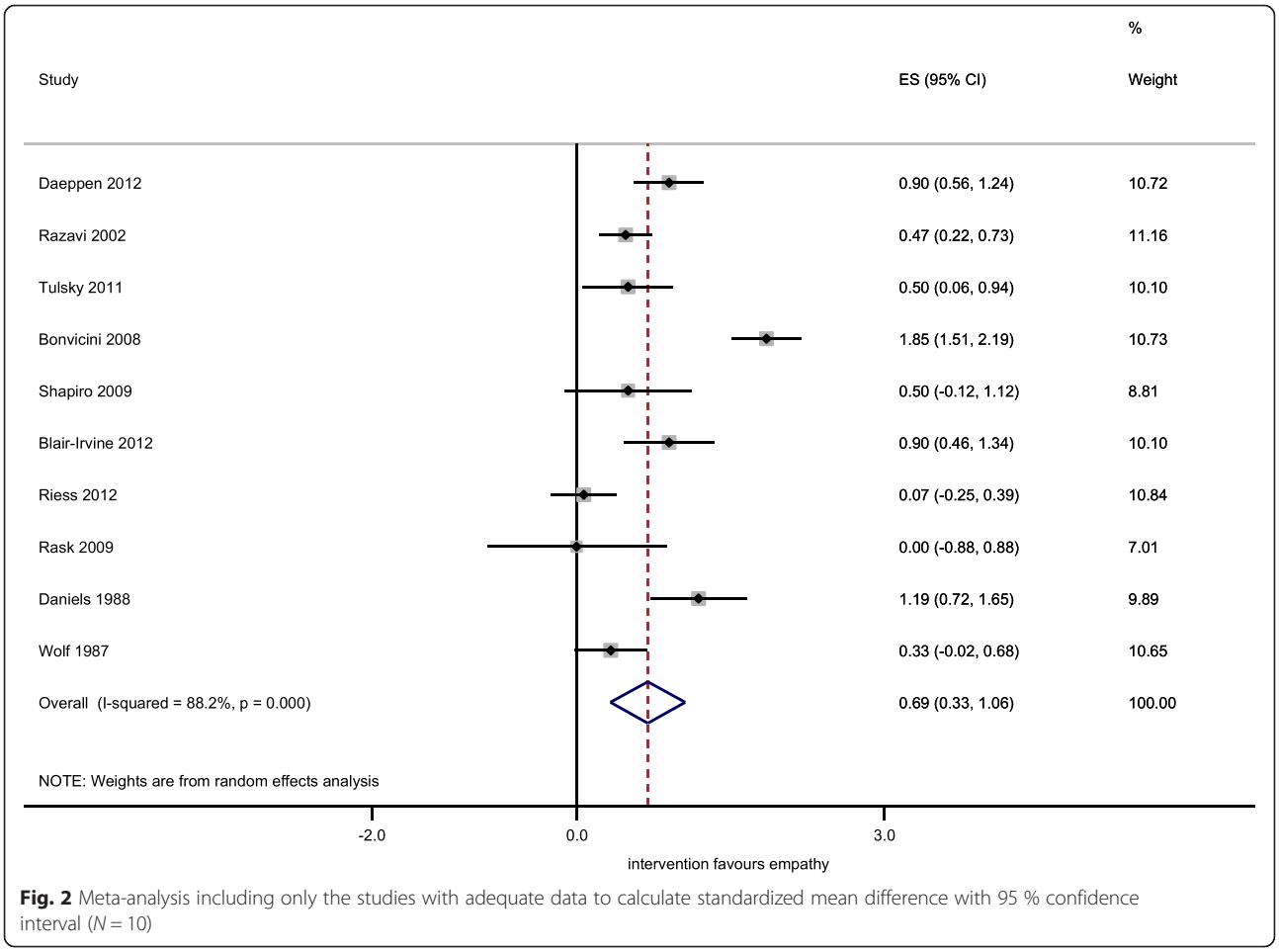

Fig. 2 Meta-analysis including only the studies with adequate data to calculate standardized mean difference with 95 % confidence interval (*N* = 10)

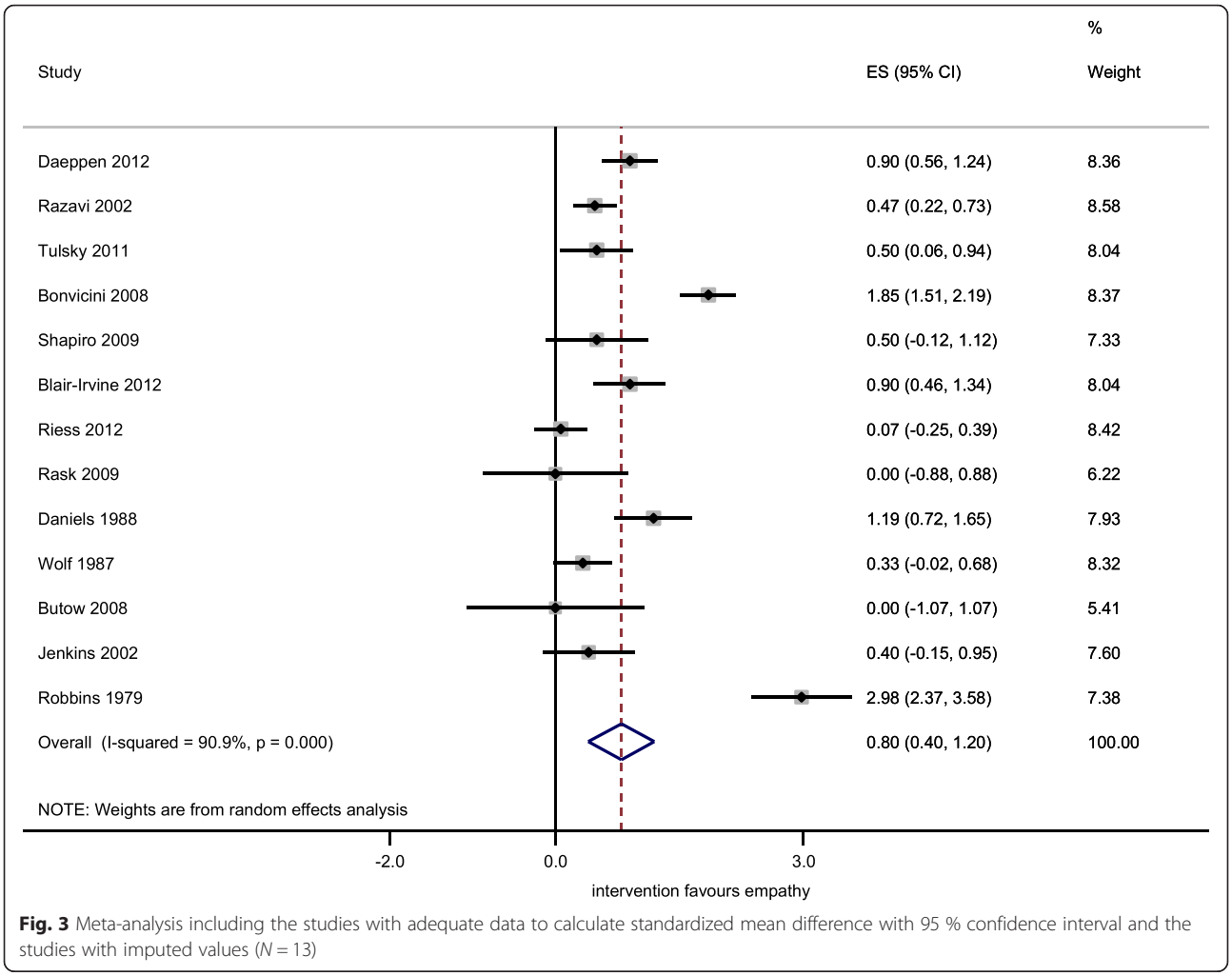

Fig. 3 Meta-analysis including the studies with adequate data to calculate standardized mean difference with 95 % confidence interval and the studies with imputed values (*N* = 13)

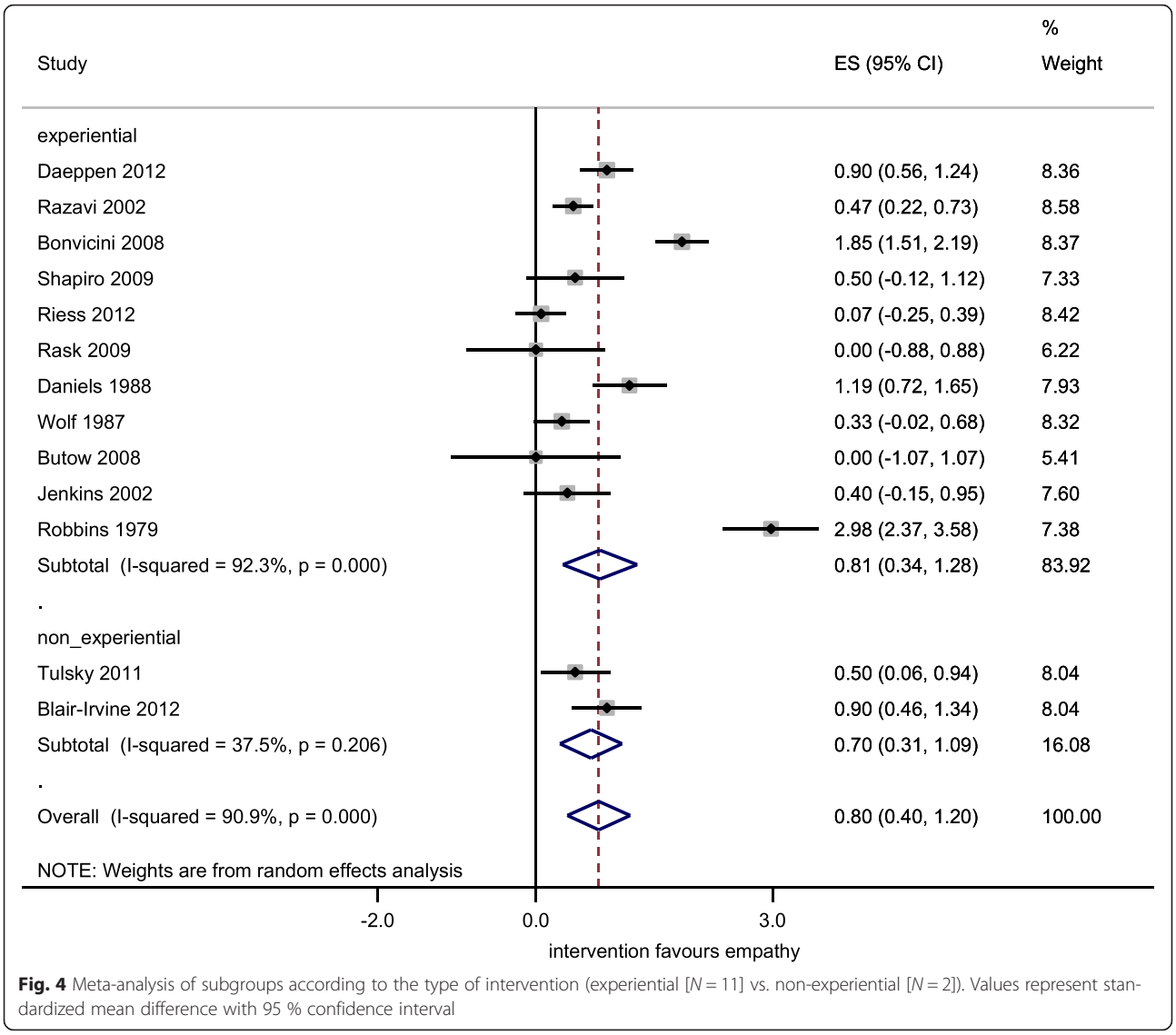

Fig. 4 Meta-analysis of subgroups according to the type of intervention (experiential [*N* = 11] vs. non-experiential [*N* = 2]). Values represent standardized mean difference with 95 % confidence interval

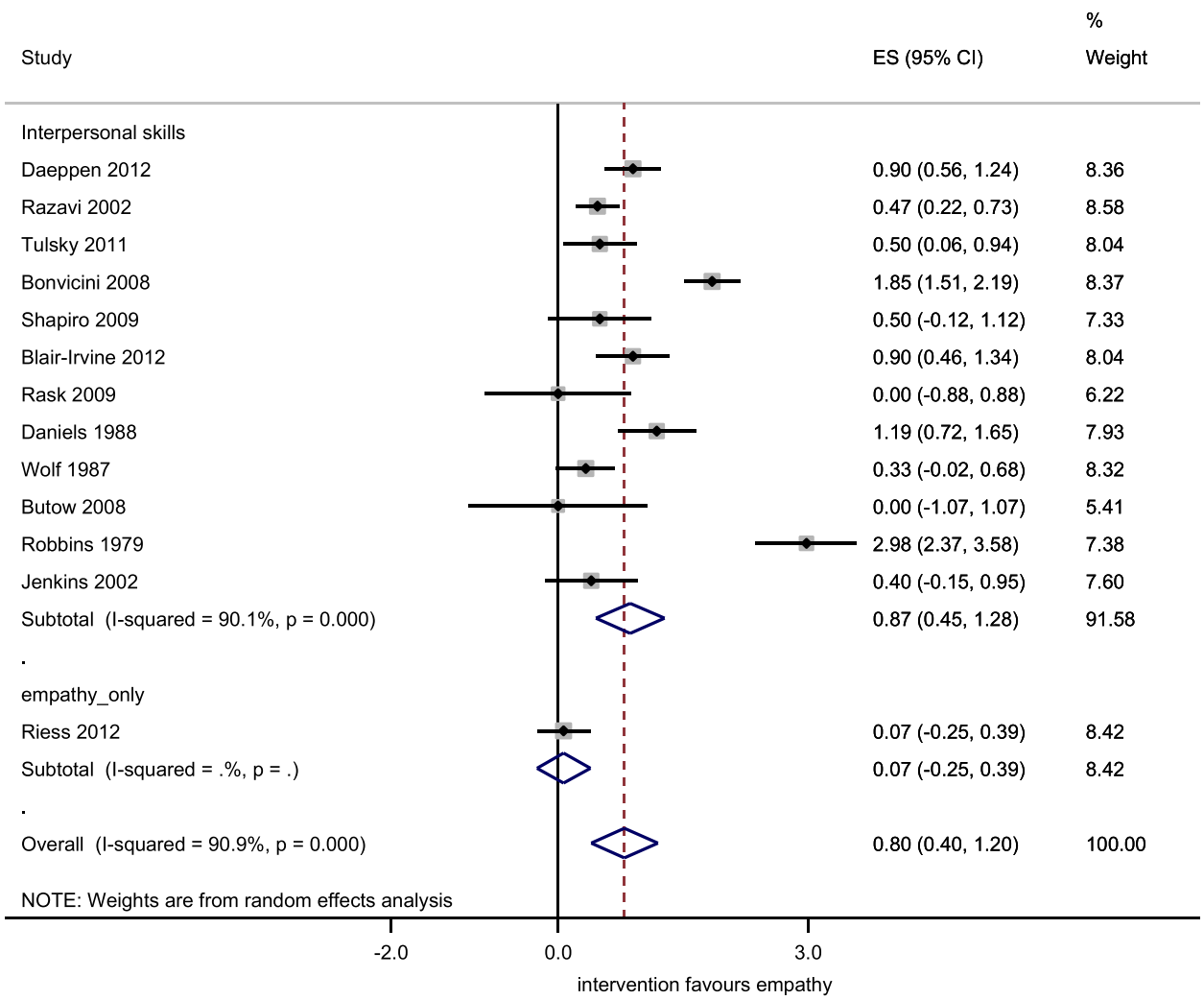

Fig. 5 Meta-analysis of subgroups according to the type of training (interpersonal skills programme [*N* = 12] vs. empathy-only training [*N* = 1]). Values represent standardized mean difference with 95 % confidence interval

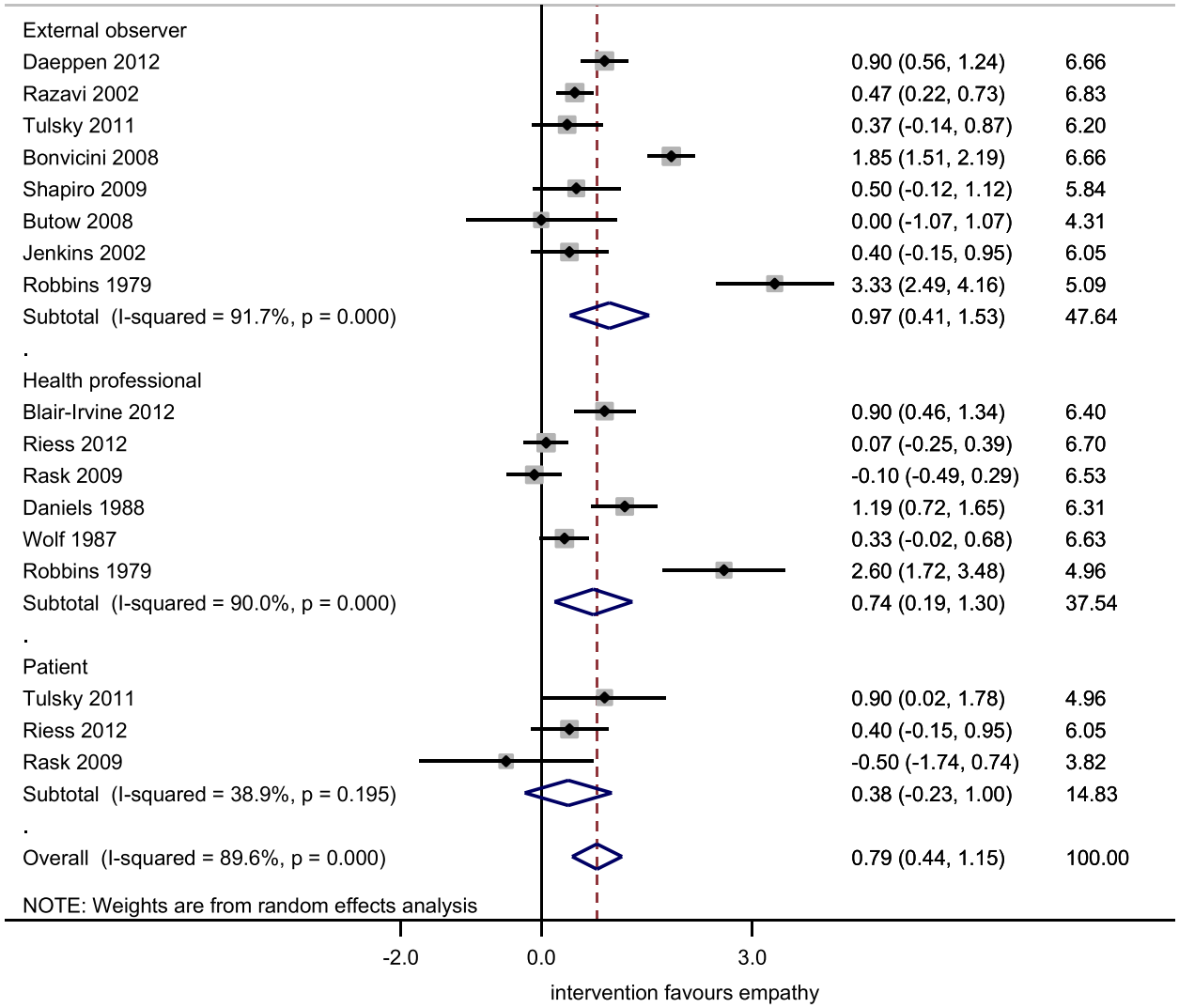

Fig. 6 Meta-analysis of subgroups according to the type of assessor (external observer [*N* = 8] vs. health professional participant [*N* = 6] vs. patient [*N* = 3]). Values represent standardized mean difference with 95 % confidence interval

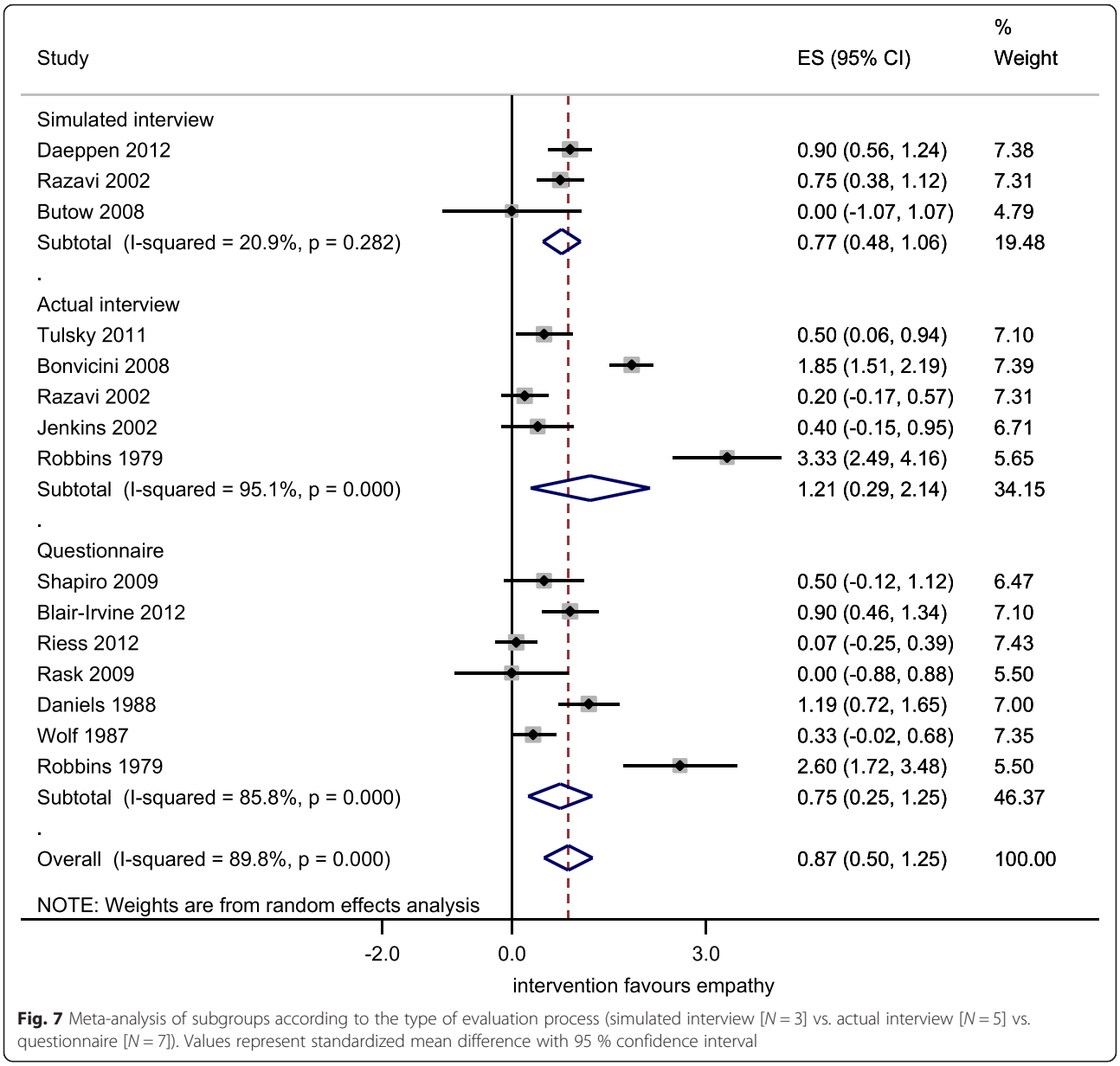

Fig. 7 Meta-analysis of subgroups according to the type of evaluation process (simulated interview [*N* = 3] vs. actual interview [*N* = 5] vs. questionnaire [*N* = 7]). Values represent standardized mean difference with 95 % confidence interval

Abbreviations
CI, confidence Interval; CONSORT consolidated standards of reporting trials; FEM, fixed effects model; IQR, interquartile range; MERSQI, medical education research study quality instrument; PRISMA, preferred reporting items for systematic reviews and Meta-analyses; RCTs, randomized controlled trials; REM, random effects model; RR, relative risk; SD, standard deviation; SMD, standardized mean difference

Acknowledgements
We thank Dr Dimitrios Mavridis, Assistant Professor in the University of Ioannina Department of Primary Education for discussing on the statistical analyses.

Funding
None.

Availability of data and materials
The datasets supporting the conclusions of this manuscript are included in Tables 1, 2 and 3 within the paper.

Authors' contributions
Dr AT had the original idea, organised the study design, and performed the statistical analyses; Mr VNK and Dr VTK collected the data; all three authors evaluated and interpreted the data and the results; Mr VNK and Dr AT wrote the first draft of the manuscript, which was critically reviewed by Dr VTK; all authors approved the final version. Dr AT agrees to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Competing interests
The authors declare that they have no competing interests.

Consent to publish
Not applicable.

Ethical approval and consent to participate
Not required.

Received: 18 April 2016 Accepted: 8 July 2016
Published online: 16 July 2016

References

1. Hojat M. Empathy in patient care: antecedents, development, measurement, and outcomes. New York: Springer; 2007.
2. Mercer SW, Reynolds WJ. Empathy and quality of care. Br J Gen Pract. 2002;52:9–12.
3. Attar HS, Chandramani S. Impact of physician empathy on migraine disability and migraineur compliance. Ann Indian Acad Neurol. 2012;1:89–94.
4. Platt FW. Empathy: can it be taught? Ann Intern Med. 1992;117(8):700. author reply 701.
5. Neumann M, Bensing J, Mercer S, Ernstmann N, Ommen O, Pfaff H. Analyzing the "nature" and "specific effectiveness" of clinical empathy: A theoretical overview and contribution towards a theory-based research agenda. Pat Ed and Couns. 2009;74(3):339–46.
6. Batson CD, Fultz J, Schoenrade PA. Distress and Empathy: Two Qualitatively Distinct Vicarious Emotions with Different Motivational Consequences. J Pers. 1987;55(1):19–39.
7. Zachariae R, Pedersen CG, Jensen AB, Ehrnrooth E, Rossen PB, von der Maase H. Association of perceived physician communication style with patient satisfaction, distress, cancer-related self- efficacy, and perceived control over the disease. Br J Cancer. 2003;88:658–65.
8. Kim SS, Kaplowitz S, Johnston MV. The effects of physician empathy on patient satisfaction and compliance. Eval Health Prof. 2004;27:237–51.
9. Vermeire E, Hearnshaw H, Van Royen P, Denekens J. Patient adherence to treatment: Three decades of research. A comprehensive review. J Clin Pharm Ther. 2001;26:331–42.
10. Hojat M, Louis DZ, Markham FW, Wender R, Rabinowitz C, Gonnella JS. Physicians' empathy and clinical outcomes for diabetic patients. Acad Med. 2011;86:359–64.
11. Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. Influence of context effects on health outcomes: A systematic review. Lancet. 2001;357:757–62.
12. Rakel D, Barrett B, Zhang Z, et al. Perception of empathy in the therapeutic encounter: Effects on the common cold. Pat Educ Couns. 2011;85:390–7.
13. Rakel DP, Hoeft TJ, Barrett BP, Chewning BA, Craig BM, Niu M. Practitioner empathy and the duration of the common cold. Fam Med. 2009;41:494–501.
14. Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. Physician–patient communication. The relationship with malpractice claims among primary care physicians and surgeons. JAMA. 1997;277:553–9.
15. Frankel RM. Empathy research: a complex challenge. Patient Educ Couns. 2009;75:1–2.
16. Manual E, Snyder SEL. JD, for the Ethics, Professionalism and Human Rights Committee*, American College of Physicians. Ann Intern Med. 2012;156:73–104.
17. Decety J, Yang C-Y, Cheng Y. Physicians down-regulate their pain empathy response: an event-related brain potential study. Neuroimage. 2010;50:1676–82.
18. Cheng Y, Lin C-P, Liu H-L, Hsu Y-Y, Lim K-E, Hung D, Decety J. Expertise modulates the perception of pain in others. Curr Biol. 2007;17:1708–13.
19. Halpern J. What is Clinical Empathy? J Gen Intern Med. 2003;8:670–4.
20. Benoit LG, Veach PM, LeRoy BS. When you care enough to do your very best: genetic counselor experiences of compassion fatigue. J Genet Coun. 2007;16:299–312.
21. McMullen L. Oncology nursing and compassion fatigue: caring until it hurts. Who is caring for the caregiver? Oncol Nurs Forum. 2007;34:491–2.
22. Najjar N, Davis LW, Beck-Coon K, Doebbeling CC. Compassion fatigue: a review of the research to date and relevance to cancer-care providers. J Health Psychol. 2009;14:267–77.
23. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996;312:71–2.
24. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009;6(7):e1000100.
25. Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. JAMA. 2007;298(9):1002–9.
26. Reed DA, Beckman TJ, Wright SM, Levine RB, Kern DE, Cook DA. Predictive validity evidence for medical education research study quality instrument scores: quality of submissions to JGIM's Medical Education Special Issue. J Gen Intern Med. 2008;23(7):903–7.
27. Turner L, Shamseer L, Altman DG, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database Syst Rev. 2012;14:11.
28. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7:177–88.
29. Engels EA, Schmid CH, Terrin N, Olkin I, Lau J. Heterogeneity and statistical significance in meta-analysis: an empirical study of 125 meta-analyses. Stat Med. 2000;19:1707–28.
30. Blair Irvine A, Billow MB, Eberhage MG, Seeley JR, MacMahon E, Bourgeois M. Mental illness training for licensed staff in long-term care. Issues Ment Health Nurs. 2002;33:181–94.
31. Riess H, Kelley JM, Bailey R, Dunn EJ, Phillips M. Empathy training for resident physicians: A randomized trial of a neuroscience- informed curriculum. J Gen Intern Med. 2012;27(10):1280–6.
32. Daeppen JB, Fortini C, Bertholet N, et al. Training medical students to conduct motivational interviewing: A randomized controlled trial. Pat Ed Couns. 2012;87:313–8.
33. Tulsky JA, Arnold RM, Alexander SC, et al. Enhancing communication between oncologists and patients with a computer-based training program. Ann of Int Med. 2011;155:593–601.
34. Rask MT, Jensen ML, Andersen J, Zachariae R. Effects of an intervention aimed at improving nurse- patient communication in an oncology outpatient clinic. Cancer Nursing. 2009;32:E1–11.
35. Shapiro SM, Lancee WJ, Richard-Bentley CM. Evaluation of a communication skills program for first- year medical students at the University of Toronto. BMC Medical Education. 2009;9:11.
36. Bonvicini KA, Perlin MJ, Bylund CL, Carroll G, Rouse RA, Goldstein MG. Impact of communication training on physician expression of empathy. Pat Ed Couns. 2009;75:3–10.
37. 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. Psych-Oncology. 2008;17:209–18.

38. Delvaux N, Merckaert I, Marchal S, et al. Physicians' communication with cancer patient and a relative: A randomized study assessing the efficacy of consolidation workshops. Cancer. 2005;103(11):2397–411.
39. Shapiro J, Morrison EH, Boker JR. Teaching empathy to first year medical students: Evaluation of an elective literature and medicine. Education Health. 2004;17(1):73–84.
40. Fallowfield L, Jenkins V, Farewell V, Saul J, Duffy A, Eves R. Efficacy of a cancer research UK communication skills training model for oncologists: A randomized controlled trial. The Lancet. 2002;359:650–6.
41. Jenkins V, Fallowfield L. Can communication skills training alter physicians' beliefs and behavior in clinics. J Clin Onc. 2002;20(3):765–9.
42. Razavi D, Delvaux S, Marchal S, et al. Does training increase the use of more emotionally laden words by nurses when talking with cancer patients? A randomized study. Brit J Cancer. 2002;87:1–7.
43. Smith RC, Lyles JS, Mettler JA, et al. A strategy for improving patient satisfaction by the intensive training of residents in psychosocial medicine: A controlled, randomized study. Acad Med. 1995;70:729–32.
44. Wolf FM, Woolliscroft JO, Calhoun JG, Boxer GJ. A controlled experiment in teaching students to respond to patients' emotional concerns. J Med Ed. 1987;62:25–34.
45. Robbins AS, Kauss DR, Heinrich R, Abrass I, Dreyer J, Clyman B. Interpersonal skills training: Evaluation in an internal medicine residency. J Med Ed. 1979;54:885–94.
46. Daniels T, Denny A, Andrews D. Using microcounseling to teach RN students skills of therapeutic communication. J Nurs Ed. 1988;27(6):246–52.
47. Koepsell TD, Diehr PH, Cheadle A. Invited commentary symposium on community intervention trials. Am J Epidemiol. 1995;142(6):594–9.
48. Barone FD, Hutchings PS, Kimmel HJ, Traub HL, Cooper JT, Marshall CM. Increasing empathic accuracy through practice and feedback in clinical interviewing course. J Soc Clin Psych. 2005;24(2):156–71.
49. Rogers C. Client-centered Therapy: Its Current Practice, Implications and Theory. London: Constable; 1995.
50. McKinstrey B, Aschcroff RE, Car J, Freeman GK, Sheikh A. Interventions for improving patients' trust in doctors and group of doctors. Cochrane Database Syst Rev. 2006;3:CD004134.
51. Moore PM, Wilkinson SSM, Rivera MS. Communication skills training for health care professionals working with cancer patients, their families and/or carers. Cochrane Database Syst Rev. 2009;1:CD003751.
52. Merzel C, d'Affilitti J. Reconsidering community- based health promotion: promise, performance and potential. Am J Public Health. 2003;93(4):557–74.