

Accepted Manuscript

Title: Patient feedback questionnaires to enhance consultation skills of healthcare professionals: a systematic review

Authors: Hiyam Al-Jabr, Michael J. Twigg, Sion Scott, James A. Desborough



PII: S0738-3991(18)30120-4
DOI: <https://doi.org/10.1016/j.pec.2018.03.016>
Reference: PEC 5918

To appear in: *Patient Education and Counseling*

Received date: 11-9-2017
Revised date: 26-2-2018
Accepted date: 15-3-2018

Please cite this article as: Al-Jabr Hiyam, Twigg Michael J, Scott Sion, Desborough James A. Patient feedback questionnaires to enhance consultation skills of healthcare professionals: a systematic review. *Patient Education and Counseling* <https://doi.org/10.1016/j.pec.2018.03.016>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Patient feedback questionnaires to enhance consultation skills of healthcare professionals: a systematic review

Hiyam Al-Jabr^a, Michael J. Twigg^a, Sion Scott^a, James A. Desborough^a

^a *School of Pharmacy, University of East Anglia, Norwich, UK*

Corresponding author at: Hiyam Al-jabr, University of East Anglia, Norwich, NR4 7TJ, UK. Tel.: +44 (0)1603 59 1996. E-mail address: h.al-jabr@uea.ac.uk (H. Al-Jabr).

Highlights

- Several patient feedback questionnaires are available.
- Most patient feedback questionnaires are designed specifically for physicians.
- Patient feedback questionnaires need to be validated with different practitioners.
- Patient feedback can help in enhancing consultation skills of practitioners.

ABSTRACT

Objective: To identify patient feedback questionnaires that assess the development of consultation skills (CSs) of practitioners.

Methods: We conducted a systematic search using seven databases from inception to January 2017 to identify self-completed patient feedback questionnaires assessing and enhancing the development of CSs of individual practitioners. Results were checked for eligibility by three authors, and disagreements were resolved by

discussion. Reference lists of relevant studies and open grey were searched for additional studies.

Results: Of 16,312 studies retrieved, sixteen were included, describing twelve patient feedback questionnaires that were mostly designed for physicians in primary care settings. Most questionnaires had limited data regarding their psychometric properties, except for the Doctor Interpersonal Skills Questionnaire (DISQ). Most studies conducted follow-up, capturing positive views of practitioners regarding the process (n=14). Feedback was repeated by only three studies, demonstrating different levels of improvement in practitioners' performance.

Conclusion: Identified questionnaires were mainly focused on physicians, however, to support using patient feedback, questionnaires need to be validated with other practitioners.

Practice implications: Several patient feedback questionnaires are available, showing potential for supporting practitioners' development. Valid questionnaires should be used with appropriate practitioners in developing more evidence for the impact they may have on actual consultations.

1. Introduction

Good consultation skills (CSs) are essential for effective patient encounters and it has been shown to drive positive outcomes including enhancing patient adherence and satisfaction [1-6]. An emphasis was given by the British Medical Association (BMA) and the General Medical Council (GMC) indicating the importance for the quality of the physician's professional work to be assessed at regular intervals, by patients and colleagues [7], thus supporting their continuous professional development [8, 9]. Numerous methods are described in the literature regarding the assessment of practitioners' consultations skills, including assessments conducted by assessors [10], peers [11-13], or by patients [14-17]. A combination of methods are suggested to provide a more holistic assessment [18-20]. However, collecting feedback from patients is probably the most suitable method [21]. Patients, as customers of the healthcare system are capable of highlighting weak areas of performance that are not usually covered by other conventional methods [6, 22], or not recognised by practitioners themselves [23].

Patient feedback can be collected using questionnaires and/or by conducting interviews [24, 25]. However, the full benefit of feedback can only be realised by using it to support professionals' development. It can help practitioners to better understand their skills, acknowledge their strengths, identify areas needing further attention, and thus directing them to where improvements are needed [26-28].

Using feedback collected from patients to enhance the CSs of individual practitioners is not thoroughly studied. Initial searches identified two systematic reviews that investigated this domain [29, 30]. While these reviews identified several feedback

questionnaires, they were focused on assessing CSs of physicians. However, patient consultations are currently conducted by a wide variety of different practitioners and not only by physicians. For example, in the UK since 2005, several practitioners are legally allowed to prescribe medications to patients, including nurses and pharmacists [31-33], thus the number of patient consultations has greatly increased. Therefore, this systematic review was conducted to identify patient feedback questionnaires used to assess and enhance the development of individual CSs of all practitioners across all settings.

1.1. Aim & objectives

This review aimed to identify and describe patient feedback questionnaires that assess the development of CSs of individual practitioners. The objectives were to describe identified studies and questionnaires according to the following: (a) name of the questionnaire, (b) practitioners assessed, (c) assessment setting, (d) questionnaire administration method (patient recruitment, individual in charge of administration, and concealment methods), (e) patient feedback reporting methods, (f) follow up to patient feedback and its impact.

2. Methods

2.1. Literature search

A systematic search was conducted to identify relevant studies using the following databases: MEDLINE, EMBASE, AMED (via Ebsco), Web of Science, SCOPUS, CINAHL, and PsycInfo, from inception of the databases up to January, 2017. A protocol was developed and registered on the international database of prospectively registered systematic reviews (PROSPERO) (CRD42017055365). Search results

were limited by two filters: English language and publication type: journal, and they were exported into the reference manager Endnote 7.2.1, where duplicates were identified and removed. An example search strategy is provided in Table 1, which was adapted appropriately when searching the other databases. Reference lists of included studies and of related systematic reviews were examined to identify additional potentially eligible studies. Open Grey was also searched for grey literature, and corresponding authors were contacted by email where necessary for missing data.

2.2. Inclusion criteria

Studies that met the following criteria were considered eligible for inclusion:

1. Patient feedback questionnaires requiring self-completion by real (not simulated) patients (> 18 years old).
2. Assessment of CSs of a practitioner in a face-to-face interaction,
3. Post-consultation feedback collection,
4. Feedback results used for individual professional development.

We excluded studies meeting the following criteria:

1. Qualitative feedback collection methods (e.g. interviews/group discussions),
2. Questionnaires not self-completed by patients,
3. Feedback collected from a third party (e.g. patient's parents, peers, or staff),
4. Assessment of medical students' CSs,
5. Assessment of patient's satisfaction only,
6. Feedback given at the organizational level of a healthcare practice,
7. Multisource feedback with indistinguishable patient input,
8. Feedback results not used for individual professional development.

2.3. Study selection

Titles were independently screened by two reviewers; (HA) and (JD), to check their eligibility against the inclusion criteria. Abstract and full text screening were also carried out by two independent reviewers (HA and JD or MT). Any arising disagreements were resolved by discussion between the two reviewers, or by consulting a third reviewer (JD or MT) when necessary. Additional studies were identified by manually checking the reference lists of included studies/related systematic reviews and they were screened for eligibility using the same criteria.

2.4. Data extraction and quality assessment

A data extraction template was designed to extract the following data from each study where possible: (1) author(s) and publication year; (2) study design, setting and country; (3) patients' information (sample size, age, gender, and response rate); (4) practitioners' information (sample size & profession); (5) patient feedback questionnaire (name, care domains covered, psychometric properties, answer scale, administration method); (6) feedback reporting methods; (7) follow up and findings. The data extraction template was piloted using a small sample of studies. Data from each study was independently extracted by HA, and verified by JD for accuracy and completeness. Disagreements were resolved by discussion, or by consulting a third reviewer (MT) where necessary.

A quality assessment of included studies was carried out independently by two authors (HA and SS), disagreements were resolved through discussion. The assessment tool used was the NIH Quality Assessment Tool for Observational

Cohort and Cross Sectional Studies [34]. It is composed of 14 criteria that are answered either by “Yes”, “No”, “Not Applicable (NA)”, or “Not Reported (NR)”. Inherent to its design, cross-sectional studies automatically score NA on criteria 6, 7, 10 and 13. Additionally, studies could also score NA to criteria 8 as per tool’s instruction. Depending on the number of criteria met, a similar approach described by a previous study [35] was used in this review with respect to categorisation of included studies. Studies were categorized of “good” quality when meeting 10-14 criteria, of “fair” quality when meeting 5-9 criteria, or of “poor” quality when meeting 0-4 criteria. The higher the rating of a study, the lower the risk of bias [34].

2.5. Dealing with missing data

Linked publications were checked and corresponding authors were contacted by email. When no response was received, studies with essential missing data (e.g. missing questionnaire) were excluded.

3. Results

The search process is presented in the PRISMA flow diagram in Figure 1. A total of sixteen studies were included in the review.

3.1. Quality assessment

Table 2 provides a summary of quality assessment of included studies. Some studies were rated as “poor” (n=7) (score range 3-4) [36-42], and some were rated as “fair” (n=7) (score range 5-9) [18, 20, 43-47]. Only two studies had an overall rating of “good” (score range 11-12) [15, 48]. Several limitations were encountered including firstly sample sizes. Most studies did not provide justification for the chosen

sample size (n=13). However, most of these studies were of cross-sectional observational design [18, 20, 36-40, 42, 43, 45, 47], where a lack of sample size calculation does not represent a “fatal flaw” since such studies are exploratory in nature [34]. Secondly, results of many studies were not adjusted for confounders [18, 20, 36-39, 41-45, 47]. Thirdly, some studies did not provide sufficient description of exposure measures [36-38, 40-42], thus creating a difficulty in identifying the presence of an association between exposure and outcome. Additionally, outcome measures were not clearly defined in three studies [36, 38, 40], which thus may affect the validity of obtained results. Some degrees of selection bias were demonstrated by some studies [20, 39, 42] as two methods were used in recruiting patients with lack of clear exclusion criteria. Finally, some items of the assessment tool were not reported across the included studies.

3.2. General characteristics of included studies

Most included studies were cross sectional observational studies (n=13/16) [18, 20, 36-45, 47] and were conducted in a primary care setting (N=9/16) [15, 36, 37, 40-43, 46, 47]. Studies were conducted in five countries (UK, Canada, USA, Australia, and Netherlands) over a 27-years period (1986-2013).

3.3. Description of questionnaires

Table 3 provides a summary of the twelve different questionnaires identified by this review. Of these questionnaires, the Physician Achievement Review (PAR) [20, 38, 40, 46] and the Doctor Interpersonal Skills Questionnaire (DISQ) [15, 39, 43, 45] were the mostly reported to be used by the included studies. A 5-point Likert scale was the standard answer scale used by all except three questionnaires, where a 4-

point Likert scale [41, 42] or a binary scale [49] were used instead. The number of items composing the different questionnaires ranged between 10-40 items, and dedicating a space for patients' comments was only encountered with three questionnaires, DISQ [15, 39, 43, 45], 360-degree evaluation questionnaire [18], and Physicians Patient survey (PPS) [41]. Seven questionnaires showed evidence for at least one type of validity, including Patient Satisfaction Questionnaire (PSQ) [48], Consultation Satisfaction Questionnaire (CSQ) [21, 50], PAR [40, 46, 51], DISQ [52], 360-degree evaluation questionnaire [18], Patient Feedback Checklist (PFC) [42], and Multisource Feedback (MSF) [47]. Of the twelve questionnaires, nine were tested for reliability [18, 21, 37, 38, 44, 47, 48, 52, 53]. Assessing the internal consistency was used by eight questionnaires for testing reliability, including PSQ [48], CSQ [21], the Medical Interview Satisfaction Scale-26 (MISS-26) [53], North Worcestershire vocational training scheme PSQ [37], PAR [38], DISQ [52], 360-degree evaluation questionnaire [18], and MSF [47]. The reliability of one additional questionnaire (patient assessment) was tested by assessing its generalizability coefficient [44]. Of all questionnaires, DISQ was the only one that was tested for reliability and for the different types of validity, whereas no data was found regarding the psychometric properties of the Patient-Doctor Satisfaction Questionnaire (PDSQ) [49] and the PPS [41].

3.4. Description of participants

Table 4 summarises the general characteristics of included studies and description of participants. Physicians were the practitioners mostly assessed by patients of included studies, however, one study assessed occupational therapists [47], and nurses were also assessed in another study [39].

Regarding patient participants, all studies except one [41] reported patient sample size (range 55-28,156). Only two studies reported recruiting new patients [46, 48], whereas a mixture of old and new patients was described by other studies [15, 36, 42, 43, 45]. The average age of participants was only reported by six studies, with 67% being females lower than 60 years old. An average of 28 patients was used to assess each practitioner, with justifications based on providing reliable results [15, 41], selecting a patient sample size sufficient for the learning experience without being a burden [42], and overcoming the effects of a stressful day that may affect participants [45].

Patients were mostly recruited using consecutive sampling technique (n=8/16) [15, 36, 37, 39, 41-43, 48]. Patient response rate was only reported by six studies [20, 36-38, 47, 48], and the calculated average was 83%. One additional study reported the mean response rate per doctor at two different times [46].

3.5. Questionnaire administration and feedback reporting

A summary of questionnaire administration and feedback reporting are described in Table 5. Questionnaires were mostly given to patients either by a third person (n=7) [18, 38, 39, 41, 43, 46, 48], or by the practitioner (n=5) [15, 37, 40, 42, 47]. In two other studies, questionnaires were administered using an electronic device (touch-tone telephone) [44], or they were posted to patients [20], in both of these studies, patients were initially identified by their practitioners. Questionnaire administration was not described in the remaining two studies [36, 45]. Questionnaires were collected using methods that are cost neutral to participants (n=5) [20, 38-40, 43], or

they were collected by a third, independent person (n=2) [42, 45]. In one study, patients were given the choice of sending questionnaires by freepost if they were not collected immediately following the consultation [41], and in another study, patients were contacted by the research assistant to complete missing information by phone [48]. Questionnaire collection in one study was performed electronically using a touch-tone telephone system [44]. No data regarding questionnaire collection was mentioned in the remaining four studies [15, 18, 36, 37, 46, 47].

Individualized reports were used in all studies except one [41] to report feedback results to practitioners. The reports included the practitioner's individual scores, and for the purpose of comparison, some studies also included anonymised scores of peers [15, 18, 20, 36-40, 43-48]. In the remaining study [41], results were distributed to practitioners in an appraisal meeting, showing the average score of each question and number of patients answering it. Patient anonymity was protected in all but one study [42], where patients were asked for their date of birth.

3.6. Follow-up and impact of patient feedback

Table 6 describes the follow-up and impact of patient feedback of included studies. A follow-up was conducted by all except two studies [37, 41], focussing mostly on identifying practitioners' views about the feedback process and on detecting whether changes to individual practices were commenced/planned following the receipt of feedback reports. Described follow-up methods included asking practitioners to complete questionnaires [20, 44, 46, 47], join focus group discussions/individual interviews [18, 45], or undergo re-assessment by new patients [15, 46, 48]. Combined methods were also described [36, 38-40, 42, 43]. Follow-up was

conducted either shortly following the receipt of feedback reports [18, 36, 39, 40, 42-45], or months to years later [15, 20, 38, 46-48].

Practitioners' reassessment was mainly conducted by three studies [15, 46, 48].

Feedback in one study [48] was conducted 6 months following the initial assessment, during which practitioners with low initial feedback scores were privately interviewed to discuss their results and identify ways to enhance their performance. Study results showed significant improvement of practitioners' scores. In the second study [15], practitioners were randomly assigned into three groups where they were exposed to pre-assessment and to various frequencies of reassessments over 15 months period, including reassessment at 15 months (control group), and every 3-6 months for the two other intervention groups, one of which received supplementary feedback from practice supervisors. Study findings showed a higher improvement in CSs of practitioners in the two intervention groups compared to the control group, with sustained improvement achieved when reassessment is conducted at regular intervals. Reassessment was conducted 5 years following the initial one in the final study [46], and improvements in CSs were also demonstrated.

The follow-up conducted by the different studies demonstrated a generally positive influence of patient feedback experience on practitioners. Some studies illustrated that changes to individual practices of practitioners have started following the receipt of their feedback reports [38, 45, 48], and the intention to develop strategies of interaction with patients was also reflected by other practitioners in other studies [20, 40, 44]. Collecting feedback from patients was considered to be a learning

experience to help in professional development [36], and some practitioners were involved in additional training to further improve their skills [43, 46]. However, in one study [42], despite being initially enthusiastic, some practitioners expressed difficulties in fitting a patient feedback programme into their practices. In another study [46], improvements in professionals' performance were demonstrated, however, the effect size was likely to be small to moderate.

4. Discussion and conclusion

4.1. Discussion

4.1.1. Summary of main results

To our knowledge, this is the first systematic review that identified patient feedback questionnaires assessing CSs of all practitioners in all settings. Sixteen studies were identified. The majority of studies were similar in terms of their design, setting, methods of patient recruitment and feedback reporting to practitioners. Results were generally positive, however, they were mostly reflecting the views of practitioners regarding feedback process, without actually measuring the extent of CSs development that is induced by patient feedback, and whether it is statistically significant or not. The methodological quality of most studies ranged from poor to fair, with only two studies rating good. This not surprising as most of included studies were of cross-sectional design which has partly contributed to the final rating. Accordingly, some degrees of bias were identified in these studies, therefore the results should be interpreted with caution.

4.1.2. Questionnaires

The twelve questionnaires identified were designed across the past four decades (late 1970s to late 2000s). Questionnaire items were generated from different sources, including collecting patients and/or practitioners' views, reviewing literature, and examining other related questionnaires. Patients were involved in the design of most of the included questionnaires, however, their involvement was not clear for others (PSQ and MSF). Patients' involvement helps in highlighting areas that might not be recognised by other assessment methods [6, 22, 23], and in supporting the questionnaire's content validity [21, 52, 54, 55].

Only three questionnaires provided space for patients' comments. Providing such space was found to be welcomed by many patients [56, 57]. Most questionnaires were constructed of less than 25 items ($n=10/12$). This is important as lengthy questionnaires are less likely to be completed [58] and guidelines for questionnaire design indicate that answers to research questions can be collected using no more than 25 items [59]. A 5-point Likert scale was the most commonly reported, where respondents were given an option to provide a neutral response. Despite the controversy of using a neutral response, removing it will force respondents to select an answer they may not want [60, 61]. Additionally, 5-point Likert scale was found to be easier for use by patients [21, 62, 63], as many patients reported encountering difficulty handling a questionnaire with many response options [64].

With respect to psychometric properties, no publications were found regarding the validation of most of the included questionnaires, or the reliability of some of them. Validity and reliability are considered important qualities of a questionnaire that increase the confidence in their results [65]. DISQ was the only questionnaire to

have more evidence regarding its psychometric properties, it is reliable with high internal consistency, and significant test-retest measures [54]. DISQ was also tested for the different types of validity including construct, criterion, concurrent, and content validity [52], and it meets most of the requirements that are favourable in a questionnaire (short, provides a space for patients' comments, and uses a 5-point Likert scale). Moreover, mixed methods were used in its design including three different patient feedback questionnaires (CSQ, PDSQ, and MISS), followed by focus group discussions with patients and GPs [36, 54], allowing it thus to be a more comprehensive questionnaire. Furthermore, DISQ was used for doctors and nurses [39], and this makes it a promising to be taken forward and used with other practitioners.

4.1.3. Participants

Physicians were the practitioners mostly assessed by patients, especially in primary care settings. This can be attributed to their historical dominant role in consulting patients, however, currently almost all practitioners are involved in such activity, and collecting feedback from their patients will help in their development.

As for patients, most studies did not report full information, patients' sample from studies which did were mostly females under 60 years of age. It is unclear whether patient's gender may have influenced their participation in completing questionnaires, as female participation was found higher in some studies [13, 66-71], and lower in others [72-74]. Increased female participation could be attributed to differences between genders in utilizing healthcare services. Females were reported to utilize healthcare services more than males [75-77], and they have higher

consultation rates [78-80], which could be due to variations between genders in symptom reporting [81-83], and differences in reproductive biology [84, 85].

Consecutive sampling approach was described by most of the included studies in recruiting patients. Such sampling approach was described of being easy to apply [15, 41] and it could help in reducing selection bias [86, 87]. As for the number of patients recruited, a range of at least 25-50 patients per practitioner (PPP) was suggested by the included studies, however, a minimum number of 25 PPP seems to be suitable to provide reliable data, specifically when using DISQ [70].

4.1.4. Questionnaire administration

Questionnaires were mostly administered by a third person, or by the practitioner. Practitioners' involvement has made them not blind to the process and might have encouraged them to behave differently, which is known as the Hawthorne effect [88]. Additionally, it might have also influenced the feedback given by patients, encouraging them to give more favourable responses that could please their practitioners. Patient responses seem to be influenced by the mode of questionnaire administration [89]. It is hence recommended for practitioners to be blinded in order to avoid biased performances and thus biased evaluations [90], and for questionnaires to be given by an independent person, as this will help in eliminating the unconscious influence of practitioner-patient relationship, thus avoids giving candid feedback by patients [89]. Patients' anonymity must also be protected, to avoid difficulties encountered with filling questionnaires that are not anonymous [42, 57], and to help in collecting more honest responses, thus reducing response or social desirability bias [91, 92].

As for collection method, questionnaires were mostly collected by an independent individual, either immediately following the encounter, or they were sent back by patients to a designated address by mail. Encouraging patients to complete questionnaires immediately following the encounter is advisable, since their recollection of details is still fresh, thus minimizing recall bias, besides, some evidence suggests that taking questionnaires home can discourage patients from completing questionnaires, or reduces the quality of collected data [57, 93]. Alternatively, using prepaid envelopes was found to encourage questionnaire return and to increase response rates [94, 95].

4.1.5. Response rate

Several factors might have influenced response rates, including questionnaire administration by the practitioner, especially that most recruited patients were not new [15, 20, 37, 38, 40, 42, 44, 47]. Using a face-to-face approach in patient recruitment was reported of providing higher response rates than by using other means of recruitment, such as using mail [96], and patients' interest in the subject of the questionnaire [94]. Included studies were aiming to enhance CSs of practitioners as guided by patients' views, and this may have given patients the sense of contribution to healthcare reforms. However, we could not establish this link, since response rates were not reported by all studies, and one study reported aggregated response rate from using three different questionnaires [36]. Moreover, the highest response rate was reported by a study that used a 40-items questionnaire, whereas the lowest response rate was associated with using a 14-items questionnaire. Both of these studies showed similarities in aspects related to questionnaire

administration and patient anonymity, but differed in their questionnaire collection, which does not explain the difference in their reported response rates. Other factors may exist to have encouraged patients to respond to one questionnaire more than the other.

4.1.6. Format of patient feedback report

Individualized reports were mostly described in reporting feedback results to the practitioner. Most reports included individual scores of the practitioner and, for the purpose of comparison, it also included the anonymised scores of their peers. Results were presented using numbers (e.g. mean and SD), graphical formats and tables. Presenting results using combined methods was found helpful for practitioners to better understand their feedback scores, especially when benchmarks for best practice were also provided [97].

4.1.7. Follow-up to patient feedback reports

Follow-up conducted by most of the included studies mainly focused on collecting practitioners' views about the feedback process, and identifying skills needing further development, and whether changes were commenced or planned. Follow-ups were conducted either immediately, or weeks to years later. Practitioners' views were generally positive and most have welcomed receiving feedback from their patients, however, most of the studies did not measure the impact of patient feedback reports on CSs development.

Three studies described repeating the whole process for the same practitioners using new patients, however, they differed in the time span for repeating the

assessment process, and in the other interventions that were included to enhance practitioners' development. The results of these studies showed an improvement in CSs of practitioners, with similar levels of improvements achieved when reassessment was repeated once 5 years later to when it was repeated several times regularly over a shorter period of time (months). Thus, similar results of improved scores of CSs of practitioners could be achieved by repeating the assessment process months or years following the initial one, however, this requires multiple points of reassessment to be conducted at regular intervals for the purpose of reinforcing skills development.

4.1.8. Agreements and disagreements with other reviews

The findings of this systematic review demonstrate that some evidence exists regarding the use of patient feedback in enhancing CSs of individual practitioners, however, further studies are needed to exactly measure the significance that patient feedback has in CSs development. This is consistent with the findings of two other systematic reviews [30, 98], however, in contrast to Evans et al. [29] systematic review, where practitioners were found to show some resistance towards seeking patient feedback, most practitioners in this review were positive about receiving patient feedback. Patients were regarded by some to be the most appropriate raters to assess their practice [40], and some practitioners have the desire to continue seeking feedback from their patients [20, 39, 44].

4.1.9. Strengths and weaknesses of the review

A number of elements exist that strengthen the confidence with the findings of this systematic review. This review used a standard approach to systematic reviews as

outlined by the Cochrane Database of Systematic Reviews (CDSR) [99]. The results of this review were based on searching for the best available evidence by using a combination of complementary key words to systematically search all related databases. The search was also widened to cover the bibliographies of all included studies and related systematic reviews, in addition to searching grey literature. Moreover, no time restriction was made by this review to capture all possible evidence across the years. However, some limitations were encountered with this review, which include missing data from the included studies. Attempts to contact corresponding authors for missing data were unsuccessful, and this prevented proper comparisons to be made, and even led to rejecting some studies that were identified from the bibliographies of other included studies [51, 100, 101]. Additionally, the search strategy employed was limited to English language, leading to possibly rejecting some useful questionnaires that were not written in English.

4.2. Conclusion

The review identified gaps in literature regarding the use patient feedback questionnaires for a wider range of practitioners and in different healthcare settings. Most included studies had a poor to fair methodological quality which hinders making firm conclusions. The evidence that is shown so far indicates that it is feasible to use patient feedback, however, the impact it has on CSs development is still not clear as it has not been thoroughly examined, thus, more higher quality studies with clearly defined methods are needed in order to identify its real impact in improving CSs of different practitioners. Additionally, most of the identified questionnaires lacked validation and/or reliability, thus hindering the confidence in their results. The

recommendations that we provide in this review can guide future studies in examining patient feedback as a tool for CSs development.

4.3. Practice implications

Based on our review, we suggest the following recommendations to be used by future studies to help in increasing the robustness of patient feedback in CSs development:

(1) Using valid and reliable questionnaire in feedback collection (e.g. DISQ with at least 25 patient per practitioner), (2) Recruiting patients by an independent person, preferably by using face-to-face approach, (3) Collecting patient feedback over more than one day, while protecting patient anonymity and maintaining practitioners' blindness, (4) Collecting questionnaires from patients immediately following the encounter or providing them with prepaid envelopes to encourage questionnaire return, (5) Reporting feedback results to practitioners by using a combined method of data presentation that allows comparison with peers, (6) Conducting follow-up that includes reassessment of practitioners by new patients at regular intervals to detect changes in CSs and measure how significant changes are.

Author contributions

HA, MT and JD contributed to the design of the review. HA and SS assessed the quality of included studies. HA drafted the manuscript and all authors revised and approved its final version.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

None

ACCEPTED MANUSCRIPT

References

- [1] C. Butler, S. Rollnick, N. Stott. The practitioner, the patient and resistance to change: recent ideas on compliance. *Can Med Assoc J* 1996; 154: 1357-62.
- [2] D.G. Safran, D.A. Taira, W.H. Rogers, M. Kosinski, J.E. Ware, A.R. Tarlov. Linking primary care performance to outcomes of care. *J Fam Pract* 1998; 47: 213-21.
- [3] P. Kinnersley, N. Stott, T.J. Peters, I. Harvey. The patient-centredness of consultations and outcome in primary care. *Br J Gen Pract* 1999; 49: 711-6.
- [4] R.C. Maly, L.B. Bourque, R.F. Engelhardt. A Randomized Controlled Trial of Facilitating Information Giving to Patients with Chronic Medical Conditions Effects on Outcomes of Care. *J Fam Pract* 1999; 48: 356-.
- [5] S. Svensson, K.I. Kjellgren, J. Ahlner, R. Säljö. Reasons for adherence with antihypertensive medication. *Int J Cardiol* 2000; 76: 157-63.
- [6] A. Bredart, C. Bouleuc, S. Dolbeault. Doctor-patient communication and satisfaction with care in oncology. *Curr Opin Oncol* 2005; 17: 351-4.
- [7] General Medical Council. Colleague and patient feedback for revalidation. http://www.gmc-uk.org/doctors/revalidation/colleague_patient_feedback.asp, (Accessed 24 July 2017).
- [8] A. Brownlea. Earning confidence: perspectives for a modernising NHS. *J Clinical Excellence* 2001; 3: 27-32.
- [9] General Medical Council, Supporting information for appraisal and revalidation, GMC, Manchester, UK, 2012.
- [10] R.J. Howells, H.A. Davies, J.D. Silverman, J.C. Archer, A.F. Mellon. Assessment of doctors' consultation skills in the paediatric setting: the Paediatric Consultation Assessment Tool. *Brit Med J* 2010; 95: 323-9.

- [11] P.G. Ramsey, M.D. Wenrich, J.D. Carline, T.S. Inui, E.B. Larson, J.P. LoGerfo. Use of peer ratings to evaluate physician performance. *J Amer Med Assoc* 1993; 269: 1655-60.
- [12] J.J. Norcini. Peer assessment of competence. *Med Educ* 2003; 37: 539-43.
- [13] J.L. Campbell, S.H. Richards, A. Dickens, M. Greco, A. Narayanan, S. Brearley. Assessing the professional performance of UK doctors: an evaluation of the utility of the General Medical Council patient and colleague questionnaires. *Qual Saf Health Care* 2008; 17: 187-93.
- [14] M. Greco, W. Francis, J. Buckley, A. Brownlea, J. McGovern. Real-patient evaluation of communication skills teaching for GP registrars. *Fam Pract* 1998; 15: 51-7.
- [15] M. Greco, A. Brownlea, J. McGovern. Impact of patient feedback on the interpersonal skills of general practice registrars: results of a longitudinal study. *Med Educ* 2001; 35: 748-56.
- [16] A.G. Espinel, R.K. Shah, M.C. Beach, E.F. Boss. What parents say about their child's surgeon: parent-reported experiences with pediatric surgical physicians. *J Amer Med Assoc* 2014; 140: 397-402.
- [17] J.M. Stausmire, C.P. Cashen, L. Myerholtz, N. Buderer. Measuring general surgery residents' communication skills from the patient's perspective using the Communication Assessment Tool (CAT). *J Surg Educ* 2015; 72: 108-16.
- [18] J. Wood, J. Collins, E.S. Burnside, M.A. Albanese, P.A. Propeck, F. Kelcz, J.M. Spilde, L.M. Schmaltz. Patient, faculty, and self-assessment of radiology resident performance:: A 360-degree method of measuring professionalism and interpersonal/communication skills. *Acad Radiol* 2004; 11: 931-9.

- [19] F. Kamangar, P. Davari, K.K. Parsi, C.-S. Li, Q. Wang, S. Mathis, N. Fazel. 360-degree Evaluations on Physician Performance as an Effective Tool for Interprofessional Teams: A critical analysis of physician self-assessment as compared to nursing staff and patient evaluations of providers. *Dermatol Online J* 2016; 22.
- [20] S.K. Vinod, D.M. Lonergan. Multisource feedback for radiation oncologists. *J Med Imaging Radiat Oncol* 2013; 57: 384-9.
- [21] R. Baker. Development of a questionnaire to assess patients' satisfaction with consultations in general practice. *Br J Gen Pract* 1990; 40: 487-90.
- [22] J. Labarere, P. Francois, P. Auquier, C. Robert, M. Fourny. Development of a French inpatient satisfaction questionnaire. *Int J Qual Health Care* 2001; 13: 99-108.
- [23] E. Zarei. Service quality of hospital outpatient departments: patients' perspective. *Int J Health Care Qual Assur* 2015; 28: 778-90.
- [24] P.D. Cleary. The increasing importance of patient surveys. *Brit Med J* 1999; 319: 720-21.
- [25] M. Wensing, E. Vingerhoets, R. Grol. Feedback based on patient evaluations: a tool for quality improvement? *Patient Educ Couns* 2003; 51: 149-53.
- [26] T.L. Delbanco. Enriching the doctor-patient relationship by inviting the patient's perspective. *Ann Intern Med* 1992; 116: 414-8.
- [27] K. Tasa, G.R. Baker, M. Murray. Using patient feedback for quality improvement. *Qual Manag Health Care* 1996; 4: 55-67.
- [28] M.N. Marshall, P.G. Shekelle, S. Leatherman, R.H. Brook. The public release of performance data. *J Amer Med Assoc* 2000; 283: 1866-74.

- [29] R.G. Evans, A. Edwards, S. Evans, B. Elwyn, G. Elwyn. Assessing the practising physician using patient surveys: a systematic review of instruments and feedback methods. *Fam Pract* 2007; 24: 117-27.
- [30] M.E. Reinders, B.L. Ryan, A.H. Blankenstein, H.E. van der Horst, M.A. Stewart, H.W.J. van Marwijk. The effect of patient feedback on physicians' consultation skills: a systematic review. *Acad Med* 2011; 86: 1426-36.
- [31] A.P. Tonna, D. Stewart, B. West, D. McCaig. Pharmacist prescribing in the UK— a literature review of current practice and research. *J Clin Pharm Ther* 2007; 32: 545-56.
- [32] R. Cooper, L. Guillaume, T. Avery, C. Anderson, P. Bissell, A. Hutchinson, J. Lymn, E. Murphy, P. Ward, J. Ratcliffe. Nonmedical prescribing in the United kingdom: developments and stakeholder interests. *J Ambul Care Manage* 2008; 31: 244-52.
- [33] Department of Health, Improving patients' access to medicines: a guide to implementing nurse and pharmacist independent prescribing within the NHS in england, Department of Health, 2008.
- [34] National Institutes of Health. National Heart Lung and Blood Institute, Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies - NHLBI, NIH. <https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/cohort.>, (Accessed 10/01/2018).
- [35] M.H. Woolford, C. Weller, J.E. Ibrahim. Unexplained Absences and Risk of Death and Injury Among Nursing Home Residents: A Systematic Review. *Journal of the American Medical Directors Association* 2017; 18: 366. e1-. e15.
- [36] M.J. Greco, A.A. Brownlea, J.M. McGovern. Utilising patient feedback in the RACGP training program. An exploratory study. *Aust Fam Physician* 1995; 24: 1077.

- [37] M. Jenkins, A. Thomas. The assessment of general practitioner registrars' consultations by a patient satisfaction questionnaire. *Med Teach* 1996; 18: 347-50.
- [38] W. Hall, C. Violato, R. Lewkonia, J. Lockyer, H. Fidler, J. Toews, P. Jennett, M. Donoff, D. Moores. Assessment of physician performance in Alberta the physician achievement review. *Can Med Assoc J* 1999; 161: 52-7.
- [39] M. Greco, K. Sweeney, J. Broomhall, P. Beasley. Patient assessment of interpersonal skills: a clinical governance activity for hospital doctors and nurses. *J Clinical Excellence* 2001; 3: 117-24.
- [40] J.M. Sargeant, K.V. Mann, S.N. Ferrier, D.B. Langille, P.D. Muirhead, V.M. Hayes, D.E. Sinclair. Responses of rural family physicians and their colleague and coworker raters to a multi-source feedback process: a pilot study. *Acad Med* 2003; 78: S42-S4.
- [41] L. Mackillop, M. Armitage, W. Wade. Collaborating with patients and carers to develop a patient survey to support consultant appraisal and revalidation. *Clinician in Management* 2006; 14: 89-94.
- [42] M.E. Reinders, A.H. Blankenstein, H.J. van Marwijk, H. Schleypen, P.L. Schoonheim, W.A.B. Stalman. Development and feasibility of a patient feedback programme to improve consultation skills in general practice training. *Patient Educ Couns* 2008; 72: 12-9.
- [43] M. Greco, S. Pocklington. Incorporating patient feedback into vocational training: An interpersonal skills development exercise for GP trainers and registrars. *Educ Prim Care* 2001; 12: 285-91.
- [44] R.S. Lipner, L.L. Blank, B.F. Leas, G.S. Fortna. The value of patient and peer ratings in recertification. *Acad Med* 2002; 77: S64-S6.

- [45] A.K. Al-Shawi, A.G. MacEachern, M.J. Greco. Patient assessment of surgeons' interpersonal skills: A tool for appraisal and revalidation. *Clinical Governance: An International Journal* 2005; 10: 212-6.
- [46] C. Violato, J.M. Lockyer, H. Fidler. Changes in performance: a 5-year longitudinal study of participants in a multi-source feedback programme. *Med Educ* 2008; 42: 1007-13.
- [47] C. Violato, L. Worsfold, J.M. Polgar. Multisource feedback systems for quality improvement in the health professions: Assessing occupational therapists in practice. *J Contin Educ Health Prof* 2009; 29: 111-8.
- [48] D.W. Cope, L.S. Linn, B.D. Leake, P.A. Barrett. Modification of residents' behavior by preceptor feedback of patient satisfaction. *J Gen Intern Med* 1986; 1: 394-8.
- [49] A. Rashid, W. Forman, C. Jagger, R. Mann. Consultations in general practice: a comparison of patients' and doctors' satisfaction. *Brit Med J* 1989; 299: 1015-6.
- [50] R. Baker, M. Whitfield. Measuring patient satisfaction: a test of construct validity. *Qual Saf Health Care* 1992; 1: 104-9.
- [51] C. Violato, A. Marini, J. Toews, J. Lockyer, H. Fidler. Using peers, self, patients and co-workers to assess physician performance. *Acad Med* 1997; 72: 582-4.
- [52] M. Greco, M. Cavanagh, A. Brownlea, J. McGovern. The Doctor's Interpersonal Skills Questionnaire (DISQ): a validated instrument for use in GP training. *Educ Gen Pract* 1999; 10: 256-64.
- [53] M.H. Wolf, S.M. Putnam, S.A. James, W.B. Stiles. The Medical Interview Satisfaction Scale: development of a scale to measure patient perceptions of physician behavior. *J Behav Med* 1978; 1: 391-401.

- [54] M. Greco, A. Brownlea, J. McGovern, M. Cavanagh. Consumers as educators: implementation of patient feedback in general practice training. *Health Commun* 2000; 12: 173-93.
- [55] B. Kitchenham, S.L. Pfleeger. Principles of survey research part 4: questionnaire evaluation. *SEN* 2002; 27: 20-3.
- [56] J. Rattray, M.C. Jones. Essential elements of questionnaire design and development. *J Clin Nurs* 2007; 16: 234-43.
- [57] L. Land, S. Sizmur, J. Harding, J.D.C. Ross. Development of a validated patient satisfaction survey for HIV clinic attendees. *Int. J. STD AIDS* 2013; 24: 201-9.
- [58] J. Fox. Designing research: basics of survey construction. *Minim Invasive Surg Nurs* 1993; 8: 77-9.
- [59] C. Passmore, A.E. Dobbie, M. Parchman, J. Tysinger. Guidelines for constructing a survey. *Fam Med* 2002; 34: 281-6.
- [60] B.A. Kitchenham, S.L. Pfleeger. Principles of survey research: part 3: constructing a survey instrument. *SEN* 2002; 27: 20-4.
- [61] N. Burns, S.K. Grove, *The practice of nursing research: Conduct, critique, and utilization*, Elsevier Saunders, 2005.
- [62] R. Grover, M. Vriens, *The handbook of marketing research: uses, misuses, and future advances*, Sage 2006.
- [63] L. Nicole. The Likert Scale: Advantages and Disadvantages. <https://psyc450.wordpress.com/2011/12/05/the-likert-scale-advantages-anddisadvantages>, (Accessed 18 May 2017).
- [64] N.K. Malhotra, *Questionnaire design and scale development*, 2006.

- [65] L.J. Burton, S.M. Mazerolle. Survey instrument validity part I: Principles of survey instrument development and validation in athletic training education research. *Athl Train Educ J* 2011; 6: 27-35.
- [66] J.L. Campbell, J. Ramsay, J. Green. Age, gender, socioeconomic, and ethnic differences in patients' assessments of primary health care. *Qual Saf Health Care* 2001; 10: 90-5.
- [67] K. Korkeila, S. Suominen, J. Ahvenainen, A. Ojanlatva, P. Rautava, H. Helenius, M. Koskenvuo. Non-response and related factors in a nation-wide health survey. *Eur J Epidemiol* 2001; 17: 991-9.
- [68] M. Oremus, C. Wolfson. Female specialists were more likely to respond to a postal questionnaire about drug treatments for Alzheimer disease. *J Clin Epidemiol* 2004; 57: 620-3.
- [69] M. Potiriadis, P. Chondros, G. Gilchrist, K. Hegarty, G. Blashki, J.M. Gunn. How do Australian patients rate their general practitioner? A descriptive study using the General Practice Assessment Questionnaire. *Med J Aust* 2008; 189: 215.
- [70] J. Campbell, A. Narayanan, B. Burford, M. Greco. Validation of a multi-source feedback tool for use in general practice. *Educ Prim Care* 2010; 21: 165-79.
- [71] M. Roland, M. Roberts, V. Rhenius, J. Campbell. GPAQ-R: development and psychometric properties of a version of the General Practice Assessment Questionnaire for use for revalidation by general practitioners in the UK. *BMC Fam Pract* 2013; 14: 160.
- [72] P. Meredith, C. Wood. Aspects of patient satisfaction with communication in surgical care: confirming qualitative feedback through quantitative methods. *Int J Qual Health Care* 1996; 8: 253-64.

- [73] K. Christensen, N.V. Holm, M. Mcgue, L. Corder, J.W. Vaupel. A Danish population-based twin study on general health in the elderly. *J Aging Health* 1999; 11: 49-64.
- [74] N. Kwak, B. Radler. A comparison between mail and web surveys: Response pattern, respondent profile, and data quality. *J Off Stat* 2002; 18: 257.
- [75] K.D. Bertakis, R. Azari, L.J. Helms, E.J. Callahan, J.A. Robbins. Gender differences in the utilization of health care services. *J Fam Pract* 2000; 49: 147-.
- [76] S. Kaur, K.M. Stechuchak, C.J. Coffman, K.D. Allen, L.A. Bastian. Gender differences in health care utilization among veterans with chronic pain. *J Gen Intern Med* 2007; 22: 228.
- [77] V. Vaidya, G. Partha, M. Karmakar. Gender differences in utilization of preventive care services in the United States. *J Womens Health* 2012; 21: 140-5.
- [78] A. Rogers, K. Hassell, G. Nicholas, *Demanding patients? Analysing the use of primary care*, Open University Press, Buckingham, 1999.
- [79] S. Rowlands, K. Moser. Consultation rates from the general practice research database. *Br J Gen Pract* 2002; 52: 658-60.
- [80] J. Hippisley-Cox, Y. Vinogradova, Final report to the NHS Information Centre and Department of Health Trends in consultation rates in general practice 1995 to 2008: Analysis of the QResearch® database, Q Research and the Information Centre for Health and Social Care, London, 2009.
- [81] K.-H. Ladwig, B. Marten-Mittag, B. Formanek, G. Dammann. Gender differences of symptom reporting and medical health care utilization in the German population. *Eur J Epidemiol* 2000; 16: 511-8.
- [82] D. Mechanic. Sex, illness, illness behavior, and the use of health services. *Soc Sci Med Part B: Medical Anthropology* 1978; 12: 207-14.

- [83] A. Oksuzyan, K. Juel, J.W. Vaupel, K. Christensen. Men: good health and high mortality. *Sex differences in health and aging. Aging Clin Exp Res* 2008; 20: 91.
- [84] I. Waldron. Sex differences in illness incidence, prognosis and mortality: issues and evidence. *Soc Sci Med* 1983; 17: 1107-23.
- [85] C.G. van Wijk, A.M. Kolk, W.J.H.M. van Den Bosch, H.J.M. Van Den Hoogen. Male and female morbidity in general practice: the nature of sex differences. *Soc Sci Med* 1992; 35: 665-78.
- [86] D.L. Maxwell, E. Satake, *Research and statistical methods in communication sciences and disorders*, Singular Publishing Group 2006.
- [87] J. Daniel, *Sampling essentials: Practical guidelines for making sampling choices*, Sage 2011.
- [88] A. Indrayan, *Varieties of bias to guard against*, 2014.
- [89] C. Cook, *Mode of administration bias*, Taylor & Francis, 2010, pp. 61-3.
- [90] S.J. Pocock, *Clinical trials: a practical approach*, John Wiley & Sons 2013.
- [91] D. Colton, R.W. Covert, *Designing and constructing instruments for social research and evaluation*, John Wiley & Sons 2007.
- [92] M.L. Mitchell, J.M. Jolley, *Research design explained*, Cengage Learning 2012.
- [93] D.L. Streiner, G.R. Norman, *Health Measurement Scales: A Practical Guide to Their Development and Use*, Medical Publications, Oxford, 2003.
- [94] P. Edwards, I. Roberts, M. Clarke, C. DiGiuseppi, S. Pratap, R. Wentz, I. Kwan. Increasing response rates to postal questionnaires: systematic review. *Brit Med J* 2002; 324: 1183.
- [95] D.L. Streiner, G.R. Norman, J. Cairney, *Health measurement scales: a practical guide to their development and use*, Oxford University Press, USA, 2014.

- [96] J. Sitzia, N. Wood. Response rate in patient satisfaction research: an analysis of 210 published studies. *Int J Qual Health Care* 1998; 10: 311-7.
- [97] M. Gysels, R. Hughes, F. Aspinal, J.M. Addington-Hall, I.J. Higginson. What methods do stakeholders prefer for feeding back performance data: a qualitative study in palliative care. *Int J Qual Health Care* 2004; 16: 375-81.
- [98] S. Cheraghi-Sohi, P. Bower. Can the feedback of patient assessments, brief training, or their combination, improve the interpersonal skills of primary care physicians? A systematic review. *BMC* 2008; 8.
- [99] Cochrane Library. Cochrane Database of Systematic Reviews.
<http://www.cochranelibrary.com/cochrane-database-of-systematic-reviews/>,
(Accessed 27 August 2017).
- [100] D. Falvo. Patient perception as a tool for evaluation and feedback in family practice resident training. *J Fam Pract* 1980; 10: 471-4.
- [101] C. Violato, J. Lockyer, H. Fidler. Multisource feedback: a method of assessing surgical practice. *Brit Med J* 2003; 326: 546.
- [102] H. Fidler, J.M. Lockyer, J. Toews, C. Violato. Changing physicians' practices: the effect of individual feedback. *Acad Med* 1999; 74: 702-14.

Figures

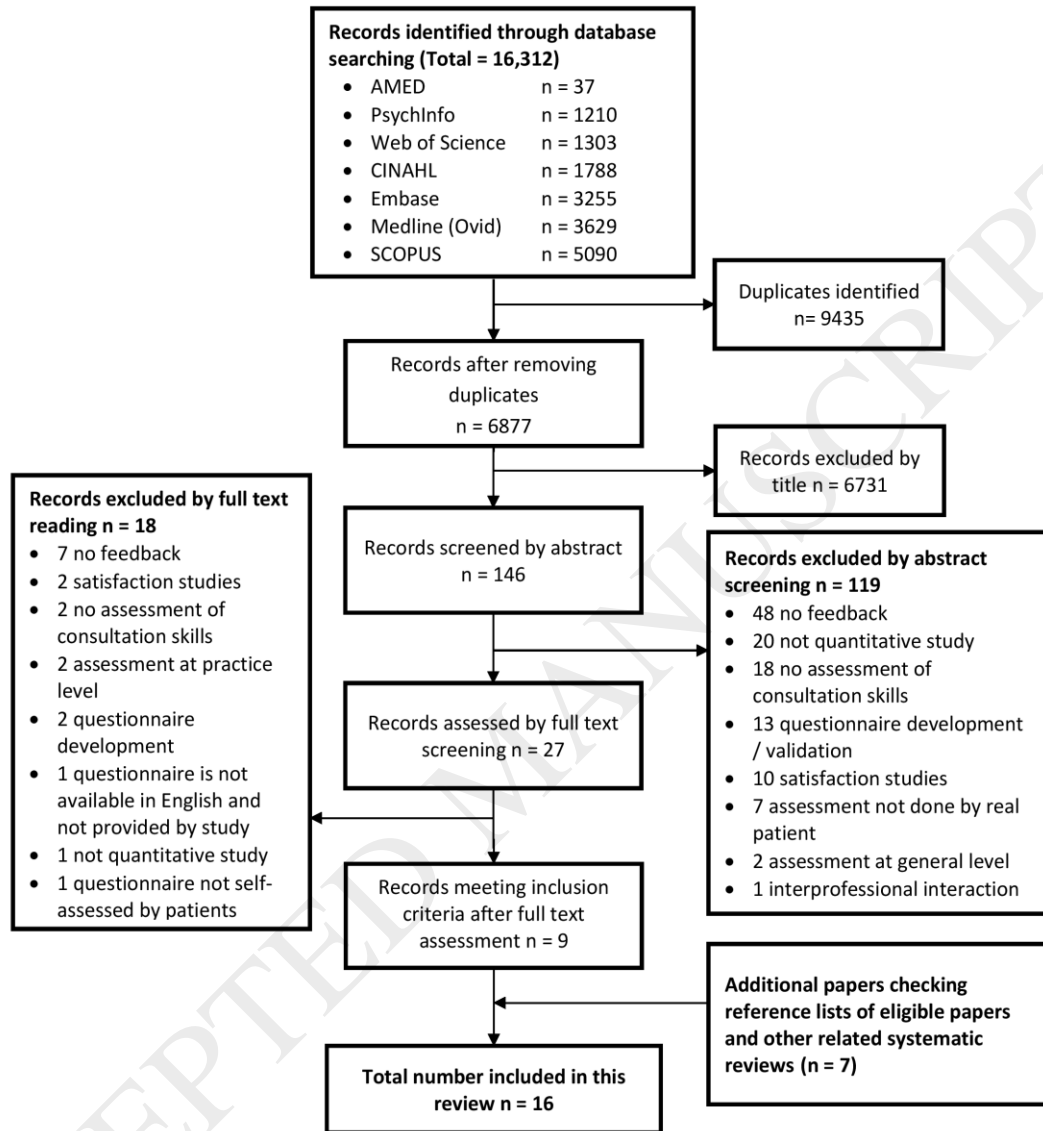


Fig.1 PRISMA Flow Diagram

Table 1

Example Search Strategy using Medline.

	Search	Results
1	"patient satisfaction".mp.	89133
2	("health?care professionals" or "general practitioner" or doctor or physician or nurse* or pharmac*).mp	1466359
3	(Feedback or questionnaire* or assessment or instrument or "evaluation tool" or survey or "performance appraisal" or "resident evaluation" or "performance feedback").mp	2158301
4	("interpersonal skills" or "communication skills" or "consultation skills" or "professional competence" or competence or consult* or communication).mp	599945
5	1 and 2 and 3 and 4	3938
6	limit 5 to (English language and journal article)	3629

Table 2
Methodological quality assessment

Study	The NIH quality assessment tool's criteria														Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Cope et al. (1986) [48]	Yes	Yes	Yes	Yes	No	Yes	Yes	NA	Yes	Yes	Yes	Yes	Yes	Yes	12/14
Greco et al. (1995) [36]	Yes	Yes	Yes	Yes	No	NA	NA	NA	No	NA	No	NA	NA	No	4/14
Jenkins and Thomas (1996) [37]	No	No	Yes	Yes	No	NA	NA	NA	No	NA	Yes	NA	NA	No	3/14
Hall et al. (1999) [38]	Yes	Yes	Yes	Yes	No	NA	NA	NA	No	NA	No	NA	NA	No	4/14
Greco et al. (2001) [39]	Yes	Yes	NR	No	NR	NA	NA	NA	Yes	NA	Yes	NA	NA	No	4/14
Greco and Pocklington (2001) [43]	Yes	Yes	NR	Yes	No	NA	NA	NA	Yes	NA	Yes	NA	NA	No	5/14
Greco et al. (2001) [15]	Yes	Yes	NR	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NR	No	Yes	11/14
Lipner et al. (2002) [44]	Yes	Yes	NR	NR	Yes	NA	NA	NA	Yes	NA	Yes	NA	NA	No	5/14
Sargeant et al. (2003) [40]	Yes	Yes	NR	NR	No	NA	NA	NA	No	NA	No	NA	NA	Yes	3/14
Wood et al. (2004) [18]	Yes	Yes	NR	Yes	No	NA	NA	NA	Yes	NA	Yes	NA	NA	No	5/14

Table 2 (Continued)

Study	The NIH quality assessment tool's criteria														Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Al-Shawi et al. (2005) [45]	Yes	Yes	NR	Yes	No	NA	NA	NA	Yes	NA	Yes	NA	NA	No	5/14
Mackillop et al. (2006) [41]	Yes	No	NR	Yes	Yes	NA	NA	NA	No	NA	Yes	NA	NA	No	4/14
Reinders et al. (2008) [42]	Yes	Yes	NR	No	No	NA	NA	NA	No	NA	Yes	NA	NA	No	3/14
Violato et al. (2008) [46]	Yes	Yes	No	Yes	No	Yes	Yes	NA	Yes	Yes	Yes	NR	NR	Yes	9/14
Violato et al. (2009) [47]	Yes	Yes	Yes	NR	No	NA	NA	NA	Yes	NA	Yes	NA	NA	No	5/14
Vinod and Lonergan (2013) [20]	Yes	Yes	Yes	No	No	NA	NA	NA	Yes	NA	Yes	NA	NA	No	5/14

Options Yes/No/CD (cannot determine)/NA (not applicable)/NR (not reported). Tool's criteria: 1. Was the research question or objective in this paper clearly stated?, 2. Was the study population clearly specified and defined?, 3. Was the participation rate of eligible persons at least 50%?, 4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?, 5. Was a sample size justification, power description, or variance and effect estimates provided?, 6. For the analyses in this paper, were the exposure(s) of interest measured prior to the outcome(s) being measured?, 7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?, 8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (e.g., categories of exposure, or exposure measured as continuous variable)?, 9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?, 10. Was the exposure(s) assessed more than once over time?, 11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?, 12. Were the outcome assessors blinded to the exposure status of participants?, 13. Was loss to follow-up after baseline 20% or less?, 14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure(s) and outcome(s)?.

Table 3

General characteristics of patient feedback questionnaires.

Study (year) [reference]	Questionnaire name, number of items, answer scale	Space for patients' comments	Validity	Reliability
Cope et al. (1986) [48]	PSQ: 14 items 5-point Likert scale	No	Criterion predictive validity	Cronbach's α between 0.81-0.92)
Greco et al. (1995) [36]	CSQ : 18 items 5-point Likert scale	CSQ : No	Content ^[21] & construct ^[50] validity	Cronbach's $\alpha = 0.91$) ^[21]
	PDSQ : 13 items Binary scale (Yes/No) ^[49]	PDSQ : No ^[49]	No	No
	MISS : 26 items 5-point Likert scale	MISS : No ^[53]	No	Cronbach's $\alpha = 0.93$ ^[53]
Jenkins and Thomas (1996) [37]	North Worcestershire Vocational Training Scheme PSQ: 11 items 5-point Likert scale	No	No	Cronbach's $\alpha = 0.84$
Hall et al. (1999) [38]	PAR: 40 items* 5-point Likert scale	No	Content ^[40] & construct validity ^[46, 51]	Cronbach's α for patients' questionnaire = 0.95
Greco et al. (2001) [39]	DISQ/NISQ: 12 items 5-point Likert scale	Yes	All types of validity ^[52]	Cronbach's $\alpha = 0.96$ ^[52]
Greco and Pocklington (2001) [43]	DISQ: 12 items 5-point Likert scale	Yes	All types of validity ^[52]	Cronbach's $\alpha = 0.96$ ^[52]
Greco et al. (2001) [15]	DISQ: 12 items 5-point Likert scale	Yes	All types of validity ^[52]	Cronbach's $\alpha = 0.96$ ^[52]

Table 3 (Continued)

Study (year) [reference]	Questionnaire name, number of items, answer scale	Space for patients' comments	Validity	Reliability
Lipner et al. (2002) [44]	Patient assessment (ABIM/CPD): 10 items 5-point Likert scale	No	No	Generalizability coefficient = 0.67
Sargeant et al. (2003) [40]	PAR: 40 items 5-point Likert scale	No	Content ^[40] & construct validity ^[46, 51]	Cronbach's $\alpha > 0.90$ ^[38]
Wood et al. (2004) [18]	360-degree: 10 items 5-point Likert scale	Yes	Concurrent validity	Cronbach's $\alpha = 0.86$
Al-Shawi et al. (2005) [45]	DISQ: 12 items 5-point Likert scale	Yes	All types of validity ^[52]	Cronbach's $\alpha = 0.96$ ^[52]
Mackillop et al. (2006) [41]	PPS: 11 items 4-point Likert scale	Yes	No	No
Reinders et al. (2008) [42]	PFC: 14 items 4-point Likert scale	No	Content and face validity	No
Violato et al. (2008) [46]	PAR: 40 items 5-point Likert scale	No	Content ^[40] & construct validity ^[46, 51]	Cronbach's $\alpha > 0.90$ ^[38]
Violato et al. (2009) [47]	MSF: 14 items 5-point Likert scale	No	Content & construct validity	Cronbach's $\alpha = 0.93$
Vinod et al. (2013) [20]	MSF/PAR: 40 items 5-point Likert scale	No	Content ^[40] & construct validity ^[46, 51]	Cronbach's $\alpha > 0.90$ ^[38]

PSQ: Patient Satisfaction Questionnaire; CSQ: Consultation Satisfaction Questionnaire; PDSQ: Patient-Doctor Satisfaction Questionnaire; MISS: Medical Interview Satisfaction Scale; PAR: Physician Achievement Review; DISQ: Doctor Interpersonal Skills Questionnaire; PPS: Physicians Patient Survey; PFC: Patient Feedback Checklist; MSF: Multisource Feedback. * PAR questionnaire is described of having 44 items [38], however we confirmed from other references [20, 40, 46, 102] that it's composed of 40-items.

Table 4

General characteristics of included studies and description of participants.

Author(s), (Year), Country, [Reference]	Study design Setting	Healthcare professionals	Patients	Patient recruitment method	PPP / Justification provided	Patients' response rate
Cope et al. (1986) USA [48]	Quasi-experimental with control group. Secondary care.	68 residents.	424 patients; (mean age 53, 67% females).	Consecutive sampling	6-7 No	73%
Greco et al. (1995) Australia [36]	Cross-sectional. Primary care.	33 GPTs.	295 patients; (average age 39, 65% females).	Consecutive sampling	10 No	<u>89%</u>
Jenkins and Thomas (1996) UK [37]	Cross-sectional. Primary care.	10 GPRs.	426 patients.	Consecutive sampling	50 No	85%
Hall et al. (1999) Canada [38]	Cross-sectional. Primary & secondary care.	308 physicians ^[102] .	6,825 patients.	Systematic sampling	25 No	89%
Greco et al. (2001) UK [39]	Cross-sectional. Secondary care.	39 (21 consultants, 10 registrars, & 8 nurses).	1,416 patients (mean age 57, 59% females).	Consecutive sampling	40-50 No	No data
Greco and Pocklington (2001) UK [43]	Cross-sectional. Primary care.	13 pairs of GPRs & GPTs.	973 patients (mean age 45.4, 66% females).	Consecutive sampling	50 No	No data
Greco et al. (2001) Australia [15]	Randomized, controlled, longitudinal study. Primary care.	210 GPRs	28,156 patients (mean age 37.5, 70% females)	Consecutive sampling	50 Yes	No data
Lipner et al. (2002) USA [44]	Cross-sectional. Secondary care.	356 physicians	8,900 patients (average age 59, 57% females)	Random selection	25 Yes	No data

Table 4 (Continued)

Author(s), (Year), Country, [Reference]	Study design Setting	Healthcare professionals	Patients	Patient recruitment method	PPP / Justification provided	Patients' response rate
Sargeant et al. (2003) Canada [40]	Cross-sectional. Primary care.	142 family physician	3,550 patients	Random selection	25 No	No data
Wood et al. (2004) USA [18]	Cross-sectional. Secondary care.	7 radiology residents	57 patients	No data	12-14 No	No data
Al-Shawi et al. (2005) UK [45]	Cross-sectional. Secondary care.	10 surgeons	402 patients	No data	35-40 Yes	No data
Mackillop et al. (2006) UK [41]	Cross-sectional. Primary care.	No data	No data	Consecutive sampling	30 Yes	No data
Reinders et al. (2008) Netherlands [42]	Cross-sectional. Primary care.	48 GPTs	878 patients	Consecutive sampling	30 Yes	No data
Violato et al. (2008) Canada [46]	<u>Uncontrolled before and after study.</u> Primary care.	250 family doctors or GPs	6,250 patients	No data	25 No	mean response rate per doctor: - 24.09 (time1) - 24.39 (time2)
Violato et al. (2009) Canada [47]	Cross-sectional. Primary care.	238 occupational therapists	2,881 patients	No data	15 No	81%
Vinod et al. (2013) Australia [20]	Cross-sectional. Tertiary care.	7 radiation oncologists	55 patients	Convenience sampling	10 No	79%

Abbreviations: PPP: Patient Per Practitioner; GPTs: General Practice Trainees; GPRs: General Practice Registrars; GP: General Practitioner.

Table 5

Mechanics of patient feedback process

Study (year) [reference]	Person(s) in charge of questionnaire administration	Questionnaire collection	Patient anonymity protected	Practitioner blindness	Feedback reporting method
Cope et al. (1986) [48]	Research assistant	Returned directly to receptionist or complete missing data by phone	Yes	No data	Individualized report
Greco et al. (1995) [36]	No data	No data	Yes	No data	Individualized report ^[30]
Jenkins and Thomas (1996)	Physician	No data	No data	No	Individualized reports
Hall et al. (1999)	Office staff	Returned to data processing centres in prepaid envelope	Yes	No	Individualized reports
Greco et al. (2001) [39]	Ward managers (setting 1) Audit department (setting 2)	Returned to a private organization	Yes	No data	Individual reports
Greco and Pocklington (2001) [43]	Reception staff	Collected by an independent research organization	Yes	No data	Individualized reports
Greco et al. (2001) [15]	Physician	No data	No data	No	Written summary of patient questionnaires ^[30]
Lipner et al. (2002) [44]	Patients used a touch-tone telephone to complete the questionnaire using a coded number, patients were identified by the physician diplomate	Diplomates monitor questionnaire completion rate through the phone system	Yes	No	Aggregated performance feedback report

Table 5 (Continued)

Study (year) [reference]	Person(s) in charge of questionnaire administration	Questionnaire collection	Patient anonymity protected	Practitioner blindness	Feedback reporting method
Sargeant et al. (2003) [40]	Physician	Collected and analysed by the Customer Information Services	Yes	No	Individualized reports
Wood et al. (2004) [18]	Patients were asked to volunteer in the study by a breast imaging technologist	No data	Yes	No data	Individualized reports
Al-Shawi et al. (2005) [45]	No data	Collected by the staff from the clinical audit department	Yes	Yes	Individualized reports
Mackillop et al. (2006) [41]	Independent person	Returned immediately to the designated person, or send back by freepost	Yes	No data	Results were given at an appraisal meeting
Reinders et al. (2008) [42]	General practitioner trainee	Patients handed over the questionnaire in an envelope to a teaching staff	No	No	Individualized reports
Violato et al. (2008) [46]	Office personnel	No data	Yes	No data	Individualized reports
Violato et al. (2009) [47]	Occupational therapist	No data	Yes	No	Individualized reports
Vinod et al. (2013) [20]	Questionnaires were mailed from the department to patients identified by radiation oncologists	Returned using a self- addressed stamped return envelope to an independent research unit	Yes	No	Individualized reports

Table 6

Follow-up and impact of patient feedback.

Study	Follow-up to practitioners	Impact of patient feedback
Cope et al. [48]	Repeat questionnaire after detailed feedback	A significant increase seen in the scores of the residents of the feedback group (changes to individual practice)
Greco et al. [36]	Focus group discussions, teleconferences & telephone interviews	Patient feedback had the potential to affect their behaviour towards patients
Jenkins and Thomas [37]	No data	No data
Hall et al. [38]	Focus group discussions and completing questionnaires	Changes in practice were planned or initiated by number of physicians, especially to communication with patients
Greco et al. [39]	Completing "Report on Interpersonal Skills" and taking part in group meetings	Patient feedback process helped <u>practitioners</u> in identifying their strengths and areas needing improvement
Greco and Pocklington [43]	Completing "Report on Interpersonal Skills"	Patient feedback process helped <u>practitioners</u> in identifying their strengths and areas needing improvement, physicians also attended a three-hour workshop to further develop their communication skills
Greco et al. [15]	Frequent reassessment and completing follow-up questionnaires	Patient feedback increased the registrars' confidence and helped in identifying areas needing improvement for future interactions with patients
Lipner et al. [44]	Completing a "Quality Improvement Plan"	Intentions to change communication strategies with patients and to continue seeking feedback from patients & peers
Sargeant et al. [40]	Program evaluation	Changes are planned especially those addressing communication with patients
Wood et al. [18]	An individual "personal quality improvement" interviews	Patient feedback increased awareness of <u>practitioners</u> of how to interact and communicate more effectively with patients
Al-Shawi et al. [45]	Focus group discussion	Patient comments had strong influences on making significant changes to the <u>practitioner's</u> consultation technique
Mackillop et al. [41]	No data	No data
Reinders et al. [42]	Group interviews and completion of an evaluation form	Patient feedback has a great potential for improving communication skills
Violato et al. [46]	Reassessment using the same questionnaire	Upward changes in performance

Violato et al. [47]	Evaluation questionnaire	Positive expressions by participants regarding MSF instruments and process
Vinod et al. [20]	Completing a survey assessing MSF	Changing aspects of practice were planned
