Information in practice

Integrating service development with evaluation in telehealthcare: an ethnographic study

Tracy L Finch, Carl R May, Frances S Mair, Maggie Mort, Linda Gask

Abstract

Objectives To identify issues that facilitate the successful integration of evaluation and development of telehealthcare services.

Design Ethnographic study using various qualitative research techniques to obtain data from several sources, including in-depth semistructured interviews, project steering group meetings, and public telehealthcare meetings.

Setting Seven telehealthcare evaluation projects (four randomised controlled trials and three pragmatic service evaluations) in the United Kingdom, studied over two years. Projects spanned a range of specialties—dermatology, psychiatry, respiratory medicine, cardiology, and oncology.

Participants Clinicians, managers, technical experts, and researchers involved in the projects.

Results and discussion Key problems in successfully integrating evaluation and service development in telehealthcare are, firstly, defining existing clinical practices (and anticipating changes) in ways that permit measurement; secondly, managing additional workload and conflicting responsibilities brought about by combining clinical and research responsibilities (including managing risk); and, thirdly, understanding various perspectives on effectiveness and the limitations of evaluation results beyond the context of the research study.

Conclusions Combined implementation and evaluation of telehealthcare systems is complex, and is often underestimated. The distinction between quantitative outcomes and the workability of the system is important for producing evaluative knowledge that is of practical value. More pragmatic approaches to evaluation, that permit both quantitative and qualitative methods, are required to improve the quality of such research and its relevance for service provision in the NHS.

Introduction

The promise of telehealthcare is that it might revolutionise the practice of medicine by enabling remote interaction between clinicians and patients, through the use of information and communications technologies such as interactive video, digital imaging, and electronic data transmission.[1][2] For policy makers and clinicians, telehealthcare offers the potential to solve problems of structural and spatial inequalities of access to specialist care, and to increase the speed of referral and management decisions.[3]

In practice, however, telehealthcare is somewhat contentious and unstable.[4] Although its proponents value the potential organisational benefits it may bring, others express concern about its implications for the practice of medicine, particularly in relation to the doctor-patient interaction.[5] Concerns about clinical risk and potential litigation[6] and, internationally, ongoing difficulties relating to licensure and reimbursement[7] may add to the resistance to telehealthcare in practice.

The production of evidence about the safety and effectiveness of telehealthcare is therefore vital for its progression. Although there have been many trials of telehealthcare in Britain and elsewhere, such services typically fail to become part of routine healthcare delivery.[8] This makes achieving sufficient levels of use of telehealthcare services to provide meaningful evaluations difficult.[9] The existing evidence base for telehealthcare is therefore not as strong as some of its champions have suggested.[10–12]

Understanding how this evidence base is constructed is important because there are concerns about the utility of applying medical models of evaluation to technological systems.[13][14] Research suggests fundamental problems in integrating telehealthcare into systems of professional practice in everyday settings.[15][16] Our study was intended to further understanding of this knowledge production process by exploring the methodological issues that arise when integrating evaluation with the (often experimental) development of telehealthcare services. We identify lessons that may improve future evaluations of telehealthcare to better inform the implementation of services.

Methods

Participants and settings

Between 2000 and 2002, we undertook an ethnographic study—using a variety of qualitative techniques to study telehealthcare projects and their development in depth and over time[17]—of factors that promote and inhibit the effective evaluation of telehealthcare. We studied seven telehealthcare evaluation projects in a variety of specialties—dermatology, psychiatry, respiratory medicine, cardiology, and oncology. Projects were chosen to represent good variability in specialties, settings, and evaluation methods (see box 1 for details). We identified project leaders from websites and databases and then asked them if we could include their project in our study. With the help of project leaders, we identified key informants within the project teams and asked them if we could record interviews with them. We obtained appropriate ethical approval and followed stringent procedures to ensure the anonymity of participants.

Data collected

We conducted interviews with key informants (n = 76), recorded steering group meetings and other meetings (n = 19), observed the projects, and analysed project documents. Interviewees included clinicians (nurses and physicians), researchers (principal investigators, research associates and assistants, health

economists, and statisticians), and technical experts associated with the specific projects. Participants were interviewed up to four times (typically two or three times). All authors conducted at least some interviews, as appropriate to each author's areas of expertise. We conducted the interviews in person at the site most convenient for the respondent (typically either an NHS or academic setting). We used a semistructured interview guide for all initial interviews, with subsequent interviews designed more specifically around the interviewee and project. Key themes for questioning included the history of the project, organisational relations between key participants, the purpose of or questions for the evaluation, methods of evaluation, and positive and negative experiences of evaluation.

Data analysis

Formal data for the analysis presented in this paper were the transcribed interviews. Although not presented here as data,

---

Box 1: Details of seven telehealthcare evaluation projects examined (medical specialties omitted to preserve anonymity)

Site 1: Store-and-forward system between primary and secondary care
Evaluation—Pragmatic service evaluation, survey of users' views
Problems—Long delays in set up, technical problems
Outcome—Project successfully completed but not intended for routine service delivery.

Site 2: Real time video based system between primary and secondary care
Evaluation—Randomised controlled trial of clinical effectiveness (intended), qualitative study of users' views (actual)
Problems—Problems with staffing and time commitments
Outcome—Randomised controlled trial not done

Site 3: Real time, video based system between hospital department and community
Evaluation—Randomised controlled trial, economic evaluation, qualitative study of users' views
Problems—Problems with the technology, cost constraints, initial professional resistance
Outcome—Project continuing

Site 4: Store-and-forward system between primary and secondary care
Evaluation—Randomised controlled trial of clinical and cost effectiveness, qualitative study of users' views
Problems—Major problems with recruitment, professional resistance
Outcome—Project completed

Site 5: System based on data transmission within secondary care
Evaluation—Pragmatic service evaluation intended (to assess utility and users' views)
Problems—Financial constraints, logistical problems of coordination
Outcome—Project failed

Site 6: Mixed system (video and data transmission)
Evaluation—Multicentre randomised controlled trial (clinical outcomes and cost measures)
Problems—Some problems with recruitment, but viewed as successful
Outcome—Project continuing

Site 7: Real time video-based system
Evaluation—Pragmatic service evaluation (assessing clinical effectiveness and users' views)
Problems—Some equipment and logistical difficulties
Outcome—Project continuing

---

observation and project documentation supported and strengthened our interpretation of interview data.

Our analysis of interview material was guided by the broad precepts of constant comparative analysis.[18] The trustworthiness of the data was established by involving all authors in data interpretation and the development of analytical themes. Six transcripts that varied greatly in perspective and content were coded individually by all authors and then discussed in team meetings. Within these meetings, we built a collaborative analysis for each of the sample transcripts by allowing individual interpretations to be offered, challenged, and resolved in order to refine thematic categories. We then merged these six collaborative analyses to form a comprehensive coding scheme for the data. The coding scheme was applied to two further transcripts by all authors for the purpose of refinement. Remaining transcripts were divided between authors for detailed analysis.

Results and discussion

Our results reveal the complexity of interaction between the evaluation of a telehealthcare service and its development. Here, we outline some of the major processes that emerged as evaluators struggled to integrate these tasks.

Evaluating the system: operationalising clinical practice

Evaluation generally involves producing knowledge about attributes of a new service or technology that relate to its potential effectiveness from various perspectives. The task of constructing measures to assess these attributes is further complicated if the research is conducted in parallel with service development. It requires not only identifying and specifying existing clinical knowledge and practices, but also anticipating ways in which practices will be changed in the new system and designing research instruments that capture such changes. Our respondents spoke of "hard" and "soft" outcomes in ways that prioritised quantitative methods of evaluation, assuming that "hard" outcomes were those that mattered most. However, participants found it difficult to define clinical knowledge and practices in terms that permitted such measurement (see box 2). This problem extended beyond measuring clinical outcomes to other effects of the telehealthcare system, such as patients' and professionals' views and particularly to the cost effectiveness of telehealthcare.

The difficulties experienced in defining clinical practice contributed to another problem for evaluation—recruitment. Although this problem is certainly not specific to telehealthcare,[19] we find here a particular impediment to recruitment, as clinicians using the telehealthcare system come to identify more and more characteristics of patients and their conditions that make them inappropriate candidates for telehealthcare (see box 3). The experience described here is common to other telehealthcare projects. It seems to reflect a growing understanding of the limited capacity of telehealthcare systems to accommodate clinical practice in the way that it is routinely enacted. Increasing exclusion criteria not only reduces the possibility of achieving recruitment targets but, more importantly, limits the claims that can be made for the effectiveness of the telehealthcare system for the broader patient population (box 3).

Managing conflicting demands of service provision and evaluation

Integrating service development with evaluation often requires clinical staff to perform additional research tasks. Evaluators expressed difficulties in ensuring complete and accurate data collection, because clinicians were inexperienced in research or

Box 2: Constructing outcome measures

"Trying to evaluate a [medical specialty] service as an outcome measure is very difficult, because we don't have very many, particularly for the majority of conditions that we look at, [names] For cancer, you can do, because you look at the recurrence rates and all the rest of it; but [names] there aren't cures for it so you're looking at things that are much softer as far as outcome, mainly around issues involving quality of life and patient satisfaction."—Clinical leader, site 1

"If you're doing HTA [health technology assessment] in the rather broader sense [than clinical trials] [the studies are] additionally complicated. You cannot usually do a randomised trial, you have to do some other sort of design, you have to create your own measures. You can't just be satisfied with, 'Oh it's all right, mortality is the measure for a cancer trial,' it's a key measure, the other ones are fairly clearly spelt out, you've got to create measures. And I think if you add to that the problems of doing informatics research in general, as opposed to telemedicine in particular, I think you've got a whole series of other issues to do with attempts to apply clinical paradigms to informatics research."—Clinical leader, site 4

Box 3: Recruitment of patients and validity of results

"Really, in the last two months or so I'd say that recruitment has slipped, certainly at our site, because the winter months tend to generate more patients [with the condition], and what I'm finding, though, is that they're either end stage … and they're just too sick, there's no way that they could manage with equipment at home, they're more respite care really. All the patients have just said no, they just don't want anything to do with it."—Research nurse, site 6

"The recruitment's a crucial issue, not just for the actual achievement of numbers but also for the external validity, which I think is very important."—Clinical leader, site 4

did not have time to complete both clinical and research duties (see box 4). Research is often not considered a priority when it competes with the demands of service provision. The importance of accommodating everyday "workability" within a trial design is shown by the dilemma posed when nurses working an "extra" telehealthcare service felt that the trial created obligations that compromised their role as care providers by placing unmanageable demands on their time (box 4).

A major part of the conflict around service provision and evaluation concerned the management of risk. In routine settings, clinical practice is based on minimising risks to patients. For clinicians in this study, the introduction of a telehealthcare system and its evaluation highlighted the possibility of increased risk from what they perceived as new forms of practice. Clinicians involved in telehealthcare evaluations were thus faced with a dilemma: they had concerns about patient safety and their own personal liability (see box 5) but needed to engage with these new technologies in order to prove their safety.

In the evaluations we examined, ensuring that the system was safe was clearly the priority. Often individual clinicians managed perceived risk to patients by reverting to the default model of service provision (box 5). At site 1 (providing the first quote in box 5), this focus on risk minimisation meant that 60% of patients had to be recalled to a conventional doctor-patient encounter. The priority of clinical safety therefore needed to be built into the research protocol: clinicians needed flexibility to exercise their judgment and revert to standard care if they con-

sidered a patient was at risk, though too much flexibility could invalidate the research. At site 3 (the second quote in box 5), it became apparent that an unexpectedly large proportion of eligible patients were being excluded because of healthcare providers' concerns about clinical risk and their lack of confidence in using the new system. Thus, professional assessments of risk attributed to the system were perceived and acted on in ways that could (and did) adversely affect evaluation.

Making sense of study findings

Accurately understanding the effects of a telehealthcare system is essential if the study results are to inform further service development. However, we observed that some evaluators found it difficult to determine how much the results of their study reflected effects of the telehealthcare system and how much they were a product of the research process and the disruption to normal practice that was caused by it (see box 6). Clinicians and researchers themselves recognised these limits, knowing that their studies reflected experimental work that was sometimes

Box 4: Managing conflicting demands of service provision and evaluation

"Partly, but I think the GPs say that there are two things that have added to their burden, and they find it difficult to say which is worse. One is the actual telemedicine, and the other is the research bit, including the ethics bit.… The telemedicine has added to their work, but it would have been easier if they'd just been implementing a telemedicine project. I can actually see that if you can produce a telemedicine project that makes GPs' lives easier they'll like it, but nobody will ever learn anything about it."—Clinical leader, site 4

"And the trouble really is, just because the equipment was late coming we were then trying to start at peak time [for this condition] which means that at [the hospital] there they're just lined up on trolleys, so it's very hard for the nurses to say, 'Sorry, we're not going to see this patient because we're doing this trial.'"—Clinical leader, site 3

Box 5: Managing risk by protocol

"But there is incorrect diagnosis all the time." (Interviewer) "But diagnosis incorrect through a telemedicine application—and should that individual have been brought in to the hospital to have it checked and was the amount of information you actually physically had to make the diagnosis sufficient without bringing the patient in? And probably misdiagnosis could be slightly higher.… Then it needs to come down on the basis of what's the protocol for lesions—and that the protocol may be that we actually say that, for the doctor to have the degree of confidence, that he would ask the patient to strip so he can actually look at the whole of the body and to ask the question 'Have you got any more lesions in any other place?' Now would that sort of process stand up then in court that 'Yes, we've gone through this process, this is the protocol for lesions, and we follow it to the law' and the doctor has a degree of confidence that the nurse actually follows that protocol."—Service manager, site 1

"We will be monitoring how often [they default to standard care], but they're allowed to do that so I don't think safety will be an issue. But the issue of safety is something we've taken extremely seriously, and that's just to do with the lack of knowledge about the legal status of telemedicine. So we'll be erring on the side of overcautious in terms of how our protocol is set up because we definitely don't want to ever be putting a patient at risk at all, in any shape or form."—Clinical leader, site 3

Box 6: Ambiguity in attribution of effects

"I think what it's not brought out, or we can't conclusively draw out, is whether it has been the trial and all the problems with the trial or whether it's the telemedicine per se, and I think that's a big problem and that's what I'm having to write up—that we're not quite sure."—Research associate, site 4

"Yes, but the thing is, you see, like all these things in life, it's not where the truth lies, it's where it's perceived to lie. They see the waiting list numbers going down, therefore it's working, that's the simple equation. Because all they're concerned about is numbers on the waiting list, because that's what they get the pressure upon."—Clinical leader, site 1

Box 7: Randomised controlled trials versus pragmatic evaluation

"I think you need a randomised study to evaluate it in general. I think you need some form of controlled comparison, and RCTs give you that. And the associated rigor, I think, is important because it's very easy for technology to develop it's own inertia, and it's happened time and again in healthcare that technologies have been adopted when really they don't stand up to a rigorous evaluation. So in that sense an RCT I would see as being the definitive evaluation."-Statistician, site 3

"I suppose there's a bit of a worry about that—that we've suddenly introduced something that's slightly different to the protocol. But I think, given the way recruitment has been running quite low, the trial again is pragmatic rather than explanatory. It's telemedicine being installed and taking the pictures and the camera—the process more than the actual pinning it down where you take the pictures in this form, that form, send them down this type of telephone line in these ideal conditions using this camera. I think we're never really going to get that; we're going to say it's the policy of telemedicine in this format that is broadly the same, or not as the case may be, as gold standard outpatients' appointment."-Statistician, site 4

---

considerably different from the experience of normal service provision. Difficulty in interpreting effects of the new systems was also sometimes a product of knock-on effects through different levels of service provision and depended greatly on which perspective of "effectiveness" was being considered (box 6). The problem of judging effectiveness was an ongoing problem for telehealthcare evaluators and contributes to the broader problem of translating research findings into everyday practice.

Randomised controlled trials versus pragmatic evaluation

The dominance of the randomised controlled trial (RCT) as the "gold standard" of medical research is clearly apparent in our research (box 7). In practice, however, many respondents re-evaluated the appropriateness of randomised controlled trials for assessing telehealthcare, having expressed disappointment about progress or uncertainty about the outcomes of their projects. Participants in studies applying such formal study designs found that trying to impose sufficient constraint on the system for the purpose of measurement conflicted with the dynamic nature of the health service environment, where some flexibility is necessary (see box 7).

The need for more pragmatic approaches to the evaluation of telehealthcare systems was thus apparent. Evaluators who had adopted non-randomised designs felt they were producing results they could use, even though they too experienced problems with integrating telehealthcare systems into existing

practice. However, because they were less restricted in their evaluation approaches, they were able to modify both clinical practice and technical systems more readily, and so improve the stability of the project overall. In doing so, respondents often drew a distinction between two types of knowledge—experimental quantitative knowledge about outcomes, and experiential qualitative knowledge about workability (see box 8). They regarded the former as having higher status, and saw its publication as the main objective of their work, but found the latter to be more useful, particularly for judging the utility of the system in practice but more generally for informing service development. These findings have important implications for commissioning processes, which must permit greater acknowledgement of the practical value of research methods that produce knowledge about processes rather than healthcare outcomes.

Conclusion

The complexity of a combined implementation and evaluation of a telehealthcare system is often underestimated in both the design and the conduct of evaluation studies. The requirement of stability for the evaluation protocol conflicts with the need for flexibility in the provision of health services to individual patients. This tension raises particular methodological issues, which centre around defining and measuring clinical practice, managing conflict between evaluation and service provision, and difficulty in interpreting study findings. Evaluating telehealthcare thus requires more pragmatic and flexible approaches to the production of evidence than those permitted within the rigid structures of controlled study designs. The issues identified in this paper, such as workability, must be given greater attention in the design of evaluation studies in order to improve both the quality of such research and its relevance for clinical practice.

Box 8: Different forms of knowledge

"It sounds like you're saying that you feel that the research won't necessarily tell you the answers you want in terms of 'Is this useful?'"—Interviewer

"No, no, I think it will. I think it's the other way round. I think it probably won't tell me 'Yes it's useful,' but hopefully it will tell me it's not dangerous. So it won't be able to demonstrate that it's useful—that will be up to me to say whether it's useful or not—but it should be able to tell me that I'm not putting patients at risk."—Clinician, site 7

"And how will you judge whether it's useful?"—Interviewer

"By seeing how positive we feel about using it. Because if it really is an effort to use it then it will turn out to be not useful, because it won't get used."—Clinician, site 7

What is already known on this subject?

Telehealthcare is a rapidly growing field of clinical activity and technical development

New technologies offer clinicians and policy makers the potential to solve structural problems around inequalities of service provision and distribution

Despite many pilot studies, telehealthcare has not yet penetrated practice in any systematic way

What this study adds

This ethnographic study of seven telehealthcare evaluation projects identified key difficulties that are experienced when evaluation and development of a telehealthcare service are combined

More pragmatic approaches to evaluation would improve both the quality of such research and its relevance for service provision in the NHS

1 Lenaghan J. *Rethinking IT and health.* London: Institute for Public Policy Research, 1998.
2 Bashshur RL. Critical issues in telemedicine. *Telemed J* 1997;3:113-26.
3 NHS Executive. *Information for health: an information strategy for the modern NHS 1998-2001.* London: NHS Executive, 1998.
4 Lehoux P, Sicotte C, Denis JL, Berg M, Lacroix A. The theory of use behind telemedicine: how compatible with physicians' clinical routines? *Soc Sci Med* 2002;54:889-904.
5 Miller EA. Telemedicine and doctor-patient communication: an analytical survey of the literature. *J Telemed Telecare* 2001;7:1-17.
6 Stanberry B. Legal ethical and risk issues in telemedicine. *Comput Methods Programs Biomed* 2001;64:225-33.
7 Sanders JH, Bashshur RL. Challenges to the implementation of telemedicine. *Telemed J* 1995;1:115-23.
8 May C, Mort M, Mair FS, Williams T. Factors affecting the adoption of telehealthcare technologies in the United Kingdom: the policy context and the problem of evidence. *Health Inf J* 2001;7:131-4.
9 Whitten P, Rowe-Adjibogoun J. Researching health communication technology intervention projects: The challenge of achieving utilization levels sufficient for evaluation. *J Technol Human Serv* (in press).
10 Urquhart C, Currell R, Lewis R, Wainwright P. Evaluating telemedicine: indications from a systematic review. In: *Current perspectives in healthcare computing 1999.* Weybridge: BHJC Books, 1999:186-91.
11 Whitten P, Mair F, Haycox A, May C, Williams T, Helmich S. Systematic review of cost effectiveness studies of telehealthcare interventions: economic benefits of new technology are assumed, rather than known. *BMJ* 2002;324:1434-7.
12 Mair F, Whitten P. Systematic review of studies of patient satisfaction with telemedicine. *BMJ* 2000;320:1517-20.
13 Heathfield H, Pitty D, Hanka R. Evaluating information technology in health care: barriers and challenges. *BMJ* 1998;316:1959-61.
14 Kaplan B. Evaluating informatics applications—some alternative approaches: theory, social interactionism, and call for methodological pluralism. *Int J Med Inf* 2001;64:39-55.
15 Mort M, May CR, Williams T. Remote doctors and absent patients: acting at a distance in telemedicine? *Sci Technol Human Values* 2003;28:274-95.
16 May C, Gask L, Atkinson T, Ellis N, Mair F, Esmail A. Resisting and promoting new technologies in clinical practice: the case of telepsychiatry. *Soc Sci Med* 2001;52:1889-901.
17 Hammersley M, Atkinson P. *Ethnography: principles in practice.* London: Routledge, 1995.
18 Glaser B, Strauss A. *The discovery of grounded theory: strategies for qualitative research.* Chicago: Aldine Publishing, 1967.
19 Prescott RJ, Counsell CE, Gillespie WJ, Grant AM, Russell IT, Kiauka S, et al. Factors that limit the quality, number and progress of randomised controlled trials. *Health Technol Assess* 1999;3(20):1-143.

Centre for Health Services Research, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 3DN
Tracy Finch *senior research associate*
Carl May *professor of medical sociology*

Department of Primary Care, University of Liverpool, Liverpool L69 3GB
Frances Mair *professor of primary care research and development*

Institute of Health Research, Lancaster University, Lancaster LA1 4YJ
Maggie Mort *senior lecturer*

School of Primary Care, University of Manchester, Manchester M14 5NP
Linda Gask *reader in community psychiatry*

Correspondence to: T Finch
Tracy.Finch@newcastle.ac.uk