Audit and Feedback and Clinical Practice Guideline Adherence: Making Feedback Actionable

Sylvia J. Hysong, Ph.D.[1]

Richard G. Best, Ph.D.[2]

Jacqueline A. Pugh, M.D.[2]

[1]Houston Center for Quality of Care and Utilization Studies

[2]Veterans Evidence-Based Research Dissemination and Implementation Center

The research reported here was supported by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service (HSR&D) (CPI #99-129). Dr. Hysong is a health services researcher at the Houston Center for Quality of Care and Utilization Studies, a VHA HSR&D Center of Excellence, and an Instructor of Medicine at Baylor College of Medicine in Houston. Drs. Best and Pugh are investigators in the Veterans Evidence-Based Research Dissemination and Implementation Center (VERDICT), a VHA HSR&D Research Enhancement Award Program; Dr. Pugh is also a Professor of Internal Medicine at the University of Texas Health Science Center at San Antonio, and a staff physician at the South Texas Veterans Health Care System. This research was conducted during Dr. Hysong's tenure at VERDICT. All three authors' salaries are supported in part by the Department of Veterans Affairs. The views expressed in this article are solely those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs, Baylor College of Medicine, or the University of Texas.

Abstract

BACKGROUND: Audit and feedback has been found to be variably effective at improving clinical practice guideline (CPG) adherence; yet audit and feedback research has not investigated the impact of feedback characteristics on its effectiveness.

OBJECTIVE: The purpose of this qualitative research is to explore how HPF and LPF differ in the way they use clinical audit data for feedback purposes.

RESEARCH DESIGN: Descriptive, qualitative, cross-sectional study of a purposeful sample of 6 Veterans Affairs Medical Centers (VAMCs) with high and low performance on 6 CPGs, as measured by external chart review audits.

METHOD: 102 employees involved with outpatient CPG implementation across the 6 facilities participated in one-hour semi-structured interviews where they discussed strategies, barriers, and facilitators to practice guideline implementation. Interviews were analyzed via grounded theory.

RESULTS: high performers provided feedback in a timely manner, specific to individual providers, and in a non-punitive fashion, whereas low performers were more variable in their timeliness, and relied on facility-level reports provided by an external source. The concept of actionable feedback emerged as the grounded theory in these data, with timeliness, individualization, non-punitiveness, and customizability as hierarchically ordered properties of actionable feedback.

CONCLUSION: facilities with a successful record of guideline adherence tend to deliver more timely, individualized, and non-punitive feedback to providers about their adherence than facilities with a poor record of guideline adherence; consistent with organizational research, feedback intervention characteristics may influence the feedback's effectiveness at changing desired behaviors.

Audit and feedback has been used for decades as a mechanism for changing the clinical practice behaviors of health care personnel. In clinical practice guideline implementation, audit and feedback has been used to attempt to increase guideline adherence across a wide variety of settings and conditions, such as inpatient management of COPD(1), test ordering in primary care(2;3), and ACE inhibitor and beta blocker usage in cardiac patients(4). Recent reviews, however, indicate that the effectiveness of audit and feedback as a strategy for behavior change is quite variable. Grimshaw and colleagues (5) reported a median effect size of audit & feedback of +7% compared to no intervention using dichotomous process measures, with effect sizes ranging from 1.3% to 16%; however, that same review reported non-significant effects of audit and feedback when continuous process measures were used. Along similar lines, Jamtvedt and colleagues (6) reported a median adjusted relative risk of non-compliance of .84 (interquartile range (IQR): .76-1.0), suggesting a performance increase of 16% (IQR: no increase to 24% increase). Such studies attribute much of the variability in effect of the interventions to a lack of attention to the characteristics of the feedback used in the intervention and/or to the conditions under which audit and feedback is more likely to be effective (6-9).

Earlier audit and feedback research has suggested that the timing of feedback delivery can impact the resulting behavior change (10), as can the credibility of the feedback source (11-13). Research from the organizational literature suggests a host of other potential explanatory phenomena as potentially impacting the effectiveness of feedback, such as its format (e.g., verbal vs. written), its sign (i.e., whether it is positive or negative) (14), and its content (e.g., whether it is task-focused or person-focused, individual or group based, normative or ipsative) (15). Our own research has noted that facilities that were more successful at implementing clinical practice

guidelines (i.e., high performing facilities, or HPF) relied more heavily on chart data as a source of feedback and placed great value on educational, non-punitive means of providing feedback to providers about guideline adherence compared to facilities with lower guideline adherence (low performing facilities, or LPF) (16). Taken together, these research findings indicate a need to further explore the characteristics of audit & feedback and their impact on the desired behavioral change. Building on our previous work on barriers and facilitators of clinical practice guideline implementation, the purpose of this qualitative research is to address this need in the audit and feedback literature by exploring how HPF and LPF differ in the way they use clinical audit data for feedback purposes.

Methods

Site Selection

The data presented in this study were part of a larger data collection effort at 15 VA facilities designed to examine barriers and facilitators to CPG implementation. The original 15 facilities were selected using a stratified purposive sample from four geographically diverse regional networks, primarily on the basis of their adherence to CPGs, which could be categorized into three groups: facilities with a sustained record of high adherence to CPGs (high performing facilities, or HPF); facilities with a sustained record of low adherence to CPGs (low performing facilities, or LPF), and facilities whose adherence record had significantly improved over a two year period (improvers). Group membership was determined via External Peer Review Program (EPRP) rankings, a random chart abstraction process conducted by an external contractor to audit performance at all VA facilities (see guideline adherence section below). Additionally, in order to be invited to participate, facilities had to be sufficiently large to accommodate at least two primary care teams, each containing at least three MD providers. In order to address the present

paper's specific research question, that is, to compare differences in how high versus low performing facilities use clinical audit data for feedback purposes, only sites from the high performing and low performing categories were included in the sample. As Daft and Lewin pointed out regarding studies of organizations: "The average organization does not exist and by definition is never on the frontier of the phenomenon under study. …The use of outlier research—studying the best and worst cases—is helpful when making prescriptive recommendations for practice. " (17), P. 6. In order to have a greater chance to discern the difference in patterns between high and low performance, the facilities in the third group, the improvers, were thus excluded from the study. Thus, the final sample for this paper consisted of three HPF and three LPF.

Measurement of Clinical Practice Guideline Adherence

Guideline adherence was measured via External Peer Review Program (EPRP) rankings; EPRP is a random chart abstraction process conducted by an external contractor to audit performance at all VA facilities on numerous quality of care indicators, including those related to compliance with clinical practice guidelines. VHA's Office of Quality and Performance (OQP) is the current curator of the guideline adherence data collected from chart abstractions via the EPRP. We obtained data from OQP for fiscal year 2001 reflecting facility-specific compliance with the guideline recommendations for each of six conditions: diabetes, depression, tobacco use cessation, ischemic heart disease, cardiopulmonary disease, and hypertension. Each condition is monitored via multiple performance indicators (e.g., tobacco use cessation is monitored via three performance measures: percent of patients screened annually for tobacco; percent of patients using tobacco who have been counseled three times in twelve months to cease tobacco use, and percent of patients currently not using tobacco); in total, 20 performance indicators were used to

describe compliance across the six conditions. Facilities were rank ordered from 1-15 (15 being the highest performer) on each performance indicator; all twenty performance indicator ranks were then summed together to arrive at an indicator rank sum (IRSUM) score (a higher IRSUM score indicating higher performance). Facilities were then rank ordered according to this IRSUM score to identify the three highest and the three lowest performing facilities.

Participants

102 employees across six facilities were interviewed. Within each facility, personnel at three different organizational levels were invited to participate: Facility leadership (e.g. facility director, chief of staff), middle management and support management (e.g., quality assurance manager, primary care chief, information technology manager), and primary care personnel (e.g., physicians, nurses, and physicians' assistants). All three levels were adequately represented in the sample (see Table 1). No significant differences in the distributions of participants were found by facility or hierarchical level ($\chi^2_{10} = 17.4$, n.s.). Local contacts at each facility assisted in identifying clinical and managerial personnel with the requisite knowledge, experience, and involvement in guideline implementation to serve as potential participants. The study was locally approved by each facility's institutional review board; all participation at each facility was voluntary. An average of nine interviews occurred at each facility, for a total of 54 interviews at the 6 facilities (Table 1).

Procedure

Three pairs of interviewers were deployed into the participating sites during the spring of 2001. The interviewers were all research investigators of various backgrounds, with in-depth knowledge of the project; most were involved with the project since its inception. Each pair traveled to a given site for two days, where together they conducted one-hour, semi-structured

interviews either individually or in small groups, depending on the participants' schedule and availability (see appendix for interview guide and protocol). Interviewers took turns leading the interview, while the secondary interviewer concentrated on active listening, note-taking, and asking clarifying or follow-up questions. To minimize interviewer bias, interviewer pairs were split and paired with different partners for their following site visit. All interviewers were trained *a priori* on interviewing and field note protocol.

Participants were asked how CPGs were currently implemented at their facility, including strategies, barriers and facilitators. Interviewers probed for additional information whenever participants raised the subject of EPRP reports or any other type of performance feedback initiatives as a tool to aid in clinical practice guideline adherence. Although we used prepared questions to guide the interview process, participants were invited to (and often did) offer additional relevant information not explicitly solicited by the interview questions. The interviews were audio recorded with the participants' consent for transcription and analysis.

Data Analysis

Interview transcripts were analyzed using a grounded theory approach (18), a well established qualitative method (19) whereby a theory or model is generated from existing data, by constantly comparing and relating coded concepts and categories in the data (known as the constant comparative method). Grounded theory consists of three analytic coding phases: open, axial, and selective coding; each is discussed below. Transcripts were analyzed using Atlas.ti 4.2, a commonly used qualitative data analysis software program (20).

Open coding. Open coding consists of the initial identification of passages containing the concepts of importance. Automated searches were conducted on the interview transcripts for instances of the following terms: "feedback", "fed back", "feeding back", "report" and its

variations (e.g., reporting, reports, reported), "perform" and its variations (e.g., performing, performed, performance), "audit" and its variations (e.g., auditing, audited, audits), and "EPRP". All word variations were captured via a truncated word search. The results were then manually reviewed for relevance; only passages that specifically discussed feedback on individuals' adherence to clinical practice guidelines were included -- all other references to feedback were excluded. Examples of excluded feedback references included feedback about the computer interface to information technology personnel, or anecdotal comments received from patients about provider adherence. This review resulted in 122 coded passages across the 54 interviews in the 6 facilities, for an average of 20 coded passages per facility.

Axial coding. In this phase of analysis, the passages identified during open coding are compared and thematically organized and related; in the present research this process resulted in four characteristics of feedback emergent from the data: timeliness, individualization, customizability, and punitiveness. Each is discussed in more detail in the findings section. References to feedback identified during open coding were categorized among these four properties, and were organized by facility according to each of these properties. Patterns in the high performing facilities were compared among each other searching for potential commonalities, as were patterns in the low-performing facilities. Once patterns were identified we relied on the corpus of field notes and informal observations from interviewers to validate the observed patterns and to provide interpretive context. For example, one respondent in a low performing facility mentioned "getting dinged on the EPRP", referring to the practice at that facility of using chart abstraction data for performance purposes, and being penalized when performance fell below a certain level. This instance was interpreted as an example of feedback that was delivered punitively, and was classified under the "punitiveness" category.

Selective coding. This phase of analysis is the process of integrating and refining the ordered categories from the axial coding phase into a coherent model or theory, usually based on a core or central category from the data. Based on the pattern of passages containing the four aforementioned properties, the "customizability" category emerged as the critical phenomenon around which a model grounded in the data was constructed, centering on the concept of actionable feedback. This is discussed in more detail in the findings section.

Findings

Feedback Characteristics

Based on the participants' reports, four characteristics were evident in the data that could be used to describe the nature of feedback received by outpatient personnel (i.e., physicians, advanced practitioners, and nurses, henceforth referred to collectively as "providers") at VHA facilities:

1. *timeliness* – the frequency with which providers receive feedback; some interviewees mentioned that they received performance measure feedback on a monthly basis, whereas others reported less frequent reports, or no feedback at all. Monthly or more frequent feedback reports were considered timely; quarterly or less frequent reports were considered untimely. We chose monthly feedback as the timeliness threshold, because given usual time intervals between appointments within VA, quarterly or less frequent feedback may not give the provider sufficient time to engage in behavior change in time for a patient's next appointment.
2. *individualization* -- the degree to which providers receive feedback about their own individual performance, as opposed to aggregated data at the team, clinic, or facility level.

3. *customizability* – some passages contained reports that individual providers could look at their performance data in a way that was meaningful to them, as illustrated in this passage:

   I've got my computer setup where I can just plug in the numbers, get a new set of numbers and then update my overall cumulative scores within ten, fifteen minutes. And that's what gets fed back very, very quickly.

   -- S.M., a clinician in a HPF

4. *punitiveness* – this concerns how the feedback is delivered. Mentions were made in some interviews that when performance was substandard, providers were approached in a non-punitive way whereas in other interviews, a much more disciplinary tone was observed:

   We're down here in the trenches and if something goes wrong, somebody pounds on our head. Otherwise, they leave us alone

   -- A.B., a clinician in a HPF

Feedback Characteristic Patterns in High and Low Performing Facilities

Table 2 summarizes the patterns of feedback use across the six facilities. Each property is discussed in more detail below.

Timeliness. All facilities reported delivering feedback in a timely manner. However, as can be seen from Table 2, the evidence for timeliness of feedback is more mixed in the low performing facilities than in the high performing facilities. Conflicting reports of timely and untimely feedback delivery were observed in the low performing facilities, whereas timely feedback delivery was clearly the dominant practice in high performing facilities:

And then we also do, what's basically called, excuse me, provider score cards for the VISN's and it will show exactly in which areas they were found lacking

throughout the entire process for all the CPG's, for all the PI's [performance indicators]. …. Q: And that's how often? A: About once a month.

-- R.R., a support management employee in a HPF

Individualization. As can be seen from the table, none of the low performing facilities provided individualized feedback to their providers. In most cases, individual providers received facility level data from the EPRP report:

To be honest, most of the monitoring has really been done through the EPRP data collection. If one looks at some of the other guidelines, such as our COPD guideline, there... we really don't have a formal system set up for monitoring that. So if one really looks at performance and outcomes, EPRP remains probably our primary source of those types of data.

--B.F., An executive level employee in a LPF

In contrast, all three high-performing facilities reported providing individual level data to their providers:

Feeding it back, the individual reports go back to the practitioners and providers so they would see for a specific patient that was reviewed for where actual outcomes were. And many of them take that information to heart and would actually look, go back to the medical records and say "oh yeah! You're right! I missed this" or "oh no! You guys didn't pick up this." And they'd go back and show us where they documented it and that would allow us to have the dialogue.

-- M.M., a support management employee in a HPF

Punitiveness. Two out of the three high performing facilities explicitly reported that they approached underperforming providers in a non-punitive way to help them achieve better adherence rates:

> *it's a little more than that. She [the chief of staff] sends out positive letters. She sends out suggestion for improvement letters. But at the same time the people on the provider fields know that Tom and I both do this review and I've offered many, many times to say if you've got a case that you don't understand why this didn't meet criteria, call me and we'll look at it together. And I think that's been a real positive for this place because if I can go over that particular case that applies to you, it's much more beneficial*
>
> *-- G.D., a support management employee in a HPF.*

❖❖❖

> *Oh, yeah by provider, by clinic, we track them by clinic. We can tell who and we don't use it punitively. We just say, we had one provider in particular that was not doing very well. And we just showed him data and "this is your comparative data" and all your other providers in the clinic are getting this done. And why are you not? And he's like "thank you for telling me" and he jumped up there and is doing as well as everybody else .*
>
> *-- M.B., a support management employee in a HPF*

In contrast, employees at one low performing facility made explicit mention of the punitive atmosphere associated with low guideline adherence rates:

> *Sometimes I almost thought that it was in the overall presentation. If it wasn't so threatening and if it was interactive and if it was, you can show me and we're going to work with you … then you can get a better buy-in than you can if just*

saying, this is it. Do it! Heads will roll! We'll chop off one finger and then we'll go for a hand and a foot, kind of thing!

-- C.C., a clinician in a LPF

For the rest of the facilities, however, there were insufficient reports in either direction to indicate the presence of a punitive or non-punitive approach to delivering feedback.

Customizability. No facilities reported having customizable reports or tools that allowed individual providers to customize their performance information to their needs. Some facilities, however, did report having some capability to customize (even though that capability was not being employed), as expressed by this respondent:

Yes, we could pull out, let's say, I could pull out all of the patients that have a reminder due with the diabetic foot that's a diabetic. And then I could see that two [providers] have 500 [patients with a reminder due]. You only have 100. Guess who's doing much better. … my reminder program can do that.

-- H.S., a support management employee in a HPF

These reports came exclusively from high performing facilities –however, there were a significant number of comments, both from high and low performing facilities, about the utility and desirability of having such information.

A Model of Actionable Feedback

From the pattern of the feedback properties reported by each of the facilities, a hierarchical ordering of these properties can be postulated to arrive at a model of actionable feedback. At a minimum, feedback must be timely in order to be useful or actionable -- one can easily imagine situations where the most thoughtful, personalized information would be useless if it were delivered too late. Next, feedback information must be about the right target; in this

case, since clinical practice guideline adherence is measured at an individual level (i.e., the data from which adherence measures are constructed concern individual level behaviors, such as ordering a test or performing an exam), feedback provided to individual providers should be about their individual performance rather than aggregated at a clinic or facility level to maximize its effectiveness (21;22). Third is non-punitiveness; feedback delivered in a nonpunitive way is less likely to be resisted by the recipient regardless of content (15;23;24), thus making it more actionable. Finally, is customizability -- the recipient's ability to manipulate the incoming feedback in such a way that it is more comprehensible or actionable. Customizability engages the individual with the data, making him/her an active participant in the service making process, rather than being a passive recipient of information. Figure 1 graphically depicts this hierarchical ordering, which constitutes the tenets of the model emergent from the data, which we have termed "actionable feedback."

The hierarchical ordering proposed above can be seen reflected in the data. As can be seen from table 2, four out of the six facilities reported using the EPRP data to deliver timely feedback to their providers. The three high performing facilities provided individualized feedback to their providers, whereas the low performing facilities indicated that they used facility level, rather than provider-specific reports as a source of feedback. Only the top two performing facilities specifically indicated that they approached the delivery of feedback in a nonpunitive way, whereas no evidence of this existed one way or the other in the other facilities (save for one low performing facility, who reported explicit instances of punitive feedback delivery). No facilities reported providing their providers with the ability to customize their own individual performance data, although all facilities did report that it would be nice if it were possible to do this. Thus, as we move up the facility rankings from the lowest to the highest performer, fewer of

the properties appear to be present. This hierarchical ordering thus leads us to postulate the underlying dimension of "actionable feedback."

Discussion

This research employed a qualitative, grounded theory approach to studying differences in how high and low performing facilities employed their clinical audit data as a source of feedback. Findings indicated that high performers provided feedback in a timely manner, specific to individual providers, and in a non-punitive fashion, whereas low performers were more variable in their timeliness, and relied on facility-level reports provided by an external chart abstractor (EPRP) as a source of feedback, with one facility reporting a punitive atmosphere toward guideline adherence. The concept of actionable feedback emerged as the grounded theory in these data, with timeliness, individualization, non-punitiveness, and customizability as hierarchically ordered properties of actionable feedback.

The emergent model described above is consistent with existing individual feedback theories and research. Feedback intervention theory (FIT) (15) posits that feedback should be timely, focused on the details of the task, particularly on information that helps the recipient of the feedback see how his/her behavior needs to change in order to improve performance, and delivered in a goal setting context in order to have a positive impact on performance. These propositions have been supported by empirical research. One of the characteristics in the emergent model is timeliness, consistent with feedback intervention theory and empirical research.(13) The need for individualized feedback has been demonstrated in the organizational literature, (21;22)consistent with the individualization characteristic in our model, customizability can be considered different components of correct solution information. Feedback delivered in a non-punitive way has been empirically linked to increased likelihood of

feedback acceptance (25), a critical moderator of the relationship between feedback and performance(26). Finally, the effect of customizable feedback on feedback acceptance and subsequent performance has not been directly examined in the literature, and should be the object for future research. However, a positive effect of feedback customizability on improved guideline adherence can be predicted based on FIT's proposition that feedback interventions that direct attention to the details of the task are likely to result in improved performance. The ability of the provider to customize his or her specific performance data into something that is meaningful to him/her is likely to direct attention to the details of the performance measure in question, thereby increasing the likelihood of subsequent performance improvements.

The present research has several implications for facilities using clinical audit data as a source of feedback for their providers. A notable feature of the EPRP abstractions used by some facilities is that performance scores are based on a 10% (sometimes lower) sampling rate, which can be problematic when trying to use these data as feedback for individual providers. In light of this, one high performing facility began abstracting 100% of their charts, in order to feedback accurate data to its providers. A second implication concerns the underused capability of electronic medical records (particularly the system used within VA) to generate provider-specific performance reports; there were several mentions in the interviews that people knew about the capacity of CPRS, the VA's electronic medical record system, to generate physician-specific reports, but that this capacity was not being used. Given this research's findings with respect to individualization of data, a more concerted effort to make this capability well understood and usable could be beneficial. VA is already taking steps in this direction, with a re-engineering of CPRS to make template creation and report generation a simpler task for the user, and by offering training on the use of these tools system wide(27).

Limitations

First, the study's relatively small sample size of 6 facilities, three in each performance condition, potentially limits the generalizability of our results. The strategies were reported by multiple individuals at each facility, however; usually between 10-25 people were interviewed at each facility, for a total of 102 respondents (see Table 1). Additionally, care was given in the initial selection of facilities in terms of size, geography, and primary care capabilities, to insure a range of types of facilities that would represent the spectrum of VA facilities.

Second, the density of reports (20 passages per facility) is somewhat low, which also potentially limits the strength of the findings. However, participants were not explicitly interviewed on the subject of performance feedback, but rather on more general strategies and facilitators of clinical practice guideline implementation. Given the large domain of other available strategies and facilitators that participants mentioned (28), the consistency with which the feedback theme repeats itself across the six facilities strengthens the credibility of these findings, despite the low report density.

Conclusion and Future Directions

We conclude that facilities with a record of successful guideline adherence tend to deliver more timely, individualized, and non-punitive feedback to providers about their individual guideline adherence than facilities with a poor record of guideline adherence, and that, consistent with organizational research, feedback intervention characteristics may influence the feedback's effectiveness at changing desired behaviors. Future research should more fully explore the nature and effects of feedback characteristics on their effectiveness in clinical settings, as well as on the utility of customizing clinical audit data so that it is meaningful to individual providers.

Abbreviations Used

CPG – Clinical Practice Guidelines

EPRP – External Peer Review Program

FIT – Feedback Intervention Theory

HPF – High Performing Facilities

IQR – Interquartile Range

LPF – Low Performing Facilities

OQP – Office of Quality and Performance

VA – Veterans Affairs

VAMC – Veterans Affairs Medical Center

Competing Interests

The research reported here was supported by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service (HSR&D) (CPI #99-129). All three authors' salaries are supported in part by the Department of Veterans Affairs. The authors declare they have no other competing interests, financial or non-financial.

Author Contributions

Dr. Hysong interviewed participants, coded interview transcripts, and was principally responsible for the research idea, design, analyses, and drafts for this manuscript. Dr. Best was involved in all aspects of the study, including project management, participant interviews, coding interview transcripts, and editing of manuscript drafts. Dr. Pugh is the principal investigator of the grant that funded the work presented in this manuscript; she was principally responsible for the research design and project management of the research grant that made this manuscript possible; she also participated in conducting interviews and editing drafts of this manuscript. All authors read and approved the final manuscript.

Reference List

(1) Harvey PA, Murphy MC, Dornom E, Berlowitz DJ, Lim WK, Jackson B. Implementing evidence-based guidelines:inpatient management of chronic obstructive pulmonary disease. Internal medicine journal 2005 Mar;35(3):151-5.

(2) Baker R, Falconer Smith J, Lambert PC. Randomised controlled trial of the effectiveness of feedback in improving test ordering in general practice. Scandinavian Journal of Primary Health Care 2003;21(4):219-23.

(3) Eccles M, Steen N, Grimshaw J, Thomas L, McNamee P, Soutter J, et al. Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. Lancet 2001;357(9266):1406-9.

(4) Goff DC, Jr., Massing MW, Bertoni AG, Davis J, Ambrosius WT, McArdle J, et al. Enhancing quality of heart failure care in managed Medicare and Medicaid in North Carolina: results of the North Carolina Achieving Cardiac Excellence (NC ACE) Project. American Heart Journal 2005 Oct;150(4):717-24.

(5) Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technology Assessment (Winchester, England) 2004;8(6):iii-iiv.

(6) Jamtvedt G, Young JM, Kristoffersen DT, Thomson O'Brien MA, Oxman AD. Audit and feedback: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2003;G:\Sylvia's My Documents from Laptop\Articles\Implementation(3):CD000259.

(7) Bordley WC, Chelminski A, Margolis PA, Kraus R, Szilagyi PG, Vann JJ. The effect of audit and feedback on immunization delivery: a systematic review. American Journal of Preventive Medicine 2000;18(4):343.

(8) Foy R, Eccles MP, Jamtvedt G, Young J, Grimshaw JM, Baker R. What do we know about how to do audit and feedback? Pitfalls in applying evidence from a systematic review. BMC health services research 2005;5:50.

(9) Thomson O'Brien MA, Oxman AD, Davis DA, Haynes RB, Freemantle N, Harvey EL. Audit and feedback versus alternative strategies: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2000;(2):CD000260.

(10) Robinson MB. Evaluation of medical audit. J Epidemiol Community Health 1994;48:435-40.

(11) Stull MK. Performance feedback: a question of source credibility. The Journal of nursing administration 1986 Apr;16(4):17.

(12) Exworthy M, Wilkinson EK, McColl A, Moore M, Roderick P, Smith H, et al. The role of performance indicators in changing the autonomy of the general practice profession in the UK. Social science & medicine 2003;56(7):1493.

(13) Bradley EH, Holmboe ES, Mattera JA, Roumanis SA, Radford MJ, Krumholz HM. Data feedback efforts in quality improvement: lessons learned from US hospitals. Quality & Safety in Health Care 2004 Feb;13(1):26-31.

(14) Vancouver JB, Tischner EC. The Effect of Feedback Sign on Task Performance Depends on Self-Concept Discrepancies. Journal of Applied Psychology 2004;89(6):1092-8.

(15) 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.

(16) Hysong S, Best R, Pugh J, Moore F. Not of One Mind: Mental Models of Clinical Practice Guidelines in the VA. Health Serv Res 2005;40(3):823-42.

(17) Daft RL, Lewin AY. Can organization studies begin to break out of the normal science straitjacket? An editorial essay. Organization Science 1990;1(1):1-9.

(18) Strauss AL, Corbin J. Basics of qualitative research : techniques and procedures for developing grounded theory. 2nd ed. Thousand Oaks, CA: Sage Publications.; 1998.

(19) Glaser BG, Strauss AL. The Discovery of Grounded Theory. Chicago: Aldine Publishing Co.; 1967.

(20) Atlas.ti - Short User's Manual [computer program]. Version 4.2 Build 061. Berlin, Germany: Scientific Software Development; 2001.

(21) Deshon RP, Kozlowski SWJ, Schmidt AM, Milner KR, Wiechmann D. A multiple-goal, multilevel model of feedback effects on the regulation of individual and team performance. Journal of Applied Psychology 2004;89(6):1035-56.

(22) Archer-Kath J, Johnson DW, Johnson RJ. Individual Versus Group Feedback in Cooperative Groups. Journal of Social Psychology 1994 Oct;134(5):681-94.

(23) Hauenstein NMA. Training Raters to Increase the Accuracy of Appraisals and the Usefulness of Feedback. In: Smither J, editor. Performance Appraisal: State of the Art in Practice.San Francisco: JOSSEY-BASS; 1998.

(24) Nease AA, Mudgett B, Quinones MA. Relationships among feedback sign, self-efficacy, and acceptance of performance feedback. Journal of Applied Psychology 1999;84(5):806-14.

(25) Leung K, Su S, Morris MW. When is criticism not constructive? The roles of fairness perceptions and dispositional attributions in employee acceptance of critical supervisory feedback. [References].
Ref Type: Generic

(26) Kinicki AJ, Prussia GE, Wu B, Kee-Ryan FM. A covariance structure analysis of employees' response to performance feedback. Journal of Applied Psychology 2004;89(6):1057-69.

(27) VHA 2005 eHealth University. Veterans Health Administration 2005 June 20 [cited 2006 Jan 4];Available from: URL: http://vaww.vehu.med.va.gov/vehu2005/Default.cfm

(28) Hysong SJ, Pugh JA, Best RG. Clinical Practice Guideline Implementation Patterns In VHA Outpatient Clinics. Manuscript submitted for publication. 2005.

Figure Legend

Figure 1. A Model of Actionable Feedback

Table 1. Number of Participants by Facility and Hierarchical Level.

	Facility	***Hierarchical Level***			***Total # of***	***Total # of***
		Primary Care Personnel	Middle/Support Management	Facility Leadership	***Participants***	***Interviews***
High Performers	1	14	2	3	19	8
	2	6	10	7	23	14
	3	7	4	3	14	9
Low Performers	4	4	8	4	16	8
	5	3	4	4	11	7
	6	7	10	2	19	8
	Total	41	38	23	102	54

Note: facilities are listed in decreasing order of performance. No significant differences in the distributions of participants were found by facility or hierarchical level ($\chi^2_{10} = 17.4$, n.s.).

Table 2. Patterns of feedback properties by facility.

	High Performers			Low Performers		
PROPERTY	1	2	3	4	5	6
Timely	E	E	E	E	C	C
Individualized	E	E	C	N	N	N
Non-Punitive	E	E	I	I	N	I
Customizable	I	I	I	N	N	I

Note: E= Evidence was observed that the property in question was present at that facility; c = conflicting evidence; I = insufficient evidence; N = negative evidence, i.e., evidence that the opposite property was present (e.g., an N for facility 4 on individualized means that there is evidence that the feedback is not individualized)

Appendix: Interview Guide and Protocol Notes

Notes on Interview Protocol

Interviewers used the guide below to conduct participant interviews, using a semi-structured format. Interviewers were not required to use the probes below, rather, they were included as aids to facilitate the interviewer's task, and to illustrate the type of information for which they were to probe. Similarly, although the questions are listed in the suggested order, interviewers were free to change the order of the questions to better fit the flow of the interview.

Interviews were scheduled to be one hour in length, with ½ hour in between interviews for interviewers to compile notes on a given interview and conduct administrative tasks (e.g., labeling the interviews on the memory card, recording interviewee information in participant record). In some cases, the interviews went somewhat over the one-hour mark, however, never more than approximately 10 minutes. In a very few number of instances, the participants' comments were concise enough that the interview ended before the one hour mark. In general, however, most interviews naturally lasted approximately one hour.

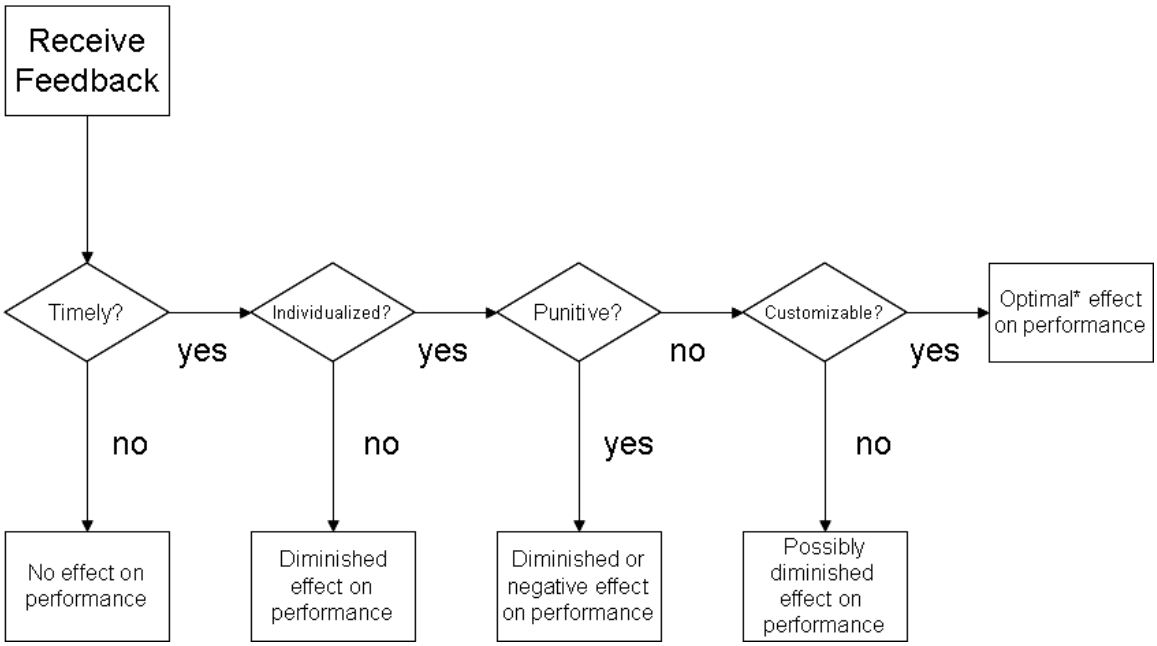
Interview Guide

CONCEPT TAPPED	PRIMARY QUESTION	POSSIBLE PROBES
Quality of Care in General	1. How do you or your staff identify quality of care issues in need of improvement for your **OUTPATIENT** primary care clinics?	*Probe for explicit processes (e.g., strategic planning, balanced score cards, data that is monitored, etc.)*
		a. Who would be responsible for initiating and carrying out such efforts?
		b. Who would be responsible for monitoring such efforts?
Mental Models of Clinical Practice Guidelines (CPG)	2. What does the term "Clinical Practice Guidelines" mean to you?	a. What role do you see for clinical practice guideline use as a method for improving quality of care?
		b. Do you believe clinical practice guidelines are effective for improving quality of care? Please explain.
		➢ If no, follow up with, "Despite your beliefs what is your experience?
	3. How do guidelines help you improve the quality of care you provide your patients?	a. As a source of data feedback?
		b. How is data collected and utilized in your facility to improve the quality of patient care (e.g., administrative "scorekeeping" or as feedback for improving the quality of care)?
		c. Was EPRP data or other data on performance distributed?
		d. Did EPRP results affect individual performance evaluations?
		e. Does the facility collect clinical outcome data (mortality, readmission, functional status) related to the guideline?

CONCEPT TAPPED	PRIMARY QUESTION	POSSIBLE PROBES
CPG Success Story	4. Could you tell us the story of a time you and your team successfully implemented a clinical practice guideline (e.g., smoking cessation, depression screening, Diabetes mellitus, hypertension, etc.)?	*Probe for the Who, What, When, Where, & How of the story.* a. What were the steps? b. Who was involved? To what extent are clinicians involved in determining how to implement guidelines? c. How was this guideline effort brought to the attention of clinicians and managers in your facility? (e.g., formal meetings, guideline champions, grand rounds, e-mail distributions, web sites, etc)? d. To what extent were committees (one steering committee for all guidelines or guideline specific committees) used to implement guidelines? e. What made it a success?
CPG Training & Development	5. Please describe the training (i.e., professional development) that clinicians have received for implementing guidelines.	a. Would clinicians say they have been provided adequate support for professional development with respect to CPG implementation? b. Any training in the use of technology (e.g., CPRS, clinical reminders, etc.)? c. CME credit?

CONCEPT TAPPED	PRIMARY QUESTION	POSSIBLE PROBES
Facilitators	6. What are the most important factors that facilitate guideline implementation?	a. Technology (CPRS, clinical reminders)?
		b. Targeted educational or training programs, patient specific reminder systems, workshops, retreats?
		c. Incentives (e.g., monetary, extra time off from work, gift certificates, etc.)?
		d. Mentoring or coaching?
		e. Additional resources (e.g., equipment, staff, etc.)?
		f. Social Factors such as teamwork or networks?
		g. Representation from a diversity of service lines?
		h. Presence of a guideline champion?
		i. Supportive leadership (i.e., VISN and/or facility)?
		j. Pocket cards or "lite" versions of the guidelines?
Barriers	7. What are the most important factors that hinder guideline implementation?	a. Lack of resources, or staff?
		b. Time (i.e., patient interactions are targeted for 20 minutes)?
		c. Lack of training?
		d. Not enough support?
		e. Financial?

CONCEPT TAPPED	PRIMARY QUESTION	POSSIBLE PROBES
Innovations	8. Were there any changes or redesigns in the clinical practices or equipment that supported the use of CPGs.	a. How were forms/procedures or reports changed to support adherence to guidelines? b. How were the responsibilities of nurses, aides, other personnel changed to support adherence? c. How were resources allocated/reallocated to support adherence?
Structural, logistic, and organizational factors	9. Please describe any other conditions that may influence CPG implementation?	a. Size of the facility? b. Academic affiliation? c. Competition with other QI initiatives? d. Location (e.g., remote vs. main facility)?



*The use of the term optimal to describe the effect on performance is relative – by this we mean optimal, given the variables in the emergent model. There are certainly other factors which could impact performance, although they are not exhibited here.

Figure 1