Empirical research on empathy in medicine—A critical review

Reidar Pedersen*

Department of General Practice and Community Medicine, University of Oslo, Oslo, Norway

1. Introduction

Empathy is generally considered important and positive to help patients in a good way, and empirical research on medical students’ and physicians’ empathy is growing. For example, many studies have shown that empathy may be stunted or reduced during medical training (see Section 3.5.1), and these tendencies have given rise to considerable concern.

Generally, empathy in medicine may be described as appropriate understanding of the patient [1]. However, there is no general agreement concerning how to define, teach, or study empathy. Some conceptual issues that have been hotly debated are whether empathy is emotional or cognitive, subjective or objective, and whether empathy includes communicating the understanding generated or acting appropriately based upon this understanding. Some researchers have argued that empathy is a multidimensional construct and have used more inclusive methods, while others have chosen to study selected dimensions. Empirical studies of empathy have been reviewed in various publications (see e.g. [2–13]). However, after reading publications in which empirical research on medical students’ or physicians’ empathy has been presented or discussed, my impression was that important methodological assumptions, ideals, and trends did not receive adequate attention. Furthermore, none of the previous reviews were systematic reviews including both qualitative and quantitative methods used to study empathy in medicine. Thus, this critical review was undertaken. The focus in this article is on the methods used to study empathy in medicine—particularly methodological limitations and challenges—and the reported results in the reviewed publications are only presented where relevant to illustrate methodological aspects. Thus, the publications reviewed include many positive contributions and interesting results not presented here.

2. Methods

A systematic literature search in Ovid MEDLINE(R), PsycINFO, EMBASE, and CINAHL was performed from May to August 2008. Publications presenting empirical research on medical students’ or physicians’ empathy were searched for (through subject headings related to empathy [AND] medical students or physicians [AND] empirical research; see Box 1. Languages included: English, German, Spanish, and the Scandinavian languages). In addition, other publications were identified

* Corresponding author at: Department of General Practice and Community Medicine, Section for Medical Ethics, University of Oslo, P.O. Box 1130, Blindern, NO-0318 Oslo, Norway. Tel.: +47 22 84 46 63; fax: +47 22 85 05 90.
E-mail address: Reidar.Pedersen@medisin.uio.no.

0738-3991/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.pec.2009.06.012
**Box 1. Subject headings used.**

<table>
<thead>
<tr>
<th>Empathy—subject heading</th>
<th>Empathy</th>
<th>Empathy</th>
<th>Empathy</th>
<th>Empathy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Medical students or physicians—relevant subject headings used</strong></td>
<td>Medicine; education, medical; education, medical, continuing; education, medical, graduate; education, medical, undergraduate; clinical clerkship; internship and residency; physicians; students; medical; physician–patient relations; psychotherapy; psychotherapeutic processes</td>
<td>Medical students; physicians; medical education; medical internship; medical residency; psychiatric training; medical personnel; psychotherapeutic processes; psychotherapists; psychotherapy; psychotherapy training; psychiatry</td>
<td>Medicine; medical school; residency education; medical student; medical education; physician; medical decision-making; doctor patient relation; medical practice; psychotherapy; psychotherapist</td>
<td>Medicine; physicians; education, medical; students, medical; education, medical, continuing; physician–patient relations; psychotherapeutic processes; psychotherapy; psychotherapists</td>
</tr>
</tbody>
</table>

| Empirical research—relevant subject headings used | Empirical methods; methodology; experimental methods; observation methods; behavioral assessment; qualitative research; quantitative methods; experimental design; between groups design; clinical trials; cohort analysis; follow-up studies; hypothesis testing; longitudinal studies; repeated measures; experimentation; psychometrics; statistical analysis; test construction; surveys; measurement; consumer surveys; mail surveys; telephone surveys; data collection; Likert scales; needs assessment; questionnaires; interviewing; case report; meta-analysis; test validity; test construction; testing; rating scales; multidimensional scaling; personality measures; test reliability; error of measurement; inter-rater reliability; test standardization; psychometrics; statistical validity; consistency (measurement); statistical correlation; statistical measurement; statistical samples; item analysis (statistical); statistical tests; statistical reliability; statistical significance; psychological assessment; galvanic skin response; experimental design; neurosciences | Research; methodology; interview; grounded theory; qualitative research; observational method; non-participant observation; participant observation; quantitative study; applied research; behavioral research; descriptive research; empirical research; ethnographic research; evaluation research; questionnaire; open-ended questionnaire; structured questionnaire; exploratory research; nursing research; delphi study; semi-structured interview; unstructured interview; psychologic test; biological model; neuroscience | Research; empirical research; research methodology; methodological research; interaction (research); evaluation research; descriptive research; action research; ethnographic research; professional practice, research-based; summative evaluation research; exploratory research; survey research; outcomes research; applied research; research, medical; clinical research; phenomenological research; research, interdisciplinary; education research; research, intradisciplinary; medical practice, research-based; ethnological research; qualitative research; predictive research; health services research; basic research; questionnaires; qualitative studies; ethnological research; ethnoursing research; grounded theory; naturalistic inquiry; psychological tests; models, biological; study design; neurosciences; quantitative studies; research, nursing; observational methods; non-participant observation; participant observation; interviews; behavioral research; open-ended questionnaires; structured questionnaires; semi-structured interview; structured interview; unstructured interview |

---

4 All subject headings were used with “explode” function and within each database, subject headings relating to medical students or physicians were combined with [OR]. The same was done with the subject headings relating to empirical research.
Box 2.
1. What kinds of methods are used to study empathy, with a special emphasis on whether the methodology used was predominantly quantitative (e.g. questionnaires with closed questions) or qualitative (e.g. qualitative interviews)?
2. Whose perspective is used to study empathy (e.g. physician, medical student, patient, relative, standardized patient, observer, peer)?
3. How many perspectives and methods are used?
4. The predominantly quantitative measures used were analyzed through the following questions:
   i. What evaluation strategy is used (e.g. definition or operationalization of empathy and items or questions used and rating system)?
   ii. Is empathy defined or operationalized as an emotional or cognitive process, and is the object of empathy the patient's feelings only, or are any aspects of the patient's experience included?
5. How are the possible influences of medical training and working conditions on empathy studied?
6. Are patients with reduced decision-making capacity included as research subjects?

Finally, 206 publications were selected for this review. These publications were analyzed through selected questions (see Box 2) that were formulated after doing a preliminary reading of the selected publications.

3. Results

3.1. What methods were used?

In the majority of the selected publications (171 of 206) predominantly quantitative methods were used. Among the 171 predominantly quantitative studies, 38 various quantitative measures were identified (see Table 1 and Section 3.4). However, 51 of the predominantly quantitative studies selected did not describe how empathy was evaluated or indicated that empathy was measured in an implicit or imprecise way [14–64]. (These studies' measures are generally not included in Table 1). However, some of these studies included one or more other measures of empathy that were more explicitly described, and these are included in Table 1, see e.g. [64].) For example, in one publication concerning satisfaction with anesthesia services it is reported (in the abstract) that one of the main items related to satisfaction was empathy from the anesthesiologist. However, in this publication, the only relevant information states that 92% of the patients had responded on the positive side of a six-point Likert-type question: “Did you feel that the anesthesiologist paid attention to your questions and comments” [53]. In another publication it is reported (again in the abstract) that physicians were evaluated through “unsystematic” database/internet searches, by reading reference lists, and through information from colleagues. This generated more than 2000 hits or publications. To select relevant publications among these, the title and abstract of every publication were examined, and when in doubt the rest of the publication was read. Publications were excluded if they did not include medical students, physicians, or their patients as informants or research subjects, dealt with group therapy only, did not present empirical research, presented or discussed personal experiences or case studies without explaining how the experiences or examples were selected or analyzed, or did not use the terms ‘empathy’ or ‘empathic’ (although there are few results from the qualitative part [98]. Another study did not provide details about how empathy was measured quantitatively and in the presentation of the results from the qualitative research, results explicitly linked to empathy were presented together with results pertaining to another category (non-judgment) and it is often not clear which of these two categories the presented results refer to [99]. In a third study using both qualitative and quantitative methods, empathy was emphasized both in the abstract, introduction, methods, discussion and conclusion – and the authors comments and conclusions seem to imply that the authors intended to study physician's empathic skills – but it is not stated explicitly how empathy was studied [100].

Although there is a growing evidence for neurobiological correlates of empathy, only two studies were found that used a biological approach to measure empathy in physician–patient interaction [101,102]. These two studies used skin conductance concordance (for therapist and patient) and the Barrett-Lennard Relationship Inventory (the Empathic Understanding Subscale, see Table 1).

3.2. Whose perspective was used to study empathy?

Approximately one-sixth of the studies used more than one perspective to study empathy (see Section 3.3).

Most frequently, empathy was studied through quantitative self-report measures probing rather general personal inclinations answered by medical students or physicians relatively far away from clinical practice (see Section 3.4.1 and Table 1). In the clinical encounter quantitative assessments performed by observers dominated. Less frequently, patients were asked to evaluate the medical student’s or physician’s empathy, and quite often simulated or standardized patients did the evaluation (see e.g. [25,103–108]). Furthermore, studies and measures that included patients’ or physicians’ concrete experiences and interpretations in practice seem to be absent. That is, when patients or physicians were used as respondents or informants, more general aspects were probed (e.g. physician’s personal inclinations/traits or questions probing whether or to what degree the physician understood the patient; not concrete details about what the patient or physician understood/ misunderstood).

Thus, studying empathy in practice was relatively rare, and when it was done it was mostly done through quantitative measures of behavior or through relatively general patient ratings of the physician's empathy.
### Table 1
The quantitative measures.

<table>
<thead>
<tr>
<th>The measure's name, and publications where the measure is used or presented</th>
<th>Short description and evaluation strategy*</th>
<th>Empathy—cognitive or emotional/affective</th>
<th>The patient's feelings or any experiences?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-report measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpersonal Reactivity Index (IRI) [41,60,113,114,122,123,125–132,135–137,153,163–167]</td>
<td>Self-report measure. Four 7-item subscales (in total 28 items): Perspective-Taking (PT), Fantasy (F), Empathic concern (EC), Personal Distress (PD). Answer scale A (Does not describes me well)—B—C—D—E (Describe me very well) A—E—rated from 0 to 4. Item examples: PT: “I sometimes try to understand my friends better by imagining how things look from their perspective”. FS: “I really get involved with the feelings of the characters in a novel”. EC: “I don’t feel very sorry for other people when they are having problems” (reverse coded). PD: “I tend to lose control during emergencies” (reverse coded).</td>
<td>Both cognitive and affective aspects involved</td>
<td>Any experiences</td>
</tr>
<tr>
<td></td>
<td>A measure of emotional empathy/Questionnaire of Emotional Empathy (QEMEE) [118,166,179–181]; further developed into the Balanced Emotional Empathy Scale (BEES) [98,138,139,182]</td>
<td>Self-report through 33 items, for example: “It makes me sad to see a lonely stranger in a group.”, “I become nervous if others around me seem to be nervous”, “I tend to lose control when I am bringing bad news to people.” “Lonely people are probably unfriendly,”, “I am able to remain calm even though those around me worry” (reverse coded), “Little people sometimes cry for no apparent reason.”, “I become very involved when I watch a movie.” Scale/answers from +4 (very strong agreement) to –4 (very strong disagreement) 7 subscales (e.g., “sensitivity to emotional contagion,” “extreme emotional responsiveness”). BEES consists of 30 items and is also answered through a nine-point agreement–disagreement scale. Example items: “Unhappy movie endings haunt me for hours afterward.” “I cannot feel much sorrow for those who are responsible for their own misery.”</td>
<td>Emotional</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Not explicitly described</td>
<td>Probably any experiences (see definition)</td>
</tr>
<tr>
<td>Hogans's Empathy Scale [58,104,107,111,134,141,142,150,151,155,193–199]</td>
<td>Self-report through 84 items, responded to on a true/false basis. The items were selected – through a rather complex procedure where some psychologists' conceptions of a highly empathic man and the definition below were cornerstones – from the California Psychological Inventory and the Minnesota Multiphasic Personality Inventory, and from testing forms used in studies at the University of California's Institute of Personality Assessment and Research. Item examples: (“empathic” answer in parenthesis): “I easily become impatient with people.” (False); “I have a natural talent for influencing people.” (True); “People today have forgotten how to feel properly ashamed of themselves.” (False) “I don't like to work on a problem unless there is the possibility of coming out with a clear-cut and unambiguous answer.” (False). Empathy is defined as the intellectual or imaginative apprehension of another's condition or state of mind without actually experiencing that person's feelings [195].</td>
<td>Not explicitly described, but emotions are not explicitly mentioned in the definition of empathy</td>
<td></td>
</tr>
<tr>
<td>Groningen Reflection Ability Scale [217]</td>
<td>Measures personal reflection (as distinct from clinical reasoning and scientific reflection) through 23 items (self-report). Three groups of personal reflection items: Self-reflection (10 items), empathetic reflection (6 items), and reflective communication (7 items). All items are scored on a five-point Likert scale (from 1 meaning ‘totally disagree’ to 5 meaning ‘totally agree’). The empathy items: “I am aware of the possible emotional impacts of information on others”, “I can empathize with someone else’s situation”, “I am aware of my own limitations”, “I reject different ways of thinking”, “Sometimes others say that I do overestimate myself”, “I am able to understand people with a different cultural/religious background”.</td>
<td>Not described</td>
<td>Not described</td>
</tr>
<tr>
<td>LaMonica's Empathy Construct Rating Scale (ECRS) [98,223,224]</td>
<td>ECRS includes 84 items, which are answered through self-report on a six-point scale (from –3 “Extremely unlike” to +3 “Extremely like”). The items describe behaviors exhibiting well-developed empathy or a lack of empathy. Item examples: “Seems to understand another person's state of being.”, “Does not listen to what the other person is saying.”, “Sometimes others say that I do overestimate myself”, “I am able to understand people with a different cultural/religious background”.</td>
<td>Both</td>
<td>Any experiences</td>
</tr>
<tr>
<td>General Empathy Scale (GES) [58]</td>
<td></td>
<td>Not described</td>
<td>Not described</td>
</tr>
<tr>
<td>Medical Empathy Scale (MES) [58]</td>
<td>12 self-report items are rated on a five-point continuum from strongly agree (+5) to strongly disagree (–1). Item examples: “I need to know ‘where my patients are coming from’ in order to treat their medical conditions adequately.”, “An important part of the care I provide to patients is emotional acceptance.”, “I don’t allow my patients to see my emotions” (reverse scored), “Patients are frequently to blame for their poor health status” (reverse scored), “There are times when I cannot pay full attention to what my patients are saying” (reverse coded).</td>
<td>Not explicitly described; probably both (see items)</td>
<td>Not explicitly described, but probably any experiences (see items)</td>
</tr>
<tr>
<td>The measure's name, and publications where the measure is used or presented</td>
<td>Short description and evaluation strategy*</td>
<td>Empathy—cognitive or emotional/affective</td>
<td>The patient's feelings or any experiences?</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Empathy and Attitudes Toward Caring for the Elderly (Modified Maxwell-Sullivan Survey)</strong> [233] “Empathy scale” [192] A similar empathy scale is used in [146]</td>
<td>A measure of empathy and attitudes toward caring for the elderly. Contains 11 items which are answered on a five-point Likert scale from 1 = strongly agree, to 5 = strongly disagree. Three items are used to measure empathy: “I can truly empathize with older patients”; “I understand what it feels like to have problems with aging”; “Understanding my elderly patients is valuable to me”. Empathy scale is one of four scales to measure physicians' attitudes to patient care. Five-point Likert-type scale. Five items in the empathy scale: “My patients often tell me things as stories that are important to my understanding of the reasons for their visits.” “My patients consider me to be a good listener.” “Patients often talk to me about how illness affects their lives.” “Even when patients present with minor symptoms, I find out how things are generally going in their lives.” “My patients assume that I am interested in them as people.” [192]</td>
<td>Not described</td>
<td>Not explicitly described, but only feelings are mentioned in the questions Not described</td>
</tr>
<tr>
<td><strong>Observer rating or coding of behaviour</strong></td>
<td></td>
<td>Observer rating or coding of behavior</td>
<td>Not explicitly described, but descriptions of the stages focus on the client's feelings</td>
</tr>
<tr>
<td><strong>A Tentative Scale for the Measurement of Accurate Empathy (Accurate Empathy Scale)</strong> [60,168–174]</td>
<td>The scale differentiates nine stages of empathy. The rater listens to and rates segments of clinical communication recordings. The scale has also been used to rate written responses to role-played statement of patient's statements. Excerpts from the descriptions of the stages: Stage 1: Therapist seems completely unaware of even the most conspicuous of the client's feelings; his responses are not appropriate to the mood and content of the client's statements. Stage 5: Therapist accurately responds to all of the client's more readily discernible feelings. He also shows awareness of many less evident feelings and experiences, but he tends to be somewhat inaccurate in his understanding of these. Stage 9: The therapist in this stage unerringly responds to the client's full range of feelings in their exact intensity. Without hesitation, he recognizes each emotional nuance and communicates an understanding of every deepest feeling. “Accurate empathy involves more than just the ability of the therapist to sense the client or patient's “private world” as if it were his own. It also involves more than just his ability to know what the patient means. Accurate empathy involves both the therapist's sensitivity to current feelings and his verbal facility to communicate this understanding in a language attained to the client's current feelings.”</td>
<td>Not explicitly described, but focuses on understanding and accuracy</td>
<td>Not explicitly described, but focuses on feelings</td>
</tr>
<tr>
<td><strong>Carkhuff’s Empathic Understanding Scale (a revision of the Accurate Empathy Scale)</strong> [57,104,107,119,175]</td>
<td>An observer scores behavior through a five-level system to rate empathic understanding. Level 1–5—where 1 is worst and 5 is best. Excerpts from the descriptions of the levels: Level 1: The verbal behavioral expressions of the helper either do not attend to or detract significantly from the verbal behavioral expressions of the helpee(s) in that they communicate significantly less of the helpee's feelings and experiences than the helpee has communicated himself. Level 3: The expressions of the helper in response to the expressions of the helpee(s) are essentially interchangeable with those of the helpee in that they express essentially the same affect and meaning. Level 5: The helper's responses add significantly to the feeling and meaning of the expressions of the helpee(s) in such a way as to accurately express or, in the event of ongoing, deep self-exploration on the helpee's part, to be fully with him in his deepest moment</td>
<td>Not explicitly described, but focuses on understanding and accuracy</td>
<td>Not explicitly described, but focuses on feelings</td>
</tr>
<tr>
<td><strong>A pencil-and-paper empathy rating test</strong> [176–178]</td>
<td>A 10-item empathy scale which requires respondents to write brief responses to trigger statements. Examples of trigger statements: (1) My parents really get me down. They insist I study physics and chemistry, when I'm not at all interested in those subjects. (2) If my exam marks don't improve I'm going to fail and lose my government allowance. I don't know what to do. (3) I just can't communicate with my parents. Whenever I try to explain how I feel about things they get all upset and call me a fool. The coding rules are as follows (modified from Carkhuff): 0 = aggressive or derogatory response. 1 = non-empathetic: does not acknowledge feeling or content of trigger; includes advice, reassurance, closed question. 2 = partially acceptable: open-ended question that does not acknowledge feeling or content of trigger. 3 = interchangeable/empathetic: acknowledges both the feeling and the content of the trigger (i.e. some variation of the classic ‘you feel… because…’). 4 = facilitative: reflects but also adds deeper feeling and meaning to the trigger statement in a way which encourages self-exploration (not really to be expected after a brief statement of the problem). Definition of empathy: “Empathy in its pure form refers to a verbal response which reflects both the emotional content of the other’s speech and the cause of the feeling, as expressed by that other.” [178]</td>
<td>Not explicitly described, but focuses on emotions</td>
<td>Not explicitly described, but focuses on emotions</td>
</tr>
<tr>
<td><strong>The Roter Interaction Analysis System (RIAS)</strong> [59,147,206–215]</td>
<td>RIAS is used to code video or audio taped doctor–patient interaction. The RIAS assigns a code to each complete thought, usually expressed as a simple sentence, clause or single word during the visit, by either patient or physician, into one of 38 mutually exclusive and exhaustive categories. There are two main types of categories: Socioemotional Exchange and Task-Focused Exchange (for example biomedical and psychosocial conversation such as questions, information giving, and counseling). Socioemotional Exchange categories include Empathy statements (Empathy), that is statements that paraphrase, interpret, name or recognize the emotional state of the other person present during the visit. Examples: “This is distressing for you, I understand”, “The pain must be very upsetting for you”, “You seem to be a little bit tense”, “You must be worried”, “You must have been nervous”, “What a relief for you!”, “I understand how you must be feeling”. The RIAS also includes six-point Likert scales (1 = low; 6 = high), where coders are asked to do “Global Affect Ratings” or to rate various affects or the emotional context of the dialogue, or their overall affective impressions of the speakers. Ratings are assigned for both the doctor and the patient. “Empathy” is one of 13 listed affects.</td>
<td>Not described</td>
<td>Not explicitly described, but the focus is on emotions</td>
</tr>
<tr>
<td>Rating Scales for the Assessment of Empathetic Communication in Medical Interviews (REM) [121,216]</td>
<td>A test inspired by motivational interviewing, which assesses both empathic and confrontational behavior of the physician in transcripts of audio taped medical interviews. Nine items, rated on a seven-point Likert scale, the two endpoints are described in behavioral terms (see each item). Items 1–6 constitute one factor, named empathy, while items 7–9 constitute the other factor, named confrontation, which is regarded as a potential empathy “neutralizer”. 1. Did the physician provide the opportunity for the patient to give his/her opinion? (no opportunity vs. a lot of opportunity); 2. Did the physician treat the patient as an equal partner? (not equal vs. completely equal); 3. Did the physician show understanding of the patient’s point of view? (no understanding vs. a lot of understanding); 4. Did the physician try to put him/herself in the position of the patient (not at all vs. a lot); 5. Did the physician show interest in the patient’s opinion? (no interest vs. a lot of interest); 6. Did the physician put the patient under pressure? (no pressure vs. a lot of pressure); 7. Did the physician “preach”? (did not “preach” vs. “preached” a lot); 8. Did the physician admonish the patient? (not at all vs. a lot); 9. Was the physician responsive to the patient? (not responsive vs. very responsive); Empathy in REM is defined as the physician’s cognitive ability to perceive and understand the patient’s perspective and the behavioral ability to communicate this understanding to the patient [121].</td>
<td>Cognitive ability</td>
<td>Not explicitly described, but probably any experiences</td>
</tr>
<tr>
<td>Liverpool Clinical Interaction Analysis Scheme (LCIAS) [234]</td>
<td>LCIAS is designed to quantify verbal communication between patients with “medically unexplained symptoms” and general practitioners. The scheme contains 25 and 30 major codes for the patient and general practitioner, respectively. Two of the physician codes are “Empathic reflection” and “Non-empathic reflection”. Empathic reflection: Reflects the problematic situation labeled by patient, such as pain, distress or worry associated with problem or indication that the nature or intensity of suffering associated with problem is understood or validated. Need not refer to immediately preceding statement, or be correct. Include: metaphors conventionally regarded as indicating feelings rather than objective characteristics of problem (e.g. ‘pressure’). Exclude: reflection of discomfort caused in examination. Non-empathic reflection: Reflection of details of problem but with no empathy. Need not refer to immediately preceding statement, or be correct. Exclude: simple reflection of treatment history.</td>
<td>Not explicitly described, but emphasize “metaphors conventionally regarded as indicating feelings”</td>
<td>Not explicitly described</td>
</tr>
<tr>
<td>Instrument Resident Communication Evaluation Form [227]</td>
<td>Modified scoring tool which originally was developed for use in theaters. The tool evaluates an observed consultation through 33 parameters within six sub scores: Empathetic communication, relating to the listener, verbal communication, nonverbal communication, respect for dignity, and overall impression. All parameters are scored with a 10-point scale. Empathetic communication parameters: 1. Social sensitivity: Displays recognition of differences in ethnicity, gender, cognition, etc.; 2. Humor: Displays ability to use or respond to humor in interpersonal interactions; 3. Patience: Allows time for full interpersonal interaction; 4. Tactfulness: Displays the ability to recognize and compensate for patient’s feelings; 5. Enthusiasm: Displays energy and interest in the topic and/or interpersonal relationship; 6. Listening while speaking: Displays an ability to recognize and adapt to patient’s cognition; 7. Integration of self: Displays ability to integrate elements of own personality into interactions; 8. Synthesis: Integrates an array of emotional, attitudinal, intellectual, and behavioral qualities.</td>
<td>Both</td>
<td>Any experience</td>
</tr>
<tr>
<td>The Four Habits Coding Scheme (4HCS) [157]</td>
<td>The 4HCS consists of 23 items/behaviors derived from the core skills referred to in the Four Habits Models. The four habits are: Invest in the Beginning, Elicit the Patient’s perspective, Demonstrate Empathy, and Invest in the End. Items/behavior used to evaluate “Demonstrate Empathy”: Clinician openly encourages/is receptive to the expression of emotion; Clinician makes comments clearly indicating acceptance/validation of patient’s feelings; Clinician makes at least one statement that reflects feelings that the patient has expressed; “Tunes in to probable feelings that are not expressed at least once”, “Explorations joint”; “Exhibits appropriate use of self (i.e. awareness of self in relationship to patient in terms of age, sex, experience)”, “Refrains from asking fact-finding questions as a retreat from dealing with feelings”; “Refrains from offering solutions before fully exploring the problem”, “Refrains from offering reassurance or a homily”, and “Refrains from denying the patient’s feelings”.</td>
<td>Not described</td>
<td>Not explicitly described, but probably any experiences</td>
</tr>
<tr>
<td>Empathetic Communication Coding System (ECCS) [159,229]</td>
<td>The ECCS is a measure of physicians’ empathic communication and has two parts: identifying patient-created empathic opportunities and coding physician responses to those empathic opportunities. The empathic opportunity begins with a clear and direct statement of emotion, progress/positive development, or challenge (e.g. negative effect or devastating life-changing event) by the patient. The empathic communication is coded in a scheme by which the physician responses are placed into one of six levels: Level 5 (the most empathic level)—Statement of shared feeling or experience; Level 4—Confirmation; Level 3—Acknowledgment; Level 2—Implicit recognition of patient perspective; Level 1—Perfunctory recognition of patient perspective; Level 0—Denial of patient perspective.</td>
<td>Not explicitly described (the test focuses on the behavioral aspects of empathy, as distinct from cognitive and emotional empathy [159])</td>
<td>Not explicitly described, but probably both</td>
</tr>
<tr>
<td>Craig’s rating scale of 43 physician/medical student’s behaviors in interviews [140]</td>
<td>A scale including a list of nine main types of interview behaviors (for example opening, empathy, and closure) (in total 43 behaviors). Video recorded interviews are rated. The behaviors are rated as either “did it” or “did not”. The empathy behaviors: “Responds to patient’s body language”, “Makes at least one statement that reflects feelings that the patient has expressed”, “Tunes in to probable feelings that are not expressed at least once”, “Explorations joint”; “Exhibits appropriate use of self (i.e. awareness of self in relationship to patient in terms of age, sex, experience)”, “Refrains from asking fact-finding questions as a retreat from dealing with feelings”; “Refrains from offering solutions before fully exploring the problem”, “Refrains from offering reassurance or a homily”, and “Refrains from denying the patient’s feelings”.</td>
<td>Not explicitly described, probably both</td>
<td>Not explicitly described, but probably only explicitly mentioned in the items</td>
</tr>
<tr>
<td>Coding of Empathic Opportunities and Continuers [231]</td>
<td>Audio recordings of clinic conversations are coded for the presence of empathic opportunities and physicians responses (continuers/offers empathy). Coding system: Empathic Opportunity (includes “Direct empathic opportunity”, defined as explicit verbal expression of emotion, “Indirect empathic opportunity”, defined as implicit verbal expression of emotion). Continuers (includes: “Name”, defined as state patient emotion; “Understand”, defined as empathizing with and legitimizing patient emotion; “Respect”, defined as praise patient for strength; “Support”, defined as show support; “Explore”, defined as ask patient to elaborate on emotion)</td>
<td>Not described</td>
<td>Emotions</td>
</tr>
<tr>
<td>Walters et al.’s Likert scale (the assessment strategy is not named) [222]</td>
<td>Consultations were videotaped, then empathy was measured on a five-point Likert scale (from 1 = the doctor’s response subtracts noticeably from affective communication, to 5 = the doctor’s response allows the patient to express feelings more fully)</td>
<td>Empathy not defined</td>
<td>Not described, but the assessment strategy mentions only feelings</td>
</tr>
<tr>
<td><strong>Patient rating measures</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Consultation and Relational Empathy (CARE) Measure [115,120,124,158,160,189–191,237]</td>
<td>Ten statements (with various examples under each of the statements) are answered by the patient on a five-point scale (from poor to excellent) plus “Does not apply”. The items: (i.e., was the doctor at “making you feel at ease”, “letting you tell your “story””, “really listening”, “being interested in you as a whole person”, “fully understanding your concerns”, “showing care and compassion”, “being positive”, “explaining things clearly”, “helping you to take control”, “making a plan of action with you”. Empathy in the clinical context involves an ability to (i) understand the patient’s situation, perspective and feelings (and their attached meanings); (ii) to communicate that understanding and check its accuracy; and (iii) to act on that understanding with the patient in a helpful (therapeutic) way [120,124]</td>
<td>Not explicitly described, but probably both [120,124]</td>
<td>Any experiences</td>
</tr>
<tr>
<td>Princess Margaret Hospital Patient Satisfaction with Doctor Questionnaire (PMH/PSQ-MD) [220,2221]</td>
<td>A patient-satisfaction-with-physician questionnaire developed within an outpatient oncology setting. Four domains: 1. Information exchange, 2. Interpersonal skills, 3. Empathy, and 4. Quality of time. Four-point Likert scale (strongly agree, agree, disagree, and strongly disagree). Six empathy items: “The doctor considered my individual needs when treating my condition”, “There were some things about my visit with the doctor that could have been better”, “It seemed to me that the doctor wasn’t really interested in my emotional well-being”, “The doctor seemed rushed today”, “The doctor should have shown more interest”, “There were aspects of my visit with the doctor that I was not very satisfied with”.</td>
<td>Not described</td>
<td>Probably any experiences (both thinking and feeling mentioned)</td>
</tr>
<tr>
<td>The Therapeutic Bond Scales-Revised (TBS-R) [228]</td>
<td>The TBS-R has 22 items constituting three scales (Role Investment, Empathic Resonance, and Mutual Affirmation). Empathic Resonance (ER) refers to the client’s sense that he or she and the therapist genuinely understand each other, and it leads to openness, genuineness, and a lack of inhibition. Includes eight items; seven items evaluate how much the participant felt frustrated, withdrawn, confused, cautious, strange, embarrassed, and inhibited (all indicating poor ER). The last item assesses to what extent the participant felt that the therapist understood what he or she was thinking and feeling.</td>
<td>Not described</td>
<td>Probably any experiences (both thinking and feeling mentioned)</td>
</tr>
<tr>
<td>SERVQUAL questionnaire [161]</td>
<td>SERVQUAL is a market research technique/questionnaire, adapted for use in a hospital setting measuring patient’s perception of quality (their expectations and perceptions, and the possible gaps between them). Includes five broad dimensions of service quality: tangibility, reliability, responsiveness, assurance, and empathy. 22 statements are scored on a nine-point scale ranging from “strongly agree” (9) to “strongly disagree” (1). The five empathy statements: “Excellent NHS hospitals would give patients individual attention (e.g., learning a patient’s specific medical history, flexibility to accommodate individual patients’ requirements, preferences, dislikes)”, “Excellent NHS hospitals would listen to patients’ requests and keep patients informed (e.g. listening to patients’ ideas, new opinions, general enquiries)”, “Excellent NHS hospitals would have 24-hour availability (e.g., evening appointments, 24-hour emergency availability)”, “Excellent NHS hospitals would have patients’ best interests at heart (e.g. building long-term relationships, providing leading-edge medical care)”, “Hospital staff of excellent NHS hospitals would understand the specific needs of patients (e.g., recognizing the importance of the patient, what the patient wants)”. Patients are also asked to rate the relative importance of the five domains and to what degree their expectations were met.</td>
<td>Not described</td>
<td>Probably any experiences (probably any experiences; see items)</td>
</tr>
<tr>
<td>Scales for patient perceived empathy and related constructs [230]</td>
<td>The scales include six scales measuring patient’s perception of physician communication skills, including cognitive and affective empathy, cognitive information exchange, partnership, physician expertise and interpersonal trust. In addition two scales for patient compliance and patient satisfaction were developed. All items were answered by the patient through five-point Likert scale (from strongly agree to strongly disagree). Cognitive empathy scale: 1. Interested in knowing what my experience means to me. 2. Still understands me when I am not clear. 3. Always knows exactly what I mean. Affective empathy scale: 1. Responds to me mechanically. 2. Tries to keep me from worrying. 3. Respects my feelings. 4. Shows interest in me. 5. Shows caring about my psychological well-being. 6. Shows great concern for my well-being. 7. Cares about me.</td>
<td>Both</td>
<td>Any experiences</td>
</tr>
<tr>
<td>Reynolds Empathy Scale [120,235,236]</td>
<td>12 items answered on seven-point scale, from “1. Always like”, to “7. Never like”. Examples of items: “Attempts to explore and clarify feelings”, “ignores verbal and non-verbal communication”, “explores personal meaning of feelings”, “judgments and opinionated”, “interrupts and seems in a hurry”, “Provides the client with direction”, “Fails to focus on solutions/does not answer direct questions/lacks genuineness”. Definition of empathy (from La Monica): “Empathy signifies a central focus and feeling with and in the client’s world. It involves accurate perception of the client’s world by the helper, and the client’s perception of the helper’s understanding” [236].</td>
<td>According to Reynolds, the definition of empathy is cognitive-behavioral [236], but the definition also includes emotional aspects</td>
<td>Any experiences</td>
</tr>
</tbody>
</table>

Table 1 (Continued)
<table>
<thead>
<tr>
<th>Table 1 (Continued)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The measure’s name, and publications where the measure is used or presented</td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td><strong>Patient perception of the doctor’s empathy</strong> [225,226]</td>
</tr>
<tr>
<td><strong>Jefferson Scale of Patient Perceptions of Physician Empathy (JSPPE) [109,110]</strong></td>
</tr>
<tr>
<td><strong>Garcia-Morillo et al.’s Likert scale [154]</strong></td>
</tr>
<tr>
<td><strong>Miscellaneous</strong></td>
</tr>
<tr>
<td><strong>Barrett-Lennard’s Relationship Inventory (RI) [101,102,104,107,111,117,200–203]</strong></td>
</tr>
<tr>
<td><strong>Hornblow’s empathy rating (inspired by Hogan) [104,107,111]</strong></td>
</tr>
<tr>
<td><strong>The Group Assessment of Interpersonal Traits (GAIT) [108,204,205]</strong></td>
</tr>
<tr>
<td><strong>The Affect Reading Scale (ARS) [156,218,219]</strong></td>
</tr>
<tr>
<td><strong>Analytic global rating form used in the OSCE (objective structured clinical examination) [105,106]</strong></td>
</tr>
<tr>
<td><strong>The Affective Sensitivity Scale [119,232]</strong></td>
</tr>
</tbody>
</table>

* The description includes short descriptions and evaluation strategies for the measures mentioned in the table. These measures are used to evaluate the empathy and affective sensitivity of health care professionals, as demonstrated by their interactions with patients. The measures are designed to assess various aspects of empathy, such as understanding, concern, and emotional awareness. The evaluation techniques include both self-report and observer ratings, as well as standardized tests and interviews. The measures are used in both clinical and educational settings to assess the quality of patient care and to guide professional development. The table provides a brief overview of the measures, their purposes, and the criteria used to evaluate them.
### Table 1 (Continued)

<table>
<thead>
<tr>
<th>The measure's name, and publications where the measure is used or presented</th>
<th>Short description and evaluation strategy*</th>
<th>Empathy—cognitive or emotional/affective</th>
<th>The patient’s feelings or any experiences?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Judged empathy [103]</td>
<td>The patient rates the physician's empathy on a Likert-type scale using the following three items: (a) “This physician was sensitive to my feelings,” (b) “This physician seemed to understand my situation/concerns,” and (c) “I felt at ease with this physician”. Furthermore, trained assessors rate the physicians’ empathy on a four-point scale (from 1 (little or no evidence of empathic ability) to 4 (considerable evidence of empathic ability). Definition of empathy: A capacity and motivation to take in patient/colleague perspective, and sense associated feelings—the ability to generate a safe/understanding atmosphere.</td>
<td>Not described</td>
<td>Probably any experiences</td>
</tr>
</tbody>
</table>

* Not all publications cited use all subscales, all items or all rating levels in the original test. All items from instruments containing fewer items than 10 are generally presented in the table, given that the items are published in some of the publications using or developing the instrument. From the other instruments, only examples or excerpts are provided in the table.

### 3.3. How many methods and perspectives were used?

33 of the selected studies used more than one type of perspective or method to study empathy in medicine [25,26,37,41,57–60,64,65,92,98–105,107,109–121]. From a methodological perspective, it is worth noting that over half of these studies indicate that the levels of empathy measured depends on both the perspective and the method used (see [25,26,37,57–60,98,103,104,107,99–112,115,117,118,121]). One example is a study reporting that the Interpersonal Reactivity Index (IRI) (measuring empathy through self-report, see Table 1) did not reflect important, relevant, and positive changes in medical students’ interpersonal skills after training; that is raised skills at eliciting patients’ social and personal concerns, increased depth of understanding and communication of this, greater ability to ask more relevant and clearer questions, and improved skills to begin and conclude interviews [60]. These changes were detected when raters used rating scales to assess video recordings (e.g. Accurate Empathy Scale, see Table 1).

However, also among the studies that used more than one method or perspective, there are also quite a few that do not adequately report how empathy was studied, or that indicate that empathy was studied in an implicit, imprecise, or peripheral way (e.g. using one or more implicit measures of empathy) (see e.g. [25,26,37,41,57–59,98–100,115,117,118,121]).

Among the qualitative studies, very few studies used more than one qualitative method [65,92], and one of these barely reported any explicit results from one of the two methods used [92]. For studies combining qualitative and quantitative methods see Section 3.1.

### 3.4. The quantitative measures used (see Table 1)

#### 3.4.1. Evaluation strategy

38 different measures that have been used to quantify empathy in medicine were identified (see Table 1). Many studies presenting or using the various empathy measures do not provide a definition of empathy—and among those that do define empathy, the definitions are quite varied (see Table 1 and Section 3.4.2).

As indicated above, the most common way to operationalize empathy is to focus on relatively general personal inclinations distanced from the physician–patient encounter or on observable behaviors (see the measures using self-report or an observer/trained rater in Table 1). Furthermore, many definitions and operationalizations of empathy explicitly or implicitly presuppose the existence of various dichotomies, for example cognitive versus emotional empathy; behavioral versus “inner” empathy, emotional or psychosocial concerns versus other concerns, socioemotional/affective versus instrumental communication, and empathy versus other aspects of clinical understanding (illustrations are provided below). Among the items or questions used to measure empathy, some are very general or of questionable relevance to empathy in clinical practice. For example: “I just can’t communicate with my parents. Whenever I try to explain how I feel about things they all get upset and call me a fool” (test subjects are asked to write down what they regard as an appropriate response. A pencil-and-paper empathy rating test); “Little children sometimes cry for no apparent reason” (reverse coded) (A measure of emotional empathy); “I sometimes try to understand my friends better by imagining how things look from their perspective” (IRI); “There were some things about my visit with the doctor that could have been better” (Princess Margaret Hospital Patient Satisfaction with Doctor Questionnaire);

“Excellent NHS hospitals would have 24-hour availability” (SERVQUAL questionnaire); “Refrains from offering reassurance or a homily” (i.e. not offering reassurance results in higher empathy ratings, Craig’s rating scale).

Some items used may even be counterproductive to the physician’s role and empathy in clinical practice. For example, if you agree to the following statement – “I am able to remain calm even though those around me worry” – this will reduce your empathy score. The same will happen if you disagree with “I tend to lose control when I am bringing bad news to people” (both items are from “A measure of emotional empathy”). However, do we want the physician to lose control in such situations?

Another item used is “I become very involved when I watch a movie” (A measure of emotional empathy). A similar item is used in the Jefferson Scale of Physician Empathy: “I do not enjoy reading non-medical literature or experiencing the arts.” (reverse coded). Using these kinds of items to measure physicians’ empathy presupposes that becoming involved in or enjoying literature, movies, or art is related to physicians being more empathic. However, is this necessarily so? Another empathy test – Davis’ IRI – includes “fantasy” as one of the subscales. One of the items is “I really get involved with the feelings of the characters in a novel.” However, Davis neither expected nor found any relationship between the fantasy subscale and measures of interpersonal functioning, and comments that “it is not apparent that a tendency to become deeply involved in the fictitious world of books, movies, and plays will systematically affect one’s social relationships” [122].

In the Jefferson Scale of Physician Empathy (JSPE) most of the items survey attitudes towards empathy and related phenomena, rather than empathy, for example: “An important component of the relationship with my patients is my understanding of the emotional status of themselves and their families.” Such items are at best only indirectly related to the physician’s empathy in practice.

A similar item is found in the test Empathy and Attitudes Toward Caring for the Elderly: “Understanding my elderly patients is valuable to me.” This latter-mentioned test also relies heavily on items that probe the subjects’ beliefs about their abilities to
empathize ("I can truly empathize with older patients"). Such self-
evaluation items also appear in other measures using self-reports.
However, the subjects’ self-perception may be more or less
accurate and more or less related to empathy in practice.

Quite a few studies and measures investigate empathy through
coding or rating of observable behavior only. Measures focusing on
observations of behavior only, often seem to contribute to a neglect
of non-observable experiences and interpretations, or contribute
to implicit assumptions about the physicians’ or patients’ concrete
experiences and interpretations. Some items and questions used
even seem to presuppose that the rater or observer knows what the
patient’s feelings or experiences are. One example is when
directors of a residency program are asked to rate the residents
by answering the following question on a 10-point Likert-type
scale: “How do you rate this resident’s empathic behavior, defined
as an understanding of the patient’s inner experiences and
perspectives, and a capability to communicate this understanding”
[64], or when instructing observers to rate “the student’s
perspectives, and a capability to communicate this understanding”
[64]. These two examples were considered as
rather implicit measures of empathy and are thus not included in
Table 1. However, similar evaluation strategies can be found in the
Roter Interaction Analysis System (RIAS), Empathic Communica-
tion Coding System (ECCS), and Carkhuff’s Empathic Under-
standing Scale (see Table 1). Coding of Empathic Communication in Medical (REM), Craig’s rating scale.

Furthermore, some of the behavioral approaches to empathy
tend to dichotomize physician–patient interaction, for example
physical symptoms (e.g. pain) and medical issues versus other
issues (ECCS) and Socioemotional Exchange (includes empathy)
versus Task-Focused Exchange (i.e. biomedical and psychosocial
conversation, which includes questions, information giving, and
counseling) (RIAS). Thus, they seem have the potential to
categorize what many patients would regard as empathic behavior
as something else than empathy (e.g. if asking about psychosocial
problems or pain symptoms, such behavior would probably not be
coded as empathic behavior in the RIAS and ECCS, respectively).

The Groningen Reflection Ability Scale distinguishes between
three types of reflection in medicine—clinical reasoning, scientific
reflection, and personal reflection (the latter includes empathetic
reflection, self-reflection and reflective communication). However,
the relationships between empathy/personal reflection and the
other two types of reflection are barely addressed. Thus, the scale
has the potential to explore physician’s personal reflection,
including empathetic reflection, isolated from clinical reasoning
and scientific reflection.

14 measures were identified that have been developed or used
to elicit the patient’s or a standardized patient’s perspective on
the physician/medical student’s empathy (see Table 1). One example is
the CARE measure, which has been developed relatively recently
within a medical context, based on qualitative and quantitative
validation procedures (including patients and physicians)
[120,124]. However, as indicated above, the measures investigating
the patient’s perspective have been relatively rarely used and
none of them explore what the patient regarded as the main
concern, or what the patient or physician specifically understood
or misunderstood. Rather, the items used are more general, for
example: “[my doctor] understands my emotions, feelings, and
concerns” (Jefferson Scale of Patient Perceptions of Physician
Empathy).

3.4.2. Is empathy defined or operationalized as an emotional or
cognitive process, and is the object of empathy the patient’s feelings
only?

As mentioned above, in many studies that present or use
empathy measures, empathy is not defined, and even among those
in which empathy is defined, it is often not explicitly stated how the
concept of empathy relates to cognitive and emotional aspects. In 26
of the 38 identified measures, it is not described explicitly whether
empathy includes cognitive and/or emotional aspects. Among, the
studies that explicitly define empathy as cognitive and/or emotional,
5 include both cognitive and emotional aspects, 4 define empathy as
cognitive, and 3 define empathy as emotional (see Table 1).

Another example is the presentation of the Jefferson Scale
of Physician Empathy (JSPE) in which empathy is defined as
cognitive. However, the questionnaire developed to measure
cognitive empathy also includes affective aspects of the empathic
process. Furthermore, studies validating the JSPE indicate that JSPE
does not only measure cognitive aspects, but also taps on
emotional aspects [113,114,116]. (Among the studies using more
implicit strategies to measure empathy (not included in Table 1),
there are also some that exaggerate and simplify the cognitive–
emotional distinction, see for example [38–40].)

Whether the scope or object of empathy is emotions or any
experiences is not explicitly described in 22 of the identified 38
measures. Only 6 measures explicitly delimit the object of
empathy to the patients’ emotions, but the operationalization of
empathy strongly focuses on the understanding of patients’
emotions or psychosocial aspects in yet another 10 measures
(see for example the Accurate Empathy Scale, Carkhuff’s Empathic
Understanding Scale, a pencil-and-paper empathy rating test, the
RIAS, and ECCS), something which tends to exclude physicians’
understanding of more cognitive or biomedically oriented con-
cerns from the study of empathy, often without any explanation or
discussion. In the RIAS, even psychosocial questions and informa-
tion exchange are separated from the coding of empathy, due to the
dichotomy between the two main coding categories (see above
and Table 1). (The researcher may of course combine various
categories to form an inclusive “empathy-cluster”, but no such
study was identified.)

3.5. How are the possible influences of medical training and
physicians’ working conditions on empathy studied?

3.5.1. Medical training

There are quite a few studies that indicate that empathy is
reduced or does not increase during medical training
[126,128,132–143], and that there may be correlations between
empathy levels and medical specialty (and medical students’
specialty preferences) [138,141,143–147], however, some report
no such patterns [35]. However, most of these studies are based on
quantitative measures using self-reports of various personal inclinations relatively far away from empathy in practice, and we know little about the possible factors contributing to the changes and variations in empathy, or their consequences in clinical practice.

There are also some studies with more variable results, comparing empathy scores and academic performance [111,128,130,131,142,148–151] or comparing physicians' or medical students' empathy with other professions or students [107,114,118,141,145,152]. Again, most of these studies are based solely on self-reported measures of empathy, and empathy in practice is very rarely measured.

Some of the qualitative and quantitative studies indicate that learning and knowing about symptoms, pathology, diagnosis, treatment, available services, and prognosis – that is some of the main parts of medical education – influences empathy, and indicate that these relationships may sometimes be of great importance [43,67,69,70,72,78,81,89,96,147]. However, in all these latter-mentioned studies, empathy or the various relationships were investigated relatively implicitly.

In some studies, aspects that may be related to or of importance to clinical reasoning have been investigated, such as tolerance for ambiguity, dogmatism, self-perceived errors, attribution of causes and guilt, number of prescribed drugs, biased questions, and rates of observable interpretations and confrontations [108,125,129,153–155]. However, most of the results are based on one empathy measure relatively far away from practice. One example is a longitudinal study that indicates that medical students' tolerance for ambiguity (e.g. tolerance for a situation that is complex, novel, or insoluble) correlates with empathy (measured using the IRI), and that empathy and tolerance for ambiguity correlate with various aspects of the student's performance in clinical examinations [129].

A few studies focus on how the informal curriculum affects empathy, for example socialization, the competing discourses of empathy, for example socialization, the competing discourses of empathy, and the objectification of patients [68,79]. However, in these studies, empathy and the possible relationships are explored relatively implicitly. Only one study designed to investigate the possible effects of (the formal or planned for) medical curricula on the students' empathy, was identified [156].

3.5.2. Working conditions

There are quite a few studies that indicate a relationship between being pressed for time and lowered empathy [18,80,91,93,109,157–160]. However, in some of these studies empathy was studied implicitly and far away from the clinical encounter. For example, in one of the studies the patients rated their physician's empathy several weeks or months after hospital discharge [159].

Some of the studies indicate that working conditions (more generally) may influence empathy [68,81,91,93,96]. However, all these studies investigated empathy or the various relationships relatively implicitly.

Furthermore, some studies did not explicitly focus on working conditions, but on possible indicators of challenging working conditions. For example, some quantitative studies indicate possible associations between empathy and physicians' fatigue, thirst, hunger, well-being, burnout, depression, and patient's waiting time for operation [126,127,161]. However, these studies evaluated empathy relatively far away from practice (i.e. IRI/self-report of empathy), or through ratings of the whole hospital or the hospital staff in general.

3.6. Patients with reduced decision-making capacity

None of the selected studies reported that patients with reduced decision-making capacity as research subjects were included (with one possible exception where patients with cognitive impairment were also included [154]).

4. Discussion and conclusion

4.1. Discussion

This review indicates that empirical research on empathy in medicine is abundant with quantitative measures and studies, and single-method approaches. The quantitative studies of empathy are often based solely on self-reports far away from medical practice and the patient, or uncritically focused on observable aspects, and in general remote from physician's and patient's concrete feelings, experiences and interpretations in practice. Some of the observer-based measures seem to presuppose that the rater is willing to make (more or less valid) assumptions about the patient's feelings or experiences.

The patient perspective has been surveyed through many measures. However, when concrete experiences and concerns are neglected, or if only the patient perspective is investigated, important aspects of empathy are still not assessed. Furthermore, quite a few of the studies and measures tend to separate the physician's more biomedically oriented experiences and perception from the empathic process and the clinical relevance of the measures' items is often questionable. Finally, the patient's more biomedically oriented or less affective experiences and concerns are often excluded. Thus, important aspects of the empathic process are often excluded from the measures used to study empathy in medicine, and the excluded aspects are rarely addressed through other methods when these empathy measures are used.

If we do not ask directly and specifically, we often do not know whether the physicians' "lack of empathic behavior" is due to poor understanding, poor communication skills, or some influences motivating the physicians to focus on other aspects than judged appropriate by the patient or the observer. Furthermore, research on other health care students indicates that self-assessment of empathy is particularly difficult and that self-report may result in overlooking the students that probably are at most need for training or supervision (i.e. those rated lowest by the observers) [162]. Since impaired or biased understanding requires other remedial measures than poor communication skills, and since self-reports often do not correspond with empathy in practice, empathy should not be studied only through quantitative behavioral approaches or self-reports of personal inclinations or abilities.

Qualitative methods seem to be largely underused in empirical research on empathy in medicine. Qualitative approaches may be particularly valuable to explore conceptual issues, concrete variations in the physician's and patient's understanding in practice (for example of the patient's main concern), what may foster and inhibit empathy, and how such modulating factors may influence empathy.

There are many studies that indicate that medical training and physician's working conditions influence empathy. Still, how medical training and working conditions may modulate empathy has barely been addressed in empirical research on empathy. For example we still have sparse knowledge about how medical training (both the formal and the informal curriculum) may promote and inhibit empathy, and in particular when it comes to clinical practice. There are several studies that have investigated the effect of targeted interventions to foster empathy among medical students and physicians (for a recent review see [4]). However, the topic pursued in this review was the possible influences of the other aspects of medical training, or medical training in general. Although there are negative correlations or no
correlation between academic performances and empathy scores, this is not sufficient to conclude that medical training does not influence empathy, or that it only influences it negatively. The same goes for the studies that show a decline in empathy or that empathy is stunted during medical training, first, because of the methods used and second, because of the possible variations behind the “big” numbers; that is, a total score that is zero or negative does not mean that there cannot be any positive numbers or influences.

We also have sparse knowledge about how various designs of medical curricula may affect the students’ empathy. Furthermore, given that the relatively consistent trend suggesting a negative relationship between time and empathy is valid, one important question seems unanswered. That is, does lack of time erode empathy, or are less empathic physicians more prone to rush when talking with the patient, or both? Furthermore, possible relationships between empathy and other aspects of working conditions, for example incentives, access, available services, clinical prioritizations, working environment, cooperation, and documentation routines, seem to be largely under researched. Finally, how medical training and working conditions in combination may influence empathy should be explored.

The lack of studies that include patients with reduced decision-making capacity indicates a lack of attention to empathy towards critically ill patients—an essential task for empathy in practice. Given the conceptual complexity and variability in definitions of empathy and the methods used, it is unfortunate that many publications do not present adequate information about how empathy was studied or what is meant by empathy. In many publications, the reader is referred to other sources or recommended to contact the authors. However, many of the sources cited do not give any further details, are unavailable, or are in a foreign language for the general reader; and personal requests are often not answered.

This critical review has some limitations. First, the selection of literature is not complete. For example, not all relevant publications are indexed in the databases used and some relevant publications may have been neglected through this review’s search strategy (e.g. if they did not use the terms ‘empathy’ or ‘empathic’). Second, the categorization of the publications is not self-evident or an exact science. For example, one could argue that the studies that use the IRI subscales measure empathy with more than one method. However, in this instance, since the subscales are all part of the same test, and all the subscales measure self-reported personal inclinations, this test was categorized as one method. Finally, this review’s focus on methodological limitations does not pay due attention to the important contributions and the pioneering work presented in many of the reviewed publications.

4.2. Conclusion

Empirical research on empathy in medicine is dominated by relatively narrow quantitative methods that include the physician’s and the patient’s concrete interpretations, feelings, and experiences to a limited extent. Furthermore, the possible influences of medical training and working conditions on empathy have not been adequately explored. In sum, the empirical studies of empathy tend to separate empathy from main parts of clinical perception, judgment, and communication. Thus, important aspects and influences of empathy have been relatively neglected.

4.3. Practice implications

Given the relational and complex nature of empathy, transparent concepts and methods, the use of more than one method and perspective, and qualitative methods, seem to be needed in future studies on empathy in medicine. Furthermore, studies should include the physician’s and the patient’s concrete experiences and interpretations, the context in which empathy is developed and practiced, and avoid peripheral or implicit approaches.

Conflict of interest statement

The author has no conflict of interest (e.g. financial, personal or other relationships with other people or organizations that could inappropriately influence, or be perceived to influence, his work).

Acknowledgments

This paper is part of a Ph.D. project funded by the Research Council of Norway and supervised by Professor Jan Helge Solbakk and Professor Anja Vetrle.

References


