

GPs perspective on the usability of the Dutch Primary Care Practice Report: a qualitative interpretative approach

Piet De Bekker (✉ p.j.g.m.de.bekker@vu.nl)

Vrije Universiteit Amsterdam <https://orcid.org/0000-0001-8560-9103>

Maarten Vink

Vrije Universiteit Amsterdam

Marianne Donker

Vrije Universiteit Amsterdam

Paul Sterkenburg

Vektis

Eric van der Hijden

Vrije Universiteit Amsterdam

Research

Keywords: Healthcare providers' performance, general practitioners, audit and feedback, benchmark information, change clinical behavior, claims-based data

Posted Date: March 13th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-17051/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background Audit and feedback informs healthcare providers and may affect professional practice and patient outcomes. The Primary Care Practice Report (PCPR) is a web based personalized feedback instrument for general practitioners (GPs) in the Netherlands, based on claims data. Its yearly uptake is limited. In order to improve the use and usability of the report this study aims to identify key criteria that GPs deem important for audit and feedback.

Methods A qualitative interpretative approach was used. We interviewed 12 GPs about their use of the practice report. These interviews followed the Three-Step Test-Interview method. Thematic content analysis was used to investigate the perception of the GP on the report and their perspective on the usability. The interviews resulted in critical items for audit and feedback, on which all tables and graphs in the report were systematically assessed.

Results From the interviews with GPs recurring criteria emerged that were identified as decisive for the effectiveness of performance feedback: content, reliability, validity and usability. The 34 tables and graphs of the PCPR were assessed using factual characteristics. Content analysis shows that the PCPR has a strong focus on costs. Assessment on validity indicates that casemix correction is always performed if relevant, but explanatory notes hardly clarify which (sub)population is measured (3/34), therefore GP's in general have difficulty interpreting the results. Assessment on usability shows that, although benchmark figures are almost always presented based on national references, the formulation of goals or any specific attention for an action perspective is lacking because the perceived comparability with the patient group of the GP is limited.

Conclusions The current PCPR does not meet key criteria for effective audit and feedback, as defined by GPs. It has a strong focus on costs instead of clinical behavior and is poorly understood when it comes to the specific population the data reflect. The results are in line with theoretical perspectives on learning and improvement of professionals. Improvement of the PCPR requires information on aspects of clinical behavior that are recognizable for GPs and they actually can influence. First steps are made to improve the method of case-mix correction in the benchmarking.

Contributions To The Literature

Literature shows that performance feedback is an effective tool for changing provider behavior and thus quality of care, if designed correctly. The results of our study build upon several earlier contributions on the effectiveness of audit and feedback and adds the specific GP perspective on elements that they deem important.

From interviews with GPs, recurring criteria emerged that were identified as decisive for the effectiveness of performance feedback: content, reliability, validity and usability.

Presented feedback data should describe actual clinical behavior, use and explain consistent case-mix correction, and deliver both guidance on the target and a strategy to improve.

Background

Performance feedback is defined as a summary of clinical performance over a specific period of time, and the provision of that summary to individual practitioners, teams, or healthcare organizations. It is widely used as a strategy to change clinical behavior (1). Literature shows that performance feedback is an effective tool for changing provider behavior and thus quality of care, if designed correctly (2).

Performance feedback might be a strategy on its own or can be used as a component of multifaceted quality improvement interventions. Healthcare professionals are prompted to modify their practice when given performance feedback showing that their clinical practice is inconsistent with a desirable target (3).

A theoretical underpinning of this behavioral change, which supports the use of performance feedback in the context of healthcare provision, is found in the Control Theory and the Feedback Intervention Theory (4). Control Theory follows the hypothesis that people change their practice when there is a gap between current behavior and a goal which they wish to attain. Theory suggests that behavior change is most likely if feedback is accompanied by comparison with a behavioral target and by action plans (5). This is supported by empirical evidence, following the Cochrane systematic review of Ivers et.al. in 2012 (3). Three Cochrane reviews over the course of 10 years came to the same conclusion: audit and feedback (A&F) generally leads to small but potentially important improvements in professional practice. Yet despite the increasing number of audit and feedback trials, uncertainty remains regarding when audit and feedback is likely to be most helpful and how to best optimize the intervention (6).

Kluger & DeNisi (4, 7) developed the Feedback Intervention Theory (FIT), as a further step towards a learning cycle. They state that the motivation to hear feedback is not equal to the motivation to improve performance. A central item in FIT is how feedback focuses one's attention. The learner should feel 'in charge', and therefore feedback should activate a locus of control (8). The central assumption of FIT is that feedback interventions (FIs) change the locus of attention among three general and hierarchically organized levels of control: task learning, task motivation, and meta-tasks (including self-related) processes. The results suggest that FI effectiveness decreases as attention moves up the hierarchy closer to the self and away from the task (4).

Building upon these insights, how can clinical audit and feedback be used as an instrument? The mechanisms for effective performance feedback are based on three main assumptions (3, 5, 8, 9). Firstly, the recipients have the wish and are able and willing to modify their behavior. They should agree on the goals of change with those delivering the performance FI. Furthermore, feedback provided should be relevant and meaningful and delivered in a way that is readily accessible to the learner. Finally, the learner is assumed to know how to interpret and act on the feedback once it is received (8). It appears that acceptance and effectiveness of performance feedback are correlated with the objective of the feedback system: does it aim to deliver feedback only, or is it for example used as a basis for contract negotiations (3).

Audit and feedback uses different techniques to invite for behavioral change, such as specifying a behavioral goal, self-monitoring, providing incentives, increasing skills through rehearsal, social support or encouragement and so forth (9). Generally, audit and feedback is characterized by a feedback loop, which represents an iterative, self-regulating process (5).

Performance feedback is suitable in primary care settings, where a high incidence of chronic conditions is managed, and a limited set of tasks is performed multiple times, providing the feedback recipient multiple opportunities to address and change the behavior in question (10). Knowledge on the perspective of GPs, regarding the characteristics of audit and feedback that effectively influence behavior, is limited, however.

In the Netherlands, initially one health insurance company started to give a selected group of GPs feedback on their claims data in order to inform them on their practice. Following up on this initiative, in 2015 all health insurers agreed to provide feedback to all GP practices based on claims data of all insured citizens. The Primary Care Practice Report (PCPR), a nationwide benchmark tool for general practitioners providing feedback on their clinical practice, is since then freely available. The PCPR is developed and distributed by Vektis, the Dutch center for information and standardization in healthcare, a research institute linked to health insurers.

From 2015 onwards, Vektis invited GP practices to download the PCPR. The PCPR was updated in December 2016, July 2017, January 2018, July 2018, November 2018 and March 2019. The benchmark information consists of about 30 pages with 34 figures and tables. The PCPR presents general information on the primary care practice, descriptive information of the patient population, and expected versus observed information on (a) total healthcare costs, both in primary care and in other facilities, (b) numbers of consultations and primary care interventions, (c) pharmaceutical costs, (d) mental healthcare use, (e) medical specialist care use, and (f) referral or use of diagnostics. Actual care provided in a GPs practice is compared with expected values for a fictional practice with a comparable – for casemix adjusted/corrected – population. Since November 2018, following requests of users, a customized version is available in which results of regionally bundled GP practices are presented as a comparison.

Although downloads are available without charge, the uptake of the PCPR among GPs is limited. According to the records of Vektis, approximately 60% of all practices (2,908 of about 5,000 GP practices) in the country have accessed the PCPR at least once in the period of 2015–2019. The number of GP practices that have ever downloaded a PCPR is slowly increasing [Figure 1]. Of this group, one third downloaded the PCPR three times or more [Figure 2]. Repeated use of the PCPR is not customary (11).

[Insert Fig. 1]

[Insert Fig. 2]

Healthcare delivery to patients may be improved using audit and feedback. The limited uptake of the PCPR is of concern. Therefore, this study aims to analyze GPs perspective on the usability of the PCPR.

Methods

This study uses a qualitative interpretative approach to understand the perspectives of GPs. In semi-structured interviews the GPs were asked about their attitude and perspective towards the PCPR. Thereafter observational data were collected on the way participants understood the PCPR following the Three-Step Test-Interview (TSTI) method (12). This method includes both think-aloud and retrospective probing techniques. The TSTI was used to investigate the experiences of GPs with specific tables and graphs in the PCPR. A reflective assessment of the PCPR was performed, based on key criteria that were derived from these interviews.

Population and procedure.

From the 9,039 GPs in the Netherlands, working in about 5,000 GP practices, 80 were selected, both randomly and related to the network of the authors or to the medical faculty. GPs were approached by email or telephone to request an interview on their use of the PCPR. Of the 80 GPs that were approached 18 were interested in participating. Finally, twelve conveniently sampled GPs were interviewed.

Data collection

General practitioners were interviewed in their GP-practice by a graduate (BvdK), who was supervised by two of the authors (PdB, MD), to gather information on perceptions, interpretation and actual use of the PCPR. The interviews were conducted between May 8th, 2018 and July 4th, 2018. An interview protocol was used as well as written informed consents, and member checks (respondent validations). Field notes were made during and after the interview. No repeat interviews were carried out.

In semi-structured interviews GPs were asked to describe their perspective and attitude towards A&F and give their opinion on what useful feedback should look like. The topic list of the interview [Appendix 1] was based on intuition and on the Control Theory and the Feedback Intervention theory. Topics included were the information GPs use to gain insight in clinical practice; the feedback used to gain insight in quality of care; their opinions on A&F and the PCPR and their expectations of the PCPR.

After the interview, GPs were asked to 'think aloud' when viewing four randomly selected tables and graphs presented in the practice report. GPs were requested to reflect on anything they saw or thought, with three basic open questions in mind: (a) what is your understanding of what you see and read?; (b) what are your thoughts on this?; (c) how does this influence you and/or your practice? The Think Aloud method uncovered understanding of the data and identified perspectives on the usefulness of the data.

Data analysis

The interviews (30–45 minutes) were audio recorded and transcribed ad verbatim. Thematic content analysis was used to analyze the interviews. In several rounds, interviews were analyzed and coded (BvdK, MD, PdB) in order to generate recurring issues and an overview of findings and particularities leading to a coding tree in which patterns of recurrent perceptions, interpretations and actual use was

listed. The qualitative computer analysis software package, MAXQDA (version 18.0.8 for Mac, VERBI Software, Berlin, Germany) was used to manage, organize, code and analyze the interview transcripts. Data saturation was discussed, but was not a goal. Participants did not provide feedback on the findings.

After content analysis, an assessment tool for the PCPR was generated to systematically study all 34 tables and graphs, following the coding tree from the interviews [Appendix 2] to explore the PCPR more in depth. We identified three key issues from the interviews, and for each category, researchers (PdB, MV and EvdH) developed a subset of criteria on which the tables and graphs of the PCPR could be assessed. The complete assessment tool is shown in Appendix 3.

Two researchers (PdB and MV) independently assessed the 34 tables and charts presented in the PCPR. Purely descriptive information on the GP practice in the PCPR was left out of consideration. Disagreement on an assessment of a specific table or chart was resolved by consensus. If no consensus could be achieved, a third researcher (EvdH) decided how to interpret a table or graph. The complete assessment of all 34 tables and figures is shown in Appendix 4.

Results

The respondents

The majority of the respondents was female (67%, compared to 47% nationwide) and the mean age (53 years) was slightly above the national average (50 years). Half of the interviewed GPs (n = 6) work in a Group Practice, the others work in duo (n = 4) or in a community health center (n = 2). One GP practice also included a local pharmacy. Eight GPs had heard of the PCPR before the invitation. A minority of the GPs had downloaded and used the PCPR before the interview.

According to a few GPs, patient feedback is the most important source of feedback received. Direct feedback provided by patients, including compliments, critiques or complaints, is always taken into account and most impactful. Less direct personal feedback from patients may include the outflow of patients over the years, and feedback provided in patient surveys. However, one GP mentioned that the extent of impersonal patient feedback is insufficient to make valid statements on the quality of delivered care.

R1: "Were patients satisfied and did they appreciate the service, that is actually the most important source of feedback indicating whether I did a good job or not."

The majority of the GPs was familiar with benchmarking as a general feedback tool. However, regional peer-to-peer sessions with affiliated general practices, in which patient cases are discussed, were mostly used and considered as most essential and helpful. Reviewing deviations and abnormalities between colleagues leads to a deeper insight into and a critical attitude towards their own clinical behavior.

GPs were interested and curious about the benchmark feedback provided by the PCPR. Two GPs expressed a certain lack of trust about the source of the PCPR. The role of health insurance companies as a supplier of data was addressed, leading to a skeptical attitude among these respondents. The main expectation was that the PCPR would provide GPs with insight into their clinical practice and to examine if they perform better than others. Nevertheless, several respondents declared beforehand, that the PCPR would not change their clinical behavior.

R2: “Are a hundred colleagues providing too many treatments or am I providing too few? Hmmm, who knows?”

Content

An important observation of GPs was that the current information in the PCPR hardly presents data on clinical behavior or actual performance. Most information represents total or average costs, usually of healthcare providers following up on referrals. However, GPs feel that they are not able to influence average costs of mental healthcare or medical specialist care. Two GPs explicitly express the wish to receive more specific feedback on their own clinical behavior regarding common disorders and the influence they have on the health of their population.

R1: “I would like to see more figures on my referrals for separate conditions. Current information is mostly cost-related, so price increases have an effect. However, I can’t control those costs”

R7: “It is unclear to me how these figures might help the patient sitting at the other side of my desk.”

GPs miss information on health gains as a possible parameter on how well they perform. With the availability of data on actual health outcomes as a consequence of GPs activities, it is easier to compare one situation with another. E.g. does health of the GP population improve when the GP makes more home visits as compared to longer consultations in the office? Are health outcomes better when GPs perform small surgery themselves, or when people are referred to medical specialist care? And what are the consequences of strict gatekeeping versus more active referral patterns?

R11: “We strive for high quality healthcare, but feedback usually reflects euro’s or percentages, instead of health gains. For how many people suffering from high blood pressure did my prescription result in a lower blood pressure? And how many of my patients have not had a stroke as a consequence? How do I relate to peers?”

Reliability and validity

The majority of the figures appeared evident to the respondents and seemed well understood and correctly interpreted. Especially diagrams containing population characteristics and care utilization of patients, subdivided per age group, led to recognition and were similar to the respondents’ perception of their clinical practice.

R3: “We have fewer elderly patients than average, so that could partly explain the difference.”

Although respondents understand the majority of the presented feedback provided by the PCPR and could provide explanations for all data presented in the figures, not all explanations were based upon a correct interpretation. For example, eight of the twelve GPs indicated that differences and deviations from the expected rates were explained by casemix differences, however figures were already corrected for population characteristics in their specific practice. According to the GPs, this could for example explain differences in applying diagnostics or minor surgery. Some of the respondents disagreed with the data and therefore did not accept the presented feedback.

R1: “Then my thoughts are: is it correct? Is it correct for our population?”

R5: “This seems to be untrue; I do not know how the benchmark is developed, this difference is too big. So, I have my reservations.”

The organization that is calculating and distributing the PCPR is mistrusted because it is associated with the health insurance companies and therefore is not seen as an independent organization. Besides, the trustworthiness of the casemix adjusted ‘expected population’ in the benchmark is questioned.

Sometimes the data is difficult to interpret for GPs, yet the GP’s do not ask for help to interpret the data or to guide them in improving clinical behavior. Several respondents indicated that they would be more alert about their registration process because the PCPR showed that they registered their claims incorrectly. Three of the twelve GPs indicate that they may have claimed too little of specific interventions they performed, which would explain a value lower than expected.

R4: “This number strikes me. I think this is due to under registration on our part”

R6: “We should look at our administrative process. Probably we sometimes forget to submit declarations. That is no medical difference, but purely administrative.”

Timeliness is an issue for half of the respondents. In the PCPR that was presented in January 2018, information had reference date July 2016.

R4: “The longer ago, the more difficult it is to apply in practice. I would appreciate feedback on my figures from last year somewhere within the first three to four months of the current year. That would really help me.”

Usability

Respondents were careful in stating implications for their clinical behavior based upon the feedback. They would rather describe scenarios than determine actions to change clinical behavior. A few GPs

made assumptions that in the future they might use a different approach for specific patients, or focus on particular patient groups. Some respondents indicated that they might change their clinical behavior by increasing or decreasing the provision of minor surgeries or the amount of primary diagnostic applications. But as guidelines or standards of care are not available, or taken into account in the PCPR, it is impossible to decide what should be the preferred action.

R5: “If you are higher than expected at one year and lower at the year after, then I think: ‘the average seems good, there’s no reason to work differently’. If I have a consultation with a patient who is coughing for half a year, I do not think ‘oh, I’ve applied too much photo so far, let’s not refer this patient.’ You still look at the patient and his needs.”

R12: “Suppose I request a huge number of hip scans although there is no scientific evidence, but with all colleagues acting the same way. The presentation of a casemix corrected expected value based on averages would not lead to an adjustment of my clinical behavior.”

The PCPR generally presents data on the practice level, in which more than one GP may be active. More individualized feedback gains support from a majority of respondents. Although GPs recognize that they are partly responsible for the fact that a breakdown to the individual level is not possible, as claims data are filed per practice.

R6: “The figures per GP in this practice are indistinct. All patients are registered on [NAME OF PRINCIPAL GP]. So, if I request diagnostics, it will appear under his name”.

R9: “In this practice, two days a week another GP is employed. He requests tests on my practice code”.

Assessment tool based on interviews to systematically analyze the PCPR

From the interviews with GPs, recurring criteria emerged that can be categorized into three themes that were identified as decisive for the perceived effectiveness of performance feedback aimed at GPs. (1) Content: does performance feedback refer to actual healthcare delivery activities. (2) Reliability and validity: is information recognizable and reliable for the addressed person. (3) Usability: is performance feedback accompanied by an articulation of (behavioral) performance goals and/or an action plan?

Following these three key themes from the interviews, we constructed an assessment tool to systematically analyze all 34 tables and graphs in the PCPR. The key criteria are operationalized using numerous variables and transformed into an effectiveness assessment tool to systematically explore the PCPR more in depth. The complete assessment tool is shown in Appendix 3, the assessment itself in a separate document, Appendix 4.

Assessment of the Praktijkspiegel

Content

Figure 3 demonstrates that of the 34 tables and graphs, 21% (7/34) reflect on actual clinical behavior, while 62% (21/34) present data on average or total costs of patients in the GP practice. Other tables and graphs for example, compare demographic characteristics of the patients in the practice to the regional average. Tables and graphs that reflect clinical behavior provide feedback on the number of applications for diagnostics (3/34, 9%), number of consultations (2/34, 6%), number of interventions (2/34, 6%) and number of referrals (2/34, 6%). In 32% (11/34) of the tables and graphs, the PCPR provides feedback on costs that are partially or completely related to activities of the GP. Costs that are more or less the result of treatment choices of medical specialists after referral by the GPs are presented in 41% (14/34) of the tables and graphs. GPs are not able to influence the treatment choices of medical specialists and the associated costs. Some tables or graphs of costs associated with mental healthcare present total costs of the care delivered in a primary care setting (mental health practice assistant) and in specialized mental health facilities combined.

[insert Fig. 3]

Reliability and validity

The PCPR summarizes claims data of the population from a GPs practice per calendar year. In 4 tables and graphs, data regarding the total number of registered people is presented. In 30 tables and graphs, a subset of the registered people is presented, while “high cost patients” are excluded. High cost patients have mental healthcare costs of €10.000 or more, or total healthcare costs of €22.500 or more.

In three descriptive tables and graphs it is stated that (all) registered patients are included, in the other 31 tables and graphs it is not described clearly to what specific (sub)population the table or figure refers and if outliers are excluded.

In most (28/34) tables and graphs casemix correction is performed, but this is not directly clear from titles or explanatory notes. In an appendix to the PCPR, casemix correction method is explained. Actual values or percentages are compared to the expected values for a practice with a population.

The data presented is usually one year old (15/34, 44%) or two years old (19/34, 56%).

Usability

One table and graph presents the feedback of individual GPs within the group practice. All other (33) tables and graphs use the GP practice as aggregation level of analysis. Also, all except one used benchmarks. Figure 4 shows the benchmarks that were used. In most tables and graphs (28/34), the casemix corrected reference population is presented as the benchmark, often (18/34) also compared with values of own clinical practice in other years.

[Insert Figure]

In some cases, as Fig. 5 shows, data is presented for one or more subgroups of patients of the GP practice. These subgroups allow to further analyze healthcare delivery or costs for specific age groups. In

20/34 tables and graphs, no subgroups are presented. In 14/34 tables and graphs a breakdown in age-groups, gender, healthcare use, or socio-economic status – based on aggregated statistical data of the average regional income – is displayed, with one table and graph combining age and gender.

[Insert Fig. 5]

Discussion

We set out to study three things. First, to explore the limited use of the PCPR by understanding how Dutch GPs look at audit and feedback in general and the PCPR in particular. Secondly, we wanted to explore how GPs interpret tables and graphs and to what extent those interpretations lead to modification of their clinical behavior, at least in their perception. Thirdly, we systematically assessed all tables and graphs based on the underlying shortcomings of the PCPR from the GPs perspective.

The current information in the PCPR hardly presents data on clinical behavior or actual performance. Most tables and graphs represent total or average costs, usually of healthcare providers following up on referrals. The high number of tables and graphs regarding average or total costs indicate a focus on the (financial) perspective of healthcare procurement/purchasing, which is actually different from the (clinical) perspective of the healthcare provider.

In most tables and graphs, “high cost patients” – mental healthcare costs of €10.000 or more, or total healthcare costs of €22.500 or more – are excluded. The main reason to exclude this subset, is that the small number of patients with high costs may cause a disproportionate distortion of the average values.

Findings in the interviews contrast the assessment as regards to the casemix correction. GPs often explained higher or lower values than expected by a lack of correction for casemix, although 82% of graphs and tables present expected values for a casemix corrected population. This limited understanding and sometimes misinterpretation may be due to unclear explanations if casemix correction is used and how this was done.

The time-interval between data of feedback by the PCPR and data presented can be explained by a delay in the settlement of declarations for some types of care. However, with two-year-old data, the usability for a GP practice is limited. This is reflected in literature (3, 13) and confirmed by GPs in several interviews.

The benchmarks that are used in tables or graphs are expected values of the casemix corrected reference population or actual values of the own clinical practice in other years. This makes both static comparison (how do I compare in one year to the expected value) and dynamic comparison (how is my development over the years) possible.

A strong element of the current PCPR is that all health insurers participate and claims data of all GPs is used to construct the feedback information. This means that the whole population of each GP is included in the calculations. Another positive characteristic is the fact that the PCPR is provided on a recurring basis and is freely available as download for all GPs. Weak elements are the lack of explanatory notes to

explicitly describe the content of tables and graphs and the way population characteristics are taken into account. The technical explanations are only given in an annex.

Following both interviews with GPs and an assessment of the PCPR, our results show that the PCPR does not meet the criteria necessary to be effective in changing clinical behavior. This may explain the limited recurring use of the PCPR by GP practices.

Strengths and limitations

Through the Three-Step Test-Interview method we were able to observe the actual interaction between the instrument and respondents. The Think Aloud technique was used to make the cognitive processing ('thinking') observable. We could identify items that explain the limited uptake and influence the usage of the PCPR. The dependability was another strength: the researchers performed iterative data collection and analysis to inform and adjust further data collection.

To our best knowledge no assessment tool for individual feedback interventions towards GPs is available. In order to score the content, reliability, validity and action perspective of the PCPR, it was necessary to generate an assessment tool based on our interviews and available literature.

The researchers kept a diary to reflect on the process and the researcher's role and influence. Documenting the steps and decisions taken in the research, and their underlying motives, led to a clear audit trail.

Furthermore, the assessment strongly confirmed the perceptions and experiences of GPs, explaining why the current PCPR leads to limited understanding and sometimes misinterpretations regarding the correction for the expected population benchmark.

Some limitations of this article should be addressed. We interviewed a limited number of conveniently sampled GPs who were older than the average GP and we interviewed more female than male GPs. This might influence the transferability of results. There is no description of diverse cases or discussion of minor themes. Although the PCPR is constructed by an independent research institution, its close connection with the health insurers made GPs question the data.

Yet, the method of how the interviews were performed can be considered a strength of our study. The use of several sources, both interviews and user information, led to triangulation. Interviews were taken by an independent researcher from a university setting (BvdK), and coding was performed by several researchers.

Interpretation

The results of our interviews and assessment are consistent with literature. Studies indicate that the effectiveness of performance feedback on healthcare quality improvement depends on how the feedback system is constructed, what information is chosen to be presented and moreover, how it is perceived and

implemented by healthcare professionals (2, 3, 5, 6). These same studies also reveal several characteristics and conditions for performance feedback, to support a change of clinical behavior. The 2012 Cochrane review of audit and feedback showed that feedback format (verbal and written), source (a respected colleague or supervisor), frequency, improvement strategies (goal setting and action planning) and baseline performance explained some variation in the effectiveness of audit and feedback (3). Colqhoun (2017) identified seventeen design elements within six categories for performance feedback in order to be effective (14). Developers of feedback reports are surprisingly unaware that an evidence base of best practices exists to guide them. This is the case even for some well-resourced, large-scale interventions (15).

Content (clinical relevance), reliability and validity (are results recognizable for me with regard to my patient population) and usability (action perspective) are mentioned in one way or another in the different reviews of design elements (1, 3, 6, 8, 14). Making a regional version of the PCPR adds to the recognition of data for GPs.

In order to generate clinically relevant performance feedback, information that is used by healthcare professionals in general, or more specifically by the GPs in our study, several recommendations can be made.

First, healthcare providers are motivated more when they receive performance feedback on their clinical behavior, instead of feedback with a focus on healthcare costs of their patients. Especially when performance feedback towards a GP refers to costs of healthcare delivery further down the care path, the results can only partially and indirectly be influenced by the GP himself. Only the number of first referrals can be influenced by the GP who receives performance feedback. Apart from that, healthcare costs are influenced by e.g. designation level of providers, practice variation between providers leading to possible over- or undertreatment, length of treatment. GPs cannot be held accountable, even if price differences between hospitals are standardized as they are in the PCPR.

Second, when a PCPR has more focus on clinical treatment behavior of a GP, we advise that the development of performance feedback information is based on an explicitly agreed view of the role of GPs in healthcare. Regarding GPs, their role in the Dutch healthcare system has recently been reconsidered in conjunction with national branch organizations. GPs role as gatekeeper is reconfirmed, and performance feedback on the number and type of referrals should be used more intensely and include more detail. On top of that Dutch GPs experience the number of tasks increasing to include taking care of the vulnerable elderly and coordinating care for patients with specific chronic conditions such as diabetes, COPD or heart failure. Inclusion of patients in integrated care programs, figures on vaccination, or participation in screening programs could enrich performance feedback to GPs and would give a more complete overview of their clinical performance.

Thirdly, the performance feedback should provide benchmark figures that are informative for GPs and action focused. For reasons of comparability - patients with high costs may cause a disproportionate distortion of the average values – high cost patients are mostly excluded from comparison. Yet,

especially these high cost patients, who frequently visit the GPs office, might be an appropriate focal point for more efficiency and more appropriate healthcare delivery. It is recommended that performance feedback to GPs also includes separate information on these high cost patients.

It helps the understanding and correct interpretation when a clear explanation is available of the performed casemix correction. Clear description of results improves understanding and increases the acceptance of findings. Besides that, a benchmark needs to give an indication of the suggested improvement: what is good clinical practice? Without any comparison to a norm or standard of care, feedback will just feed a debate about numbers without a clear goal.

The current PCPR delivers audit and feedback on numbers and costs associated with consultations, referrals, pharmaceutical costs and use of diagnostics. Next level audit and feedback will support modern bundled payment models, in which GPs work in integrated networks of healthcare providers, where incentives such as shared savings and shared risks are aimed at rewarding for the delivery of high value healthcare. New PCPRs are crucial for the success of these models, of which first experiences are gained currently in the Netherlands.

Conclusions

Performance feedback is an important instrument to work towards high quality and efficient healthcare provision. The PCPR is a claims-based instrument, providing free and yearly feedback to GPs in the Netherlands. It leads to awareness and reflection on the population in a GPs practice and the care provided, but it does not fulfill some basic requirements for effectiveness. In order to guide GPs to more appropriate healthcare, it is necessary that the presented feedback data describes actual clinical behavior. Financial feedback information is considered to be more useful for health insurers procurement procedures but has limited value for feedback to healthcare providers.

When interpreting performance feedback, context matters. Casemix correction may add to the reliability and usability of feedback, but one should be aware that patients have different healthcare needs, depending on local characteristics. Usability increases when feedback appeals to the motivation of healthcare providers and delivers both guidance on the target and a strategy to improve.

Abbreviations

A&F Audit and feedback

FIs Feedback Interventions

FIT Feedback Intervention Theory

GPs General Practitioners

PCPR Primary Care Practice Report (Praktijkspiegel)

Declarations

Declarations

Not applicable

Ethics approval and consent to participate

Not applicable

Consent for publication

Consent forms of all interviewed GPs are available.

Availability of data and material

Data sets used and/or analyzed during the current study, such as transcriptions and categorizations are available from the corresponding author on reasonable request. Data generated or analyzed during this study as a result of the assessments tool are included in this published article [and its supplementary information files].

Competing interests

The authors declare that they have no competing interests

Funding

Not applicable

Authors' contributions

PdB analyzed the transcriptions, developed the draft assessment tool, performed the assessment of the Praktijkspiegel and was a major contributor in writing the manuscript. MV delivered recommendations for the assessment tool, performed the assessment of the Praktijkspiegel and was a contributor in writing the manuscript. MD did preparatory work for the interviews and delivered feedback on the qualitative research methods that were used. EvdH contributed to the logic of the findings. PS was responsible for data on the downloads and delivered feedback on the interpretation. All authors read and approved the final manuscript.

Acknowledgements

As part of his master's thesis, Bas van de Kolk (BK) performed and transcribed the interviews with GPs. PdB and MD were Bas' mentors during his research. We are grateful for the time and effort Bas put into this project.

Authors' information (optional)

Mr P.J.G.M. de Bekker holds a Master's Degree (Msc) in General Economics (Tilburg University) and has gained experience as a Health Economist since 1998. He is founder and advisor in his own consultancy, focused on analysis of the healthcare system. In 2015 he started his part-time PhD-research work at VU University (Amsterdam).

References

1. Brehaut JC, Colquhoun HL, Eva KW, Carroll K, Sales A, Michie S, et al. Practice Feedback Interventions: 15 Suggestions for Optimizing Effectiveness. *Annals of Internal Medicine*. 2016;164(6).
2. Hysong SJ. Meta-Analysis: Audit and Feedback Features Impact Effectiveness on Care Quality. *Medical Care*. 2009;47(3):356-63.
3. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database of Systematic Reviews*. 2012.
4. Kluger AN, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*. 1996;119(2):254-84.
5. Gardner B, Whittington C, McAteer J, Eccles MP, Michie S. Using theory to synthesise evidence from behaviour change interventions: The example of audit and feedback. *Social Science & Medicine*.

2010;70(10):1618-25.

6. Ivers NM, Grimshaw JM, Jamtvedt G, Flottorp S, O'Brien MA, French SD, et al. Growing literature, stagnant science? Systematic review, meta-regression and cumulative analysis of audit and feedback interventions in health care. *J Gen Intern Med.* 2014;29(11):1534-41.
7. Kluger AND, Angelo. *Feedback Interventions: Toward the Understanding of a Double-Edged Sword.* American Psychological Society. 1998;7(3):67-71.
8. Larson EL, Patel SJ, Evans D, Saiman L. Feedback as a strategy to change behaviour: the devil is in the details. *J Eval Clin Pract.* 2013;19(2):230-4.
9. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. *Health Psychology.* 2008;27(3):379-87.
10. Hysong SJ, Smitham K, SoRelle R, Amspoker A, Hughes AM, Haidet P. Mental models of audit and feedback in primary care settings. *Implementation Science.* 2018;13(1).
11. Vektis. Personal communication, dd. October 10th 2019. 2019.
12. Hak TV, K. van der; Jansen, H. The three steps test interview (TSTI). ERIM REPORT SERIES RESEARCH IN MANAGEMENT. 2004;ERS-2004-029-ORG.
13. Payne VL, Hysong SJ. Model depicting aspects of audit and feedback that impact physicians' acceptance of clinical performance feedback. *BMC Health Services Research.* 2016;16(1).
14. Colquhoun H, Michie S, Sales A, Ivers N, Grimshaw JM, Carroll K, et al. Reporting and design elements of audit and feedback interventions: a secondary review: Table 1. *BMJ Quality & Safety.* 2017;26(1):54-60.
15. McNamara PS, D.; De La Mare, D.; Ivers, N. Confidential physician feedback reports: designing for optimal impact on performance. *AHRQ Pub.* 2016(No. 16-0017-EF).

Figures

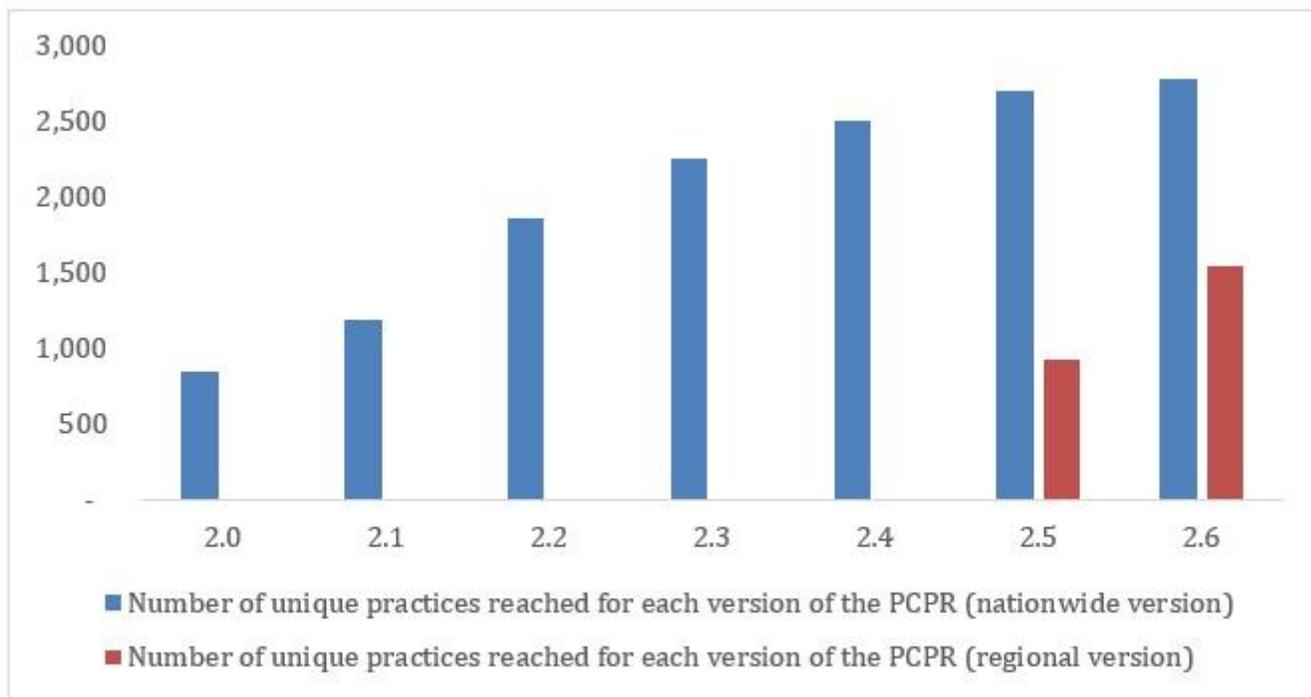


Figure 1

Number of practices that have downloaded a PCPR at least once increases

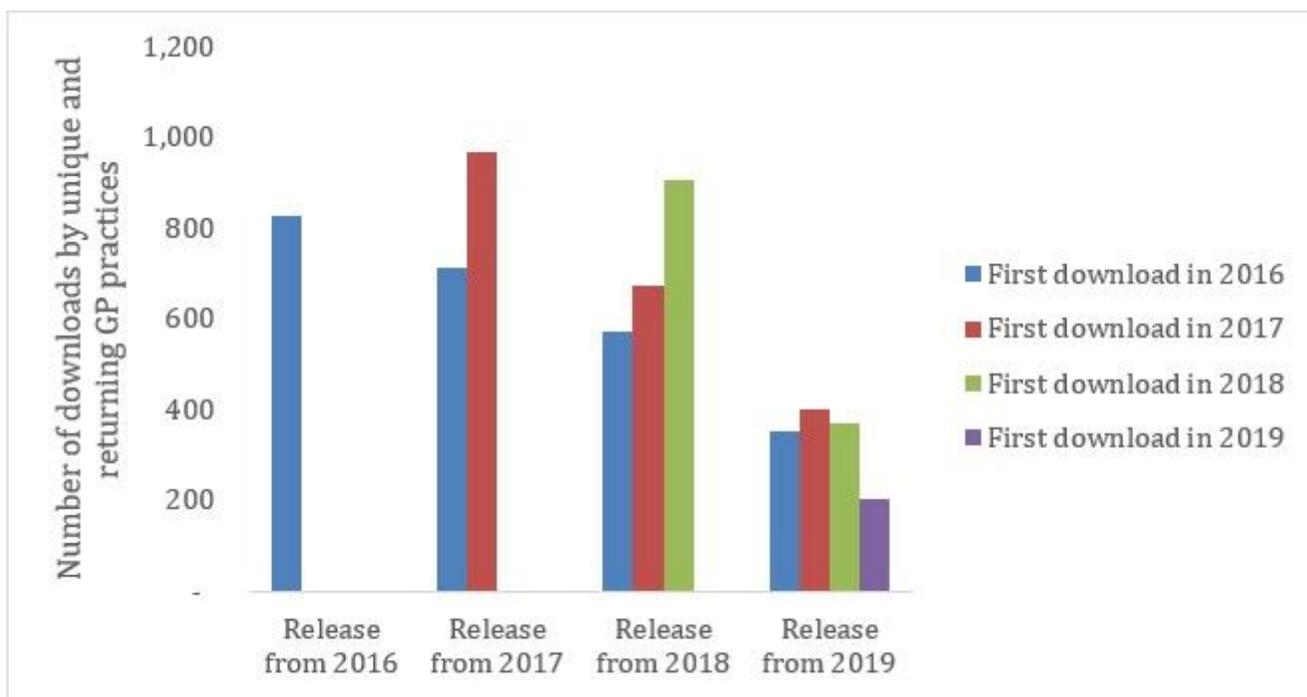


Figure 2

Number of practices that have multiple/repeated downloads is limited

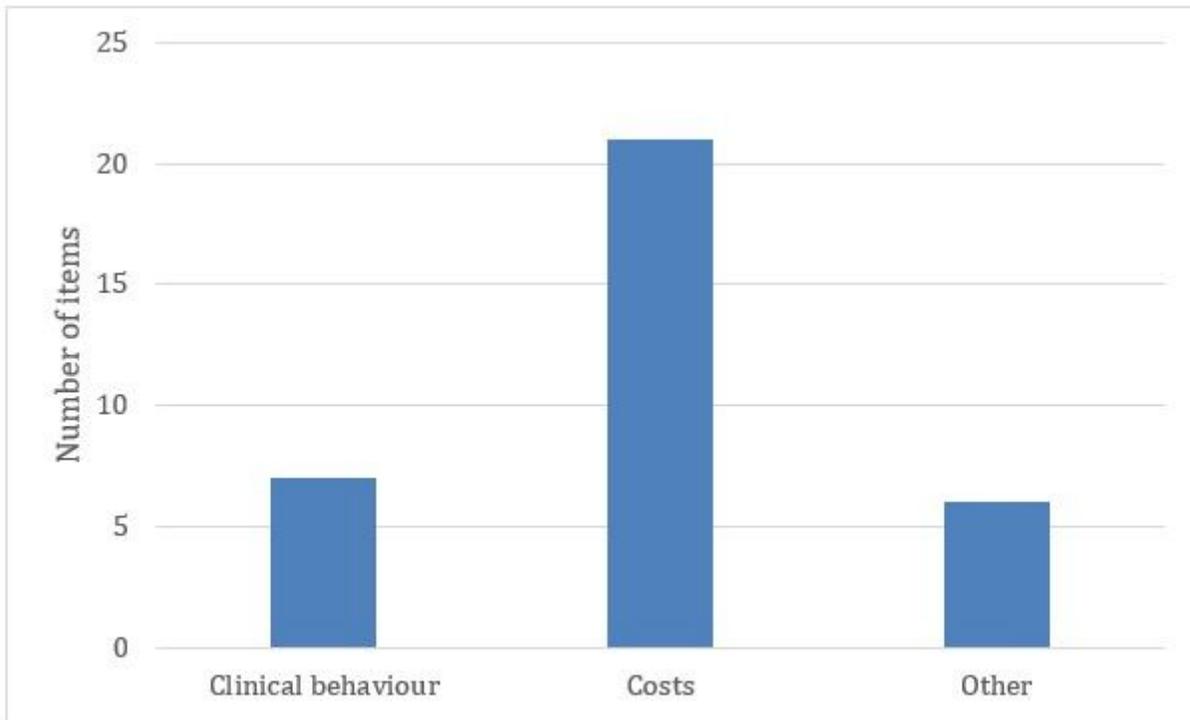


Figure 3

Subject of the feedback provided

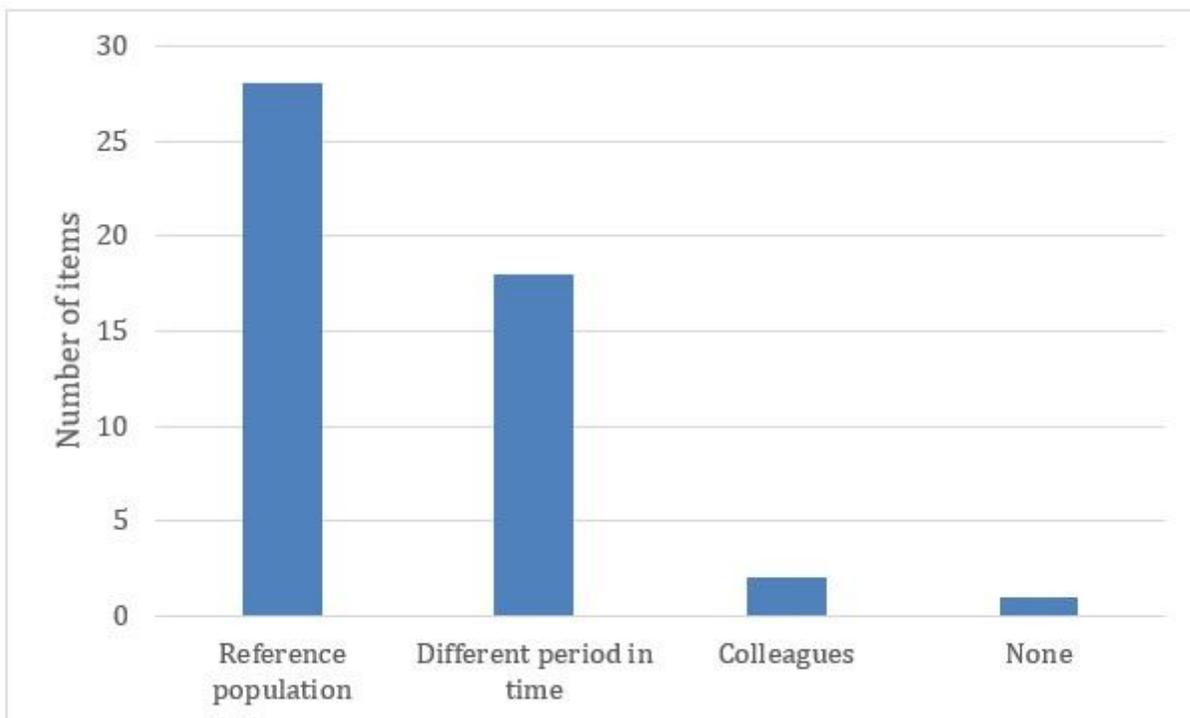


Figure 4

Source of comparison

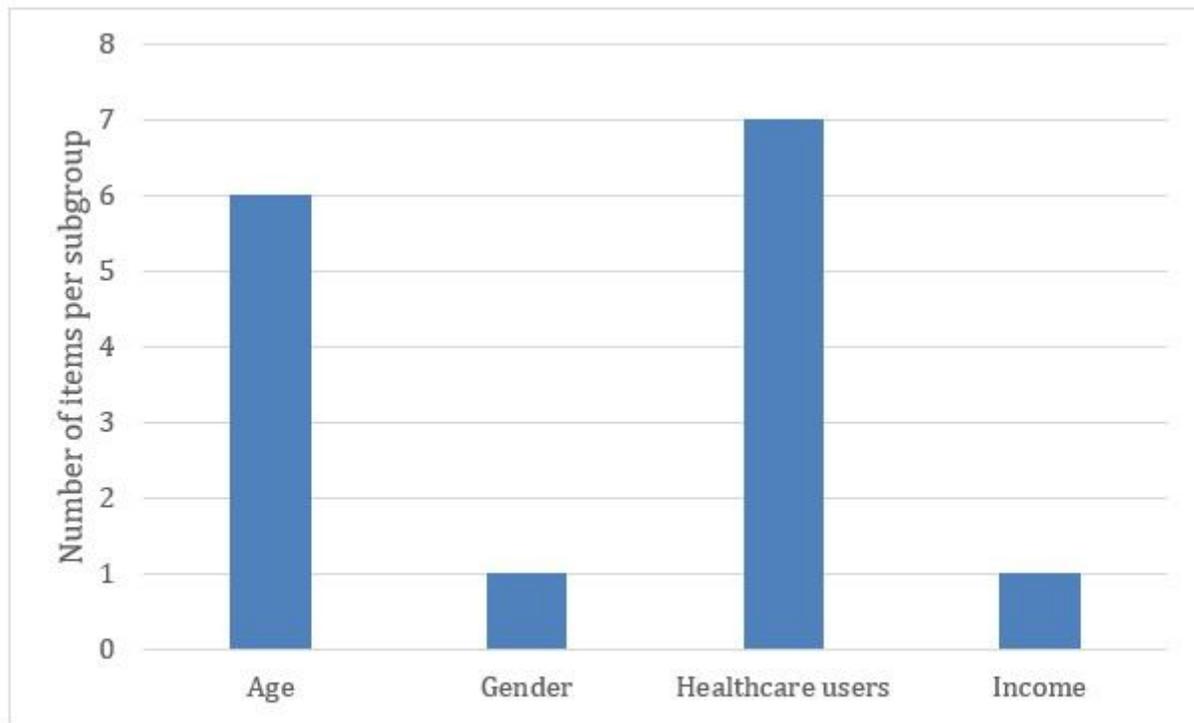


Figure 5

Breakdown of data presented for relevant subgroups (more than one subgroup possible)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [200303ISSMCOREQChecklist.pdf](#)
- [Appendix4scoreformtranslateddef.xlsx](#)
- [Append.docx](#)