The educational potential of patient feedback on physicians’ consultation skills: a systematic review

Marcel E. Reinders
Bridget L. Ryan
Annette H. Blankenstein
Henriëtte E. van der Horst
Moira A. Stewart
Harm W.J. van Marwijk

Submitted
Abstract

Objective
The effectiveness of patient feedback assessments as a method of improving physicians’ consultation skills is equivocal. Research is scarce and shows a wide variety in methods and rigour. In this systematic review, we analysed the evidence for the educational potential of real-patient feedback on consultation skills of practicing physicians.

Methods
Five electronic databases were searched (PubMed, EMBASE, Cochrane, PsycInfo and Eric) and screened against eligibility criteria: 1) practicing physicians, 2) real patient feedback, 3) general consultation skills, 4) general health care, 5) physicians being actually confronted with individual feedback. Empirical studies of all study designs (randomized (controlled), quasi-experimental, cross-sectional and qualitative designs) were eligible for inclusion. Two authors selected the articles for inclusion independently against eligibility criteria. Articles were assessed for level of educational impact: views on learning experience (level 1); a change in knowledge or skills (level 2); an intended change in the physicians’ performance (level 3); and an actual change in performance (level 4), and quality according to a best evidence medical education (BEME) coding scheme.

Results
Fifteen studies were identified (10 studies in primary care and five in other specialties) conducted in the United States, the Netherlands, the United Kingdom, Australia and Canada. The method of tailored reports with aggregated patient feedback results for transferring feedback was most commonly described. Eighty percent of the studies showed an improved outcome (effect) as a result of patient feedback. However, positive results were more common among studies assessing, for example, perceptions of participants’ knowledge (considered a low level of educational impact) than studies assessing actual
change in consultation skills (considered the highest level of educational impact) where four out of seven studies found beneficial effects.

**Conclusions**

There is evidence for the educational potential of patient feedback on physicians’ consultation skills, however, the evidence for actual improvement is limited. The implication therefore is that the rationale for implementation of patient feedback into practice depends on the level of aimed impact.
Introduction

The importance of patient assessments and feedback as a stimulus for improving physicians’ consultation skills, or as an incentive for efforts to improve quality of medical care, is gaining increased attention\textsuperscript{1,2}. Considering patients’ views on the delivery of medical care seems to be self-evident and is also related to improved patient health outcomes and satisfaction\textsuperscript{3-5}. The therapeutic essence of a doctor-patient relationship should include the patient perspective obtained either from ratings or surveys after encounters\textsuperscript{1}. However, to what extent patient feedback is effective in improving physician performance remains equivocal.

Patient feedback on communication or interpersonal skills of physicians commonly involves assessments by means of surveys\textsuperscript{6,7}. This feedback is then transferred to physicians through a variety of methods such as tailored reports or coaching sessions, in which the collated results of patient surveys are presented, or through focus groups in which patient feedback assessments are discussed. Patients appear comfortable with the various methods of involvement\textsuperscript{8}. Formal feedback on the assessment of physicians from various sources (e.g. mentors, trainers, colleagues) is effective in improving clinical performance\textsuperscript{2}, but studies lack rigorous methodology\textsuperscript{6}. The evidence for the effectiveness of feedback assessments from patients is scarce and seems to be hampered by the heterogeneity of studies\textsuperscript{9,10} and the scarcity of rigorously validated instruments\textsuperscript{7,11}. In two recent systematic reviews about instruments and patient feedback methods\textsuperscript{12}, and about patient feedback assessments and brief training interventions in general practice\textsuperscript{13}, no conclusive evidence was found for the effectiveness of patient feedback interventions. Additionally, these reviews only included randomized trials, which are scarce in this field. What is still missing, therefore, is a review of studies that focuses on other aspects of educational potential of patient feedback interventions, in addition to those that focus on actual change in skills. Due to the complex nature of the effectiveness and educational potential of patient feedback, these studies are often conducted in various, non- randomized designs\textsuperscript{6,14}. The effectiveness of patient feedback on
consultation skills is found on different levels of educational impact. In the Kirkpatrick hierarchy\textsuperscript{15,16}, four levels on which educational interventions can have impact are identified: 1. views on learning experience; 2. change in knowledge or skills; 3. intended change in performance; and 4. actual change in performance.

In this systematic review, we addressed the following research question: What is the evidence for the educational potential (from physicians’ valuation of learning experience to physicians’ behavioural change) of individually directed, real-patient feedback on general consultation skills of practicing physicians? Therefore, we included all studies (randomized trials, or quasi-experimental studies and qualitative research) with different levels of aimed educational impact. We explored participant appreciation and physician behavioural change, assessed by patients and/or by physicians. Furthermore we performed a quality assessment of included studies and made suggestions for practical implications and for future research.

**Methods**

**Search strategy**

The search strategy, based on the search definition ‘physicians receiving feedback of real patients on their consultation skills’, used the following controlled vocabulary and free text words: ‘feedback’, ‘clinical competence’ and ‘health care’ (full search strategy available from MR). A medical information specialist applied the search (last update on April 23\textsuperscript{rd} 2009), to the following databases: PubMed, EMBASE, Cochrane, PsycInfo and Eric, without restrictions to language or publication year. Other sources for eligible articles were reference lists of related systematic reviews identified in this search, and the articles included for final analysis.
Eligibility criteria
We included all empirical studies regardless of study designs meeting the following eligibility criteria:
- practicing physicians, including physicians in postgraduate training (but not medical students)
- real patient feedback (but not standardized simulated patients or other parties, like health workers, nurses or peers only, giving feedback)
- general consultation skills, (but not satisfaction of received care or practice organization, or articles limited to one aspect of the consultation, e.g. giving bad news, shared decision-making, palliative care)
- feedback on communicative aspects in general health care (but not only related to a specific disease, or to patients or physicians concerned with a specific disease)
- physicians being actually confronted with individual feedback (but not conceptual studies of the value of patient feedback or validation studies of surveys, without feedback to physicians).

Selection process
One reviewer (MR) screened the titles and abstracts of all citations to determine those matching the search definition (first step, 'low-threshold' strategy). Two independent authors (MR, BLR) further selected articles on the basis of title and abstract fulfilling the eligibility criteria (second step, 'eligibility criteria'), which led to a selection of articles for full text review. Disagreement about articles fulfilling eligibility criteria was resolved by discussion. Full text articles that met the eligibility criteria were included for final analysis (third step, 'final selection').

Data extraction and analysis
For all included studies, two reviewers (MR, BLR) performed data extraction (two of the studies were written by the first author; these articles were reviewed by an independent assessor). Discrepancies were resolved by discussion. The template used in this study is based on the standard best evidence medical education (BEME) organisation coding sheet\textsuperscript{17}, and contains the following areas:
(1) aim of the study, (2) study design, (3) study population, (4) intervention type and intensity of patient feedback interventions, (5) level of educational impact according to the Kirkpatrick hierarchy: views on learning experience (level 1); a change in knowledge or skills (level 2); an intended change in performance (level 3); and an actual change in the physicians’ performance (level 4); (6) learning outcomes; (7) effectiveness of patient feedback; and (8) quality assessments. We chose not to pool results of the studies, because we expected beforehand that studies would be too heterogeneous methodologically. In Figure 1 a schematic overview is presented of the several steps from patient feedback assessments to outcome assessments of the learning effects.

**Quality assessment**

Quality assessments were conducted separately for quantitative research and qualitative research. Studies with mixed methods were analysed according to the core method that was used. Quality assessments of quantitative studies were based on four major items: prevention of bias (participant selection, patient selection, adherence to the intervention and validity of the instruments), participant sample size, intensity of the intervention (number of involved patients, tailored coaching sessions and duration of the intervention), and the appropriateness of the intervention and adequacy of description (based on a standard sheet to extract the quality of clinical trials)\(^\text{18}\). Quality assessments of qualitative studies were based on: well defined goals, appropriateness of the design, the appropriateness of data collection, and rigour of conduct, analysis and reporting\(^\text{19}\). Studies were not excluded for analysis because of low quality scores.

**Classification of study effects by study characteristics**

The outcome of the effects on consultation skills (classified as positive or negative) as a result of patient feedback was related to: intervention type, aimed level of impact, assessment level (patients or physicians), stage of education, setting (country, medical specialty), and publication year.
Figure 1. A model of patient assessments, feedback procedure and outcome assessments

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Method</th>
<th>Subject of Study</th>
<th>Assessment of Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real-patient assessments</td>
<td>Questionnaires</td>
<td>Change</td>
<td>Self-assessment</td>
</tr>
<tr>
<td>(surveys, interviews)</td>
<td>Tailored reports</td>
<td>Intentions of behavioural change</td>
<td>Patient assessment</td>
</tr>
<tr>
<td></td>
<td>Small group discussions</td>
<td>Modification of knowledge or skills</td>
<td>Teacher/researcher assessment</td>
</tr>
</tbody>
</table>
Results

Search

Figure 2 shows a flow-diagram of the results of the selection process. After removing 815 duplicates, a first broad screening of 8452 citations (8410 from electronic data-bases and 42 from hand-search) on titles and abstracts matching the search definition yielded 287 possible articles. After a second screening on title and abstracts against eligibility criteria, 108 full text articles were retrieved, of which 15 studies met the eligibility criteria and were included \(^{20-34}\). Two studies were based on the same data-set \(^{33,34}\) but explored different outcome measures and were therefore both included.

Reasons for exclusion were: 42 patient assessments studies but no actual feedback \(^{8,35-75}\); 26 validation studies of patient assessment questionnaires \(^{76-101}\); five non-empirical studies \(^{102-106}\); 12 reviews \(^{1,3,9,10,12,13,107-112}\); and eight studies involved no real patients, no reflection on individually addressed general consultation skills, or involved undergraduate students \(^{113-120}\).

Description of Studies

Table 1 describes each of the included studies.

Setting

The included studies were conducted in the USA \(^{21-24,32}\), the Netherlands \(^{30,31,33,34}\), United Kingdom \(^{20,27,29}\), Australia \(^{26,28}\), and Canada \(^{25}\). Ten studies concerned postgraduate training (general practice \(^{24,26-28,30,31}\), internal medicine \(^{22,23,32}\), and pediatrics \(^{21}\)); five studies concerned professional education (general practice \(^{25,29,33,34}\) and orthopaedics \(^{20}\)).

Instruments for acquiring patient feedback

All but one \(^{29}\) study made use of patient feedback questionnaires (mostly validated), focusing on individual physician performance. Only the Patient Satisfaction Questionnaire was used by more than one study group \(^{21,23,32}\). Two
studies made use of structured patient interviews\textsuperscript{24,29}, one study did this in combination with written questionnaires\textsuperscript{24}.

\textit{Patient feedback presentations}

The commonest method of transferring feedback from patients to physicians was by means of aggregated/collated results of questionnaires in tailored reports made by preceptors/researchers\textsuperscript{20,24-28,32-34}. Another method was tailored coaching, in which the feedback results were mediated by a researcher/preceptor, varying from one single coaching session to quarterly meetings\textsuperscript{21-24}. The method of consultation-specific feedback questionnaires directly presented to the physician was applied in two studies\textsuperscript{30,31}. In two studies, small group sessions discussing patient feedback were conducted\textsuperscript{27,29}, either organized monthly during one year, or once at the end of the study period.

\textit{Additional measures to enhance the educational potential of patient feedback}

Some studies used physician self-assessment reports of their consultation skills, in addition to patient feedback questionnaires. This challenged the participants to reflect on their performance and the learning possibilities. Physicians’ self-assessment reports were made after each single consultation\textsuperscript{30,31}, or as a concluding report\textsuperscript{20,21,26-28}. The possibility of discussing the results and meaning of patient feedback with a preceptor\textsuperscript{27} or GP trainer\textsuperscript{29-31} was meant to further intensify the learning effect.

\textbf{Outcomes of Studies}

\textit{Outcome measures and their assessments}

The outcome measures were defined as: consultation skills\textsuperscript{30,31}, communication skills/behaviour\textsuperscript{21,23-25,34}, interpersonal skills\textsuperscript{20,26-28}, humanistic qualities\textsuperscript{32}, and non-technological skills\textsuperscript{22}. Communication skills and interpersonal skills form an integrated competence\textsuperscript{1}. Outcome measures were assessed by patients\textsuperscript{23,33}; by physicians\textsuperscript{25,32}, by a third party\textsuperscript{24,29,30,34}, or in combination\textsuperscript{20-22,26-28,31} (Figure 1).
Study designs and the intended level of educational impact

Figure 3 shows a model in which the level of educational impact is integrated with the study design (qualitative or quantitative research). The four randomized controlled trials all aimed at level 4 impact (actual change). The quasi-experimental studies aimed at various levels of impact (from level 4 to level 1, of which one study (Fiddler) explicitly aimed at both level 3 (intended change) and level 4). Of the four cross-sectional studies and the two qualitative studies, the aimed impact was at level 2 or level 1.

Effect size (data synthesis)

Twelve of the 15 studies (80%) showed improved outcome (effect) as a result of patient feedback. Table 2 describes the outcomes and significance for each study. The effects by specified level of educational impact were:

- Level 4 (observed change):
  Brinkman et al described 35% increase of ratings of consultation skills as a result of patient feedback\(^1\) (outcome reported by nurses and not by patients’ (parents) assessments). Fiddler et al described that approximately 12% of physicians initiated changes on communication and support of patients\(^2\) (self-assessments). Cope et al described a ‘large effect’ on improved patient valuation of consultation skills\(^3\), but Greco et al found only a small effect at the upper end of a patient satisfaction scale\(^4\). In contrast, in the studies of Vingerhoets and Wensing, and of Reinders, no difference in improved consultation skills was reported\(^5,6\) (patients and preceptor assessments), however, in the latter study sub-group analysis showed that active adherence to the study programme was related to more improvement in consultation skills\(^7\).

- Level 3 (intended change):
  Fiddler et al\(^8\) reported that 19% of physicians contemplated a change in communication and patient support as a result of patient feedback (versus 12% initiated change, see also level 4).
• Level 2 (modification of knowledge or skills):
In four studies (Al-shawi, Greco (1995, 2001) and Radcliffe), a majority of participants self-reported a modification in knowledge or skills as a result of patient feedback\(^{20,27-29}\), especially if the learning instrument gave insight into their deficiencies\(^{27}\). Among issues that were identified as important for change were listening, explaining and investing time\(^{29}\), awareness and more concern with the patients\(^{22}\), or communication skills in general\(^{20,28}\).

• Level 1 (appreciation/valuation):
Falvo et al described that patient feedback was considered by physicians as a useful reflection on strengths and weaknesses\(^{24}\). In the study of Thomas\(^{32}\), only patient feedback on humanistic qualities was considered relevant. In the study of Reinders\(^{30}\), the usefulness of patient feedback was limited to the more complex consultations (content related). The high patient scores suggested to the physicians that patients were providing socially desirable answers. Physicians viewed this as unhelpful and this hampered physicians’ adherence to the programme. Some quantitative studies also presented valuation data but this was not the main focus of these studies. For example, Greco et al described that GP registrars perceived better understanding of interpersonal skills\(^{26}\), whereas in the study of Wensing\(^{34}\), physicians found patient feedback to be time-consuming and difficult to learn from (due to high-levels of patient satisfaction).

Quality of the included studies
Table 3 shows the quality assessments of the quantitative studies (Table 3 a) and the qualitative studies (Table 3 b), separately.

Classification of study effects by study characteristics
Positive changes were found throughout the professional stages (experienced clinicians as well as postgraduates). Positive change did not depend on the intervention method (structured reports, tailored coaching, or otherwise), or assessment level (patients, participants). All studies that aimed at lower level of educational impact (1 and 2) found positive results; at impact level 4, however,
only four out of seven studies found a positive result. Positive outcomes were demonstrated in studies conducted in English-speaking countries (throughout the last three decades, among all specialties), whereas interventions were shown ineffective, as described in the Netherlands (in the last decade, in general practice).

**Figure 2. Flow-diagram of study selection**

- **Identification**: 9225 records through database screening, 42 additional records identified through hand search.
- **Screening**: 8452 records after duplicates removed screened on title and abstract, 8344 records excluded.
- **Eligibility**: 108 full-text articles assessed for eligibility, 93 full-text articles excluded.
- **Included**: 15 studies included.
<table>
<thead>
<tr>
<th>Study</th>
<th>Study aim</th>
<th>Intervention</th>
<th>Study design</th>
<th>Setting and population</th>
<th>Patients per physician</th>
<th>Level/stage of education</th>
<th>Validated patient feedback instruments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brinkman (2007)</td>
<td>Assessment of effectiveness of multisource feedback on consultation skills and professionalism</td>
<td>Tailored coaching with collated results of parent, nurse and self-assessment evaluations</td>
<td>Randomized (controlled) trial</td>
<td>18 paediatric residents (18 controls), USA</td>
<td>Not specified</td>
<td>Post-graduate training (PT)</td>
<td>Patient Satisfaction Questionnaire (PSQ)</td>
</tr>
<tr>
<td>Brody (1980)</td>
<td>Assessment of effectiveness of a project to improve recognition of non-technological aspects of medical care</td>
<td>Tailored coaching with collated results of patient interviews</td>
<td>Quasi-experimental</td>
<td>43 residents internal medicine, USA</td>
<td>4-6</td>
<td>PT</td>
<td>Derived from questionnaire of Hulka and Ware</td>
</tr>
<tr>
<td>Cope (1986)</td>
<td>Assessment of preceptor feedback of patient ratings of technical and non-technical quality of care</td>
<td>Tailored coaching with collated results of patient surveys</td>
<td>Quasi-experimental, with control group</td>
<td>9 residents internal medicine (9 controls), USA</td>
<td>≥ 6</td>
<td>PT</td>
<td>PSQ</td>
</tr>
<tr>
<td>Falvo (1980)</td>
<td>Assessment of feasibility and educational value of patient feedback as a tool for developing patient teaching skills</td>
<td>Tailored coaching with collated results of patient surveys</td>
<td>Qualitative analysis (interviews)</td>
<td>7 GP residents, USA</td>
<td>50</td>
<td>PT</td>
<td>12-item questionnaire about patient-physician interactions</td>
</tr>
<tr>
<td>Fidler (1999)</td>
<td>Assessment of initiated changes as a result of patient (multisource) feedback; valuation of educational efforts</td>
<td>Report with collated results of aggregated multisource feedback</td>
<td>Quasi-experimental</td>
<td>218 GPs, Canada</td>
<td>25</td>
<td>Professional education (PE)</td>
<td>40-item questionnaire on communication, support and practice organisation</td>
</tr>
<tr>
<td>Greco (1995), Greco (2001), Al-Shawi (2005)</td>
<td>Feasibility of introducing patient feedback into the vocational training of GPs or into the professional training of surgeons</td>
<td>Individualised structured feedback report of collated results of patient surveys</td>
<td>Cross-sectional, both quantitative and qualitative</td>
<td>33 GP residents, Australia</td>
<td>10</td>
<td>PT</td>
<td>(precursor of) Doctors’ Interpersonal Skills Questionnaire (DISQ)</td>
</tr>
<tr>
<td>Greco (2001)</td>
<td>Assessment of impact of different models of patient feedback on interpersonal skills development</td>
<td>Colated summary of patient surveys; additional preceptor discussions</td>
<td>Quasi-experimental, 2 control groups</td>
<td>70 GP residents (140 controls), Australia</td>
<td>50</td>
<td>PT</td>
<td>DISQ</td>
</tr>
<tr>
<td>Study</td>
<td>Study aim</td>
<td>Intervention</td>
<td>Study design</td>
<td>Setting and population</td>
<td>Patients per physician</td>
<td>Level/stage of education</td>
<td>Validated patient feedback instruments</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------</td>
<td>---------------------------------</td>
<td>------------------------</td>
<td>--------------------------</td>
<td>-----------------------------------------------------</td>
</tr>
<tr>
<td>Radcliffe (1997)</td>
<td>Reflection on invoke of patients' accounts of care as audit tool</td>
<td>Small group discussions about the results of patient interviews</td>
<td>Qualitative analysis (group interviews)</td>
<td>±40 GPs, UK</td>
<td>10</td>
<td>PE</td>
<td>Semi-structured interviews</td>
</tr>
<tr>
<td>Reinders (2008)</td>
<td>Development of a patient feedback programme as a tool for improvement of consultation skills</td>
<td>Experiential learning with consultation-specific patient feedback questionnaires</td>
<td>Cross-sectional, both quantitative and qualitative</td>
<td>58 GP residents, The Netherlands</td>
<td>20</td>
<td>PT</td>
<td>Patient Feedback Questionnaire on Consultation Skills (PFC)</td>
</tr>
<tr>
<td>Reinders (2009)</td>
<td>Assessment of effectiveness of patient feedback on consultation skills of GP trainees</td>
<td>Experiential learning with consultation-specific patient feedback questionnaires</td>
<td>Cluster-randomized controlled trial</td>
<td>23 GP resident (30 controls), The Netherlands</td>
<td>18</td>
<td>PT</td>
<td>PFC</td>
</tr>
<tr>
<td>Thomas (1999)</td>
<td>Assessment of valuation of patient feedback</td>
<td>Collated summary of patient surveys</td>
<td>Quasi-experimental, with control group</td>
<td>77 residents internal medicine (including controls), USA</td>
<td>7</td>
<td>PT</td>
<td>PSQ</td>
</tr>
<tr>
<td>Vingerhoets (2001)</td>
<td>Assessment of effectiveness of patient feedback on improvement of care</td>
<td>Individualised structured feedback report of collated results of patient surveys</td>
<td>Cluster-randomized controlled trial</td>
<td>30 GPs (30 controls), The Netherlands</td>
<td>30</td>
<td>PE</td>
<td>Chronicle Ill Patients Evaluate General Practice (CEP)</td>
</tr>
<tr>
<td>Wensing (2003)</td>
<td>Assessment of effectiveness of patient feedback on communication behaviour</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Outcomes

<table>
<thead>
<tr>
<th>Study</th>
<th>Outcome measures</th>
<th>Assessments</th>
<th>Significance (documented improvement or valuation)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Impact level 4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brinkman (2007)</td>
<td>Physicians' communication skills</td>
<td>Parent feedback questionnaires, nurse evaluation questionnaires</td>
<td>Nurse evaluations, but not parent evaluations, showed 35% change in communication with family</td>
</tr>
<tr>
<td>Cope (1986)</td>
<td>Patient satisfaction (friendliness and respect; understanding and non-judgmental attitude; communication skills; technical quality of care)</td>
<td>Parent feedback questionnaires</td>
<td>A ‘large’ effect on improved patient satisfaction (art of care and technical quality)</td>
</tr>
<tr>
<td>Greco (2001)</td>
<td>Interpersonal skills: listening, reassurance, eliciting concerns and fears, time given Evaluation of residents concerning patient feedback (perceived usefulness; intentions to modify interpersonal skills)</td>
<td>Patient feedback questionnaires; Follow-up survey on residents</td>
<td>A small improvement was found for: listening skills, time given to patients, and eliciting concerns and reassurance Intensity of the patient feedback module was related to improvement in interpersonal skills 63.4% of the residents had intentions to change behaviour</td>
</tr>
<tr>
<td>Reinders (2009)</td>
<td>MAAS-global score of general consultation skills</td>
<td>Preceptor observation of video-taped consultations</td>
<td>No change in improved consultation skills, although sub-group analysis showed moderate improvement in actively participating GP residents</td>
</tr>
<tr>
<td>Vingerhoets (2001)</td>
<td>General dimensions of care (patient relation and communication; information and advice; support)</td>
<td>Patient evaluation questionnaires</td>
<td>No change in the dimensions of care GPs self-reported changes in their professional performance and organisation of care (no empirical data shown)</td>
</tr>
<tr>
<td>Wensing (2003)</td>
<td>MAAS-global score of general consultation skills List of motivations and barriers to learn from patient views and change behaviour</td>
<td>Researcher observation of video-taped consultations; Post-intervention survey on GPs</td>
<td>No change in consultation skills GPs are open minded to learn from patients views, but find it difficult to meet all patients’ needs</td>
</tr>
<tr>
<td><strong>Impact level 3 (and 4)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fidler (1999)</td>
<td>Contemplated or initiated changes in relation to patient evaluations (communication and support of patients; office staff and systems) Evaluation of preferred learning method</td>
<td>Follow-up survey on participating GPs</td>
<td>Feedback by patients, rather than by peers, was most likely to induce initiated changes; especially among physicians with the lowest patient ratings 12.3% of physicians initiated change on communication and support of patients (19.3% contemplated a change)</td>
</tr>
</tbody>
</table>
## Table 2. Outcomes (continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Outcome measures</th>
<th>Assessments</th>
<th>Significance (documented improvement or valuation)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Impact level 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Al-Shawi (2005), Greco (1995), Greco (2001)</td>
<td>Physicians’ report on strengths and weaknesses of interpersonal skills and intended actions</td>
<td>Patient feedback questionnaires; Physician self-reports (after receiving feedback); Group discussions with participants</td>
<td>60% of the orthopaedic surgeons, 90% of the GP residents, and 67% of the GPs (highly) valued patient feedback. Some of the surgeons, 75% of the residents and 67% of the GPs tried to modify their interpersonal skills.</td>
</tr>
<tr>
<td>Brody (1980)</td>
<td>Patient satisfaction and attitude towards illness</td>
<td>Patient interviews (including Patient Satisfaction and General Health questionnaires); Assessing physicians’ reports of consultations and chart reviews; Assessing physicians’ reports of consultations and chart reviews</td>
<td>76% of the physicians reported change in communication style and content; 89% received new and unknown information. No change in patient satisfaction or in recognition of non-technological problems. 79% of the physicians reported possibilities for improved patient care, 85% valued the educational experience.</td>
</tr>
<tr>
<td><strong>Impact level 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Falvo (1980)</td>
<td>Communication and interpersonal skills (needs were met; listening skills and understanding)</td>
<td>Authors’ personal interviews with participating residents</td>
<td>Residents found the project to be helpful in identifying their strength and weaknesses in communicating skills.</td>
</tr>
<tr>
<td>Radcliffe (1997)</td>
<td>Patients evaluations and GPs reactions on: communication and explanation, doctor-patient relationship</td>
<td>Transcripts of group meetings with GPs</td>
<td>Individual examples of valuations of patient feedback by the participants.</td>
</tr>
<tr>
<td>Reinders (2008)</td>
<td>Patients’ perception of patient-centered consultation skills</td>
<td>Patient feedback questionnaires; Group discussions and evaluation forms of GP trainees</td>
<td>45.6% of the residents found patient feedback to be beneficial for improving consultation skills.</td>
</tr>
<tr>
<td>Thomas (1999)</td>
<td>Evaluation of patient feedback by residents, concerning humanistic qualities and clinical skills</td>
<td>Follow-up survey on residents</td>
<td>A sustained valuation of patient feedback on humanistic qualities; but a decline in valuation of patient feedback on other clinical skills.</td>
</tr>
</tbody>
</table>
Figure 3. An integrated model of research type (quantitative, qualitative) and level of educational impact (Kirkpatrick’s hierarchy level 1-4)
### Table 3 a. Quality of quantitative studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Prevention of bias</th>
<th>Sample size participants</th>
<th>Intensity of intervention</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>Al-Shawi (2005)</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>Brinkman (2007)</td>
<td>1.0</td>
<td>0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Brody (1980)</td>
<td>0</td>
<td>1.0</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Cope (1986)</td>
<td>1.0</td>
<td>0.5</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Fidler (1999)</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
<td>0.5</td>
</tr>
<tr>
<td>Greco (1995)</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0</td>
</tr>
<tr>
<td>Greco (2001)</td>
<td>0.5</td>
<td>1.0</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>Reinders (2009)</td>
<td>1.0</td>
<td>0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Thomas (1999)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>Vingerhoets</td>
<td>1.0</td>
<td>1.0</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>(2001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wensing (2003)</td>
<td>1.0</td>
<td>0</td>
<td>0.5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

A: randomized group of participants; B: randomized group of patients; C: controlled for adherence of participants to the intervention; D: use of validated instruments; E: sample size participants > 25; F: number of involved patients per physician > 10; G: tailored coaching of patient feedback intervention; H: duration of the intervention > 6 months; I: appropriateness of intervention; J: well described and equivocal presentation of results and conclusions; 1.0 = yes; 0 = no; 0.5 = description was unclear or missing or only partly met our criteria.

### Table 3 b. Quality of qualitative studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Well defined goals</th>
<th>Appropriateness of design (focus groups)</th>
<th>Appropriateness of data collection</th>
<th>Rigour of conduct</th>
<th>Rigour of analysis</th>
<th>Rigour of reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Falvo (1980)</td>
<td>1.0</td>
<td>1.0</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>Greco (2001)</td>
<td>1.0</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>Radcliffe (1997)</td>
<td>0.5</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Reinders (2008)</td>
<td>1.0</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

1.0 = yes; 0 = no; 0.5 = description was unclear or missing or only partly met our criteria.
We have identified 15 studies in which the educational potential of real patient feedback on general consultation skills has been investigated. These studies cover three decades of research and vary in study design, levels of educational impact, and quality. A majority of the 15 studies observed a positive effect of real patient feedback on consultation skills of physicians; however, there are some caveats. First, the positive results are clustered among studies using an outcome measure at the low end of Kirkpatrick’s hierarchy of educational impact, and were non-randomized and qualitative studies. Where the impact level was high in the hierarchy (actual change in consultation skills), four of seven studies found a positive effect. Second, only one of four randomized trials found positive results.

A possible reason for this phenomenon is that the assessment of actual change in general consultation skills can be difficult, due to the lack of precision in defining consultation skills and the lack of responsiveness of the outcome instruments. Greco et al report that changes at the high end scale (i.e. near the ceiling) of a measuring instrument and a limited effect size are related\textsuperscript{26}. Baseline scores at the high end of the scales, a concern in most of the studies, appear to make changes in behaviour less urgent, when in reality these high scores may reflect a difficulty in measurement rather than true high satisfaction with physician communication. An important consideration in measurement is that learning (and talking) is a subjective process. Authenticity and person-centeredness can be objectified and counted, but perhaps in the process of transferring them from the subjective to the objective, some of meaning, validity and relevance are lost.

In this review, some correlations between study characteristics were studied. The assumption that the intensity of patient feedback interventions was related to increased changes in outcomes, could not be confirmed or rejected, mainly because of the plethora of the applied methods and the lack of description of the adherence of participants to the programme. Furthermore, it is not clear why in the English-speaking countries (USA, Canada, United
Kingdom, Australia) more positive results were found, in contrast to the studies in the Netherlands, although of course, a relatively small number of studies is involved. Moreover, there were no indications that younger doctors (residents) were more liable to change than more experienced clinicians, who might have acquired routine or fixed patterns of consultation. This suggests that patient feedback may be as useful in faculty development and continuing medical education as it is in post-graduate education.

In contrast to the vast number of patient satisfaction studies, effect studies of patient assessments are less common; and randomized controlled trials on the educational impact of patient feedback are a rarity. Finding evidence for the educational value of the assessments and feedback assessments of practicing physicians was hampered by the heterogeneity of studies. This heterogeneity made effect sizes difficult to standardize and compare. The possibilities of an integrated presentation of qualitative research have been demonstrated by several others. In this review, we focussed on the aimed level of educational impact of the included studies, as a way of a combined and integrated presentation of the qualitative and quantitative studies.

Future patient feedback effect studies should focus on the highest level of educational impact, that of actually achieved changes in consultation skills. Qualitative studies with high rigour are also needed, in order to achieve: better understanding of how physicians can use patient feedback to improve communication; valid assessment of participants (patients and physicians, and teachers); preferred methods in which patient feedback should be acquired and presented; and high levels of adherence to training programmes.

**Conclusion**

Beyond the principle that patient involvement in health care is self-evident, and morally justifiable, we need a better understanding of how physicians can use patient feedback to learn to communicate better with their patients. This review has demonstrated that there is evidence for the effectiveness of real patient feedback interventions on lower levels of educational impact such as perceived
change in knowledge, but few studies and little evidence for an impact on the higher level such as actual change physicians’ consultation skills.
References


5. Stewart MA. Effective physician-patient communication and health outcomes: a review. CMAJ 1995 May 1;152(9):1423-33.


13. Cheraghi-Sohi S, Bower P. Can the feedback of patient assessments, brief training, or their combination, improve the interpersonal skills of primary care physicians? A systematic review. BMC Health Serv Res 2008;8:179.


36. Branch WT, Malik TK. Using 'windows of opportunities' in brief interviews to understand patients' concerns. JAMA 1993 April 7;269(13):1667-8.


