JOHN ØVRETVEIT | PAULO SOUSA

obras avulsas
08

EDITORS
JOHN ØVRETVEIT | PAULO SOUSA

Quality and Safety Improvement Research:

Methods and Research Practice from the International Quality Improvement Research Network (QIRN)

EDIÇÕES DE SAÚDE

Quality and Safety Improvement Research:
Methods and Research Practice from QIRN

08

QUALITY AND SAFETY IMPROVEMENT RESEARCH: METHODS AND RESEARCH PRACTICE FROM THE INTERNATIONAL QUALITY IMPROVEMENT RESEARCH NETWORK

Original title: *Quality and Safety Improvement Research:*
Methods and Research Practice from the International Quality Improvement Research Network

Editors: John Øvretveit and Paulo Sousa

Legal deposit: 273 496/08

ISBN: 978-972-98811-8-3

Escola Nacional de Saúde Pública — Universidade Nova de Lisboa
Av. Padre Cruz — 1600-560 Lisboa, Portugal
Phone: 21 751 21 00
Fax: 21 758 27 54
e-mail: ensp@ensp.unl.pt
http: www.ensp.unl.pt

Printed by: Alfanumérico, Lda.

John Øvretveit
Paulo Sousa
(EDITORS)

QUALITY AND SAFETY IMPROVEMENT RESEARCH: METHODS AND RESEARCH PRACTICE FROM THE INTERNATIONAL QUALITY IMPROVEMENT RESEARCH NETWORK

UNIVERSIDADE NOVA DE LISBOA
ESCOLA NACIONAL DE SAÚDE PÚBLICA
KAROLINSKA INSTITUTET

Lisbon
2008

Contents

PART 6

COMPLETED RESEARCH AND LESSONS FOR OTHERS

APPENDICES

Preface

As directors and researchers we would both like to thank the contributors to this book and the many who have contributed to the International Quality Improvement Research Network (QIRN).

The network started in 2001 to enable researchers studying quality improvement to share and improve their research. It has provided over the years inspiring and extremely useful assistance to researchers and stimulated and developed quality improvement research in many countries.

In 2007 the National School of Public Health in Lisbon hosted the meeting and was able to support this publication.

We believe this will help and encourage quality and safety improvement researchers and especially those who have less support from colleagues working on similar issues.

Quality improvement research is still not strongly supported outside the USA and the open and frank discussion of the issue will, we believe, be of help and inspiration to others. This is especially important as the time has come when more managers, policy makers and clinical professionals are increasingly interested in using research evidence to improve their practices and programmes — one of the greatest tests for research.

Professor Constantino Sakellarides
Professor and Director
National School of Public Health
New University of Lisbon, Lisbon, Portugal

Professor Mats Brommels
Professor and Director, Medical Management Centre,
Karolinska Institutet, Stockholm, Sweden,
Professor of Health Services Management,
Department of Public Health,
University of Helsinki, Finland

FOREWORD

Research for Improvement

DONALD M. BERWICK*

From the viewpoint of external evaluation and analysis, the greatest strength of successful efforts to improve complex systems is also their greatest weakness: namely, that most of these efforts are inescapably local. They resist generalization. This can sound like a *"political"* statement — that more autonomy needs to be given to local actors, or a vague commitment to *"empowerment"* as a value in organizational life — but it is not political; it is technical. Complex systems, like health care, especially *human* complex systems have dynamics that defy simplistic cause-and-effect explanations, and that are highly sensitive to context and local details. When 30 hospitals try to make care safer, for example, the single story of averages, ranges, and generalizations for the group as a whole cannot capture the texture and lessons of the 30 *different* stories that group comprises. Researchers who use inappropriate methods to study this complexity impoverish their own learning and others.

The people actually trying to make changes and improvements in local settings understand this. They experience the contextual nature of knowledge as a daily challenge — a fact of working life. They experience it in part as frustration because the roadmaps to improvement that they read about in guidelines or journals seem so often not to lead them to the same destinations that the writers and researchers describe. *"But it worked so well there,"* they say, *"Why can't we make it work here?"* And, they will stay frustrated until, with patience and learning not very different from what a child who is learning to ride a bicycle, does they eventually, in the favorable cases, master what first confused them by making it their own. In complex systems, almost no change succeeds unchanged; some assembly is required.

There are some exceptions, some improvements, the fortunate few, are much more "technical" than "socio-technical," and much more linear than complex. Effective drugs are the best examples. The systems that get the right drugs into the right patients safely are complex ones, but the drugs, themselves, are not. You can prove that a vaccine prevents poliomyelitis, and that proof is as pertinent to small clinics as to

* President and CEO, Institute for Healthcare Improvement — IHI (dberwick1@ihi.org).

large hospitals and as informative to London as to Sydney. With proper stratification of patients and a few other variables, linear models of the effectiveness of surgical procedures and diagnostic maneuvers also allow useful generalizations from relatively simple study designs.

But, sadly the simple assumptions that evaluations can ignore contexts, that effects are approximately linear, and that broad summaries are useful do not allow students of improvement in health care to learn what they most helpfully can from the increasingly widespread improvement movement, itself. To the great disadvantage of both the community of evaluators and the community of improvers, many studies of system change do not take sufficient notice of the difference between evaluating simple systems and evaluating complex ones. They use designs and statistics that illuminate the former well but obscure important lessons in the latter. In reaching simplistic conclusions about complex, textured settings, they not only confuse themselves, but they inadvertently demoralize and confuse the clinicians, managers, and others in local settings who are trying to learn and to use what they can to do better over time.

This gap, between the world of research and the world of active, widespread, local improvement, is not inevitable. Indeed, closing it is a worthy, potentially exciting, and totally feasible developmental challenge for both. For the research community, bridging the gap will involve embracing a wider range of research and evaluation methods. Some of these methods already exist in ethnography, qualitative evaluation, and other fields abutting health services research. Others remain to be created. For the community of improvement activists and practitioners, disciplined measurement and serious introspection can help guide them efficiently toward better changes with better results.

The convener of the International Quality Improvement Research Network, John Øvretveit, and the host of the 2007 meeting, Paulo Sousa, are leaders who know well what gems lie ready for discovery in a productive, mutually respectful, and open-minded relationship between the two worlds. They have encouraged observation, innovation, and productive risk-taking among students of improvement, and they have embraced narrative, as well as quantitative, methods. In their words, *"Some of the papers in this collection use and develop scientific methods necessary to studying quality improvement where controlled trails or comparative studies are not possible, or where qualitative methods are often more appropriate for the subject and objectives of the research"*. And, refreshingly, their guidance to contributors include this: *"This book is not about perfect finished research, it is about research methods and challenges you face or faced on the journey, so that others can learn from you"*.

Encouragement like this will lead careful observers to glean the greatest harvest from systematic study of the widespread practical experience and local experiments

that constitute the health care improvement movement of our time. In a happy recursion, the editors of this book teach that the study of improvement can, itself, improve, and they show the way in both their own teaching and in their invitations to others. In this book, as in their larger work, they open a wide door for thoughtful people to report in many voices what they see when they try to understand and interpret the good-hearted, informative efforts of other thoughtful people. Together — those who make improvements and those who study them — can more effectively help health care become what it should.

PART 1
INTRODUCTION

Quality and Safety Research: Introduction by the Editors

JOHN ØVRETVEIT*
PAULO SOUSA**

This chapter explains the purpose of the book and gives an overview of its main parts. The international quality improvement research network (QIRN) started in 2001 to allow often isolated researchers to come together with others working on similar issues so that they could learn from each other and improve their research. One element is annual meetings which allow researchers to present their plans or work in progress and then get constructive advice and suggestions about literature or projects which they did not know about but which would help the research. The aim is to encourage the cumulative growth of knowledge by helping researchers relate their research to that of others, and to ensure they know of related research, often in fields or areas different to their own.

One purpose of this book is to continue this aim for those who could not come to the meetings. As a reader you cannot present research and get advice, but you can use the discussions presented in the papers in this book to strengthen both the scientific and practical contributions of your research.

As well as practical advice, the meetings boost our motivation and energy to continue the research and its dissemination to others. One aim of this book also is to stimulate research into quality improvement. This is especially in subjects which have been little studied but where there is a great potential for research to inform more effective action and draw attention to neglected issues. Examples are research into cross-professional and cross-organisation communication, research into the consequences for professionals and patients of being involved in an adverse event causing death or disability, or patient perceptions and experiences of the quality of health services in other countries. It is especially also in relation to scientific methods suited to studying some quality issues and which are less know or used in medical and

* Director of Research at Medical Management Centre, Karolinska Institutet, Stockholm, Sweden (jovret@aol.com).

** Assistant Professor at Lisbon School of Health Technology and Researcher at National School of Public Health, New University of Lisbon, Portugal (paulo.sousa@ensp.unl.pt).

health services research. These include ethnographic methods, qualitative synthesis, comparative cases study designs, and action evaluation. Some of the papers in this collection use and develop scientific methods necessary to studying quality improvement where controlled trails or comparative studies are not possible, or where qualitative methods are often more appropriate for the subject and objectives of the research. There are also papers using traditional methods which show how the researchers have worked to maximise the scientific contribution by achieving as many of the requirements of conventional methods such as those of evaluation design, survey research or use of hospital statistics.

The papers are a small selection from the many presentations at QIRN meetings over the years. We selected papers which show an area of research where we thought the researcher's experience planning and doing the research, and their methods would be of interest and help to other researchers. Our instructions to the authors was,

"The book, like the QIRN meetings, is not a conventional scientific conference or book of proceedings approach. The book, like the meetings, is to help researchers and stimulate them in what can be a difficult and lonely task, and to help them make their research plan and work better. They and other academics can get papers reporting findings and methods fairly easily now. This book is not about perfect finished research, it is about research methods and challenges you face or faced on the journey, so that others can learn from you. The paper you now write will be assessed for acceptance in the book in terms of its openness and honesty about the challenges you faced and the help that your discussion of this could provide to other researchers on the same journey you are making"

The format suggested for the papers is given in appendix 1 for others to use to help researchers to prepare for seminars or similar collections of papers. There are a number of more conventional papers, most of which are presented in the last part of the book.

The aim of the book therefore is to inspire, give ideas for other researchers to strengthen their research, and also to show ways to make the research more useful to decision-makers.

Overview

Because the book is primarily for researchers, the papers are grouped into the stages of research following Part 1 — Introduction; 2) Planning Q&SI research; 3) Initial data gathering; 4) After data gathering and doing the analysis; 5) Writing thesis and publication; 6) Completed research and lessons for others, from those looking back over their journey and completed papers. The authors were asked not

to write a conventional scientific paper, but to reflect on the practical and scientific issues and make suggestions for other researchers in a way which is not possible in a scientific format. Some authors have taken this opportunity and give very open and revealing reflections which will be of great use to other researchers following similar paths.

The papers cover some of the wide range of quality improvement research subjects. Patient safety research is one of the most important sub-categories of quality improvement research (QIR) and sometimes draws on different domains of knowledge and research methods to those in other quality research fields. Over the years some of the fields of quality improvement research which QIRN presentations have covered include:

By level of object of study (often the "unit" which interventions aim to change):

— patient or individual practitioner level (e.g. patient perceptions of services or practitioner views of barriers and assisters in guideline implementation);
— team or unit level (e.g. study of team communications or interventions to improve teamworking);
— facility level (e.g. hospital quality programme, or comparison of nursing home use of nurse training to improve quality);
— health system level (e.g. integration interventions to a system of hospitals and primary care providers to improve safety);
— regional or national level (e.g. accreditation programme or breakthrough collaborative);
— cross-national level (e.g. patient perceptions of health service quality in countries where they are not resident).

By subject area;

— the epidemiology of quality and safety problems (how widespread and serious is this problem, are there large variations?);
— quality costing and the economics of quality;
— development of measurement or indicators (e.g. safety culture surveys);
— evaluation of small scale interventions (e.g. team project);
— evaluation or case studies of large scale complex interventions (e.g. quality programmes in hospitals, health systems or regional or national programmes);
— qualitative research into provider or patient experiences;
— surveys of provider or patient perceptions of quality and safety;
— reviews of research or theories;
— cross case comparisons and syntheses.

The papers in this book cover many of these subjects presented at earlier QIRN meetings and also a sample from a few of the wide range of countries involved in the network: Brazil, Germany, Norway, Pakistan, Portugal, Saudi Arabia, Spain, Sweden, Switzerland, The Netherlands, Thailand, United Kingdom, and United States of America.

1) Introduction section

This section adds to this introduction other introductory papers. The next considers how to make evaluation research of quality intervention effectiveness more useful. There are two proposals: first that decisions makers should not wait for evidence from controlled trials to decide whether to introduce a large-scale or complex quality programme or change. They do need information from research, but should not delay making changes for interventions for which there is some adequate evidence. The key issues are, which evidence is sufficient to justify acting or not acting and which evidence do they need about implementation, costs and possible harm. The second proposal is that researchers do not only try to provide evidence of effectiveness but also need to provide decision makers with information about the cost of implementation and the risks of harm or waste. It proposes that the burden of proof required should be proportionate to the likely benefits and costs in terms of possible harm and cost and ease of implementation. It describes a decision process to assess these aspects where evidence is absent or weak.

The next chapter on strengthening the scientific value of quality improvement research concentrates on what the author calls the research "the middle ground" — this is the area between the one extreme of randomised controlled trial with its high internal but low external validity, and the other extremes of pragmatic trials or qualitative studies of patient's experiences. It is the area where research seek to study quality interventions and implementation (e.g. guidelines or national quality programmes) in their natural context, without controls, and also both provide knowledge of practical use and advance scientific theory about these interventions. It provides ideas for how to assess the quality of the research and strengthen it scientifically and makes seven recommendations, which include a variety of theory testing and theory building research methods to describe QIPs, study outcomes, understand how QIPs work to produce outcomes, and understand the assisting and hindering context factors.

The next chapter turns to how to increase the practical usefulness of quality improvement research. It considers what "actors" — clinicians, managers, and policy-

makers — need from this type of research and addresses these questions: what is "actionable evidence"? What is evidence? Should we carry out this improvement change? Which improvement should we do? and, How should we carry out the improvement? It concludes that actors will never have complete evidence but can make better use of the research, which is relevant; and that decisions should never be fully rational anyway — values and non-rational preferences do and should be part of the assessment.

Two other papers in the introduction consider new fields of QI research: patient safety research as a challenge for Public Health domain; and research into how to improve the management of chronic illnesses. Patient safety research is part of quality improvement research and a rapidly growing field. Health services are different in many respects from industries where safety methods have been applied. Therefore safety theories and methods do need to be adapted for different health service situations. To achieve this aim it is essential to develop research in patient safety domains. One of the challenges is how to adapt the interventions and how to involve staff in the implementation process. This chapter considers the research that is needed and some approaches as it is essential to raise awareness and mobilize resources towards patient safety research, which can only be strengthened through a workforce capable of both planning and undertaking research — because their multidisciplinary nature public health team are in good position to leadership this approach. The last chapter in the introduction discusses the implications for research of a study of quality improvement for chronic health conditions, noting some of the challenges and solutions that are also relevant for other QI research studies.

2) Planning research

Part one of the book on planning quality and safety improvement research starts with a discussion of how to decide the focus of the research. The paper by Essen is on the subject of the sometimes conflicting pressures of standardization and customization which characterises many quality improvements. How can standardization also allow the customisation which most patients want? Which parts of their pathway should be standardized and which customised? Can you standardise customisation? These are some of the questions addressed in the research thesis which the author is presenting in 2008. But the paper in this book is really about how to define the research problem and the search for a fruitful question. In this essay the researcher honestly and revealingly reflects on the steps she took in her research in a way which is valuable to other researchers in the early part of their research. It also

describes the practical problems often encountered in studying quality improvement: the timescale in healthcare is different from the timescale for research. In this case there was a two year delay in the implementation of the information technology in the elderly care organisation which would allow the data collection: how the researcher adjusted and made the best of the situation and her journey to find the fruitful question is shared for all to learn from.

The paper by Aloqbi presents the plan for research which the QIRN meeting helped the researcher to formulate — in this case the subject is the use of total quality management methods in Saudi Arabian Medical Laboratories. It shows how the researcher defined the problem, set the questions and chose the design and data gathering methods and was selected for its thorough discussion of research choices and methods. It also shows one way to present a plan for research and to ensure the plan has sufficient detail about data gathering but also will allow some flexibility to follow up discoveries during the research which this approach requires.

3) Initial data gathering

The paper by Franx was written as the research team was starting data gathering in a study of a national quality improvement breakthrough collaboratives in mental health — specifically care depressed patients. It describes the design choices the researcher made and the issues they faced up to this stage. It demonstrates well the research issues when studying large scale complex interventions to social systems and shows how the researchers planned and designed an innovative quasi-experimental, controlled before and after study which includes a process evaluation and a cost-effectiveness assessment. The discussion also notes how events and requirements in the care sector reduces the strength of evidence of the study — 8 of the 18 teams in the collaborative could not give time to be involved in the study, *"As a consequence of this suboptimal inclusion, our study risks to be underpowered. Moreover, we wonder if the included teams will generate the number of patients needed in the study".*

Duckers et al., also describes the research issues and choices in studying national breakthrough collaboratives in the Netherlands, at the stage of having collected some of the data and trying to analyse these data. It gives a interesting discussion of the practical issues in studying large scale quality programmes, showing again how the practical issues often constrain the scientific choices which the researchers would have liked to have made. It also shows how the researchers have continually attempted to relate their research to the existing literature on the subject and build on this knowledge and sometimes contribute to it.

4) After data gathering and doing the analysis

The next paper by Mendes et al., is a pilot study in a larger programme of research but also a completed study which has already had an impact in Brazil. It studied adverse events through experts using a standard tool to assess a random sample of charts of adult patients of a teaching hospital. It found the Canadian Adverse Events Study (CAES) assessment tool was a feasible for Brazilian and other similar hospitals in other countries. It also found the proportion of preventable AEs was far higher than that reported in others studies using similar methods in higher resource settings. One important aspect of the study is the balanced self-criticism of the researchers and their assessment of the limitations of their research, including its generalization.

Mira reviews the research on patients perceptions about safety and involvement in improving safety. They describe two studies which are near completion. The first considers the relationship between patient satisfaction and the frequency of adverse events. The second is to developing a method for gathering patient's perceptions of hospitals' safety. It describes some of the practical and scientific challenges and solutions which the researchers faced. The results show patients often blame themselves for inadequate treatment reactions, or hold themselves accountable for not achieving certain results after a surgery.

5) Writing thesis and publication

This part of the book starts with advice for how to write-up your research, turn your data into a publication and get published. It notes that most researchers and research units have data and ideas which could be published but which lie unused because of lack of time and know-how about how and where to publish. It makes practical suggestions for how to publish and shows practical ideas in the different steps for drafting and submitting a paper.

The paper by Olsen is presented by the researcher at the stage of completed research and whilst writing their thesis. It is about a subject which is of practical and scientific significance: how to study safety culture in a meaningful way, specifically the validity of measures of safety culture using questionnaire surveys. The study of shows well how to deal with the challenges of choosing the right statistical methods to avoid Type I and Type II errors. It also shows how to contribute to cumulative knowledge in the field — just one example being,

> *"Forty-five percent of the sample did not report any (adverse) events. Contrary to Sorra and Nieva, we see no reason to believe that the lack of association with*

this outcome variable is due to a lack of variability or extreme skewness in the number of events reported. A more probable reason is that "Number of events reported" does not capture the actual risk level due to the poor culture of reporting in health care".

Kirsh presents the findings and reflections from research into interprofessional learning and patient participation for increasing self-care in diabetes management. One of the personal lessons which others have also reported is *"This project is of great interest and passion for me and I have learned how important this fact is, for my interest and passion have sustained me when confronting challenges"*. This and other papers also show how to study a subject of passion objectively: indeed the best way to serve the passion is to use scientific methods to reduce bias introduced by personal practical or emotional involvement.

6) Completed research and lessons for others

This section presents reflections and lessons for others from those looking back over their journey as well as some completed research papers.

Stains present a mixture of findings and reflections from his case study research into successful hospital and system quality programmes. The essay describes the steps in his journey and returns to the question addressed by Essen earlier: how a researcher formulates and focuses a question through a combination of personal interest, a review of research and a desire to help solve practical issues. It also gives an excellent description of the case study method he used and the details of the procedures to collect and analyse triangulated data.

Thor, in a searching and honest paper, reflects on his experience over many years of being a practitioner-researcher. He was deeply involved in leading many of the quality improvement changes which he was also studying, and the paper considers the insights he was able to follow up, the access to rich data, but also challenges in managing bias and ensuring objectivity and full documentation as if he was an outsider.

Does the collaborative breakthrough method work in developing countries and what do working personnel think about the method? The QIRN network is particularly rich in researchers studying large-scale programmes and in particular breakthough collaboratives and communities of practice. The paper by Unahalekhaka presents completed research from the first major study of a breakthrough collaborative outside of a developed western country. It describes a collaborative to prevent ventilator-associated pneumonia in 18 Hospitals in Thailand and shows that researchers and

practitioners from middle income (and from my own experience, lower income countries) can exceed the results of the west and produce research which is of higher quality.

Another paper from a developing country describes a study of the challenges and solutions in introducing an accreditation system in Pakistan. The study shows the value of a simple method — the strengths weaknesses opportunities and threats model can be used with selected group to generate useful data for science and practical purposes. This is just one of the papers from the developing world which we would have liked to include showing what the west can learn from these countries and their improvement efforts.

The last paper is about how ergonomics and occupational health can contribute to patient safety. It shows how the conditions surrounding health workers profoundly influence their performance. Improving the organization of healthcare can reduce errors and raise safety, but research is needed to show how we can improve the system and help health professionals to perform the highest standards, that can contribute to a correct risk management. The understanding of how the environment of care impacts the ability of providers to improve safety and how interactions with the physical healthcare environment (e.g. facility design, aesthetics, ergonomics conditions) influence the provision of safe high quality care, have been underestimated by researchers and policy makers. This article also allows us to understand and change conditions that can increase risks to patients.

Appendices

The appendices give headings for presentations of Quality Improvement Research for preparing for seminars or similar collections of papers. A second gives headings we use in the QIRN to explain your research to others and get advice in the earlier stages. A third appendix gives a checklist for assessing your research plan or completed research and a framework of 8 steps for research. Appendix 4 lists subject areas, which the QIRN proposes as priorities for future quality and safety research.

We think these papers will help and inspire researchers, especially those who have less support in their own research centres. Quality improvement research is still not strongly supported outside of the USA and the open and frank discussion of issues is not provided in other publications and is needed to help and inspire others. This is especially important because more managers, policy makers and clinical professionals are now more interested in using research evidence to improve their practical changes and programmes.

Strengthening the Scientific Value of Quality and Safety Improvement Research

JOHN ØVRETVEIT*

Introduction

Since the start of the QIRN there has been an increase in the number, variety and quality of research into the quality of health care and especially of studies of safety issues. Of particular interest is the increase in studies occupying the "middle ground" between, at the one extreme health practitioners' observational reports of their own projects and, at the other extreme, randomized controlled trials of interventions. This short chapter considers how to strengthen the scientific value of this type of research, which includes evaluations of quality and safety improvement programmes and policies over time. The next chapter considers how to strengthen the practical value of quality and safety improvement research.

The chapter will not review the range of types of study, or discuss areas for future research — a list of key areas for safety research generated in different meetings is provided in the appendix. Neither will it give an overview of methodological issues in improving each type of research, ranging from patient satisfaction to complex long term nor costly controlled experimental intervention studies. These are well covered in other papers and books noted at the end of this chapter. Rather, it will consider how to improve studies in this "middle ground", especially those which seek to:

— study quality interventions and implementation (e.g. guidelines or national quality programmes) in their natural context, without controls, and
— both provide knowledge of practical use and advance scientific theory about these interventions.

This chapter considers criteria for assessing the scientific value of a research study, (which leads to a checklist in the appendix for researchers to self-assess their

* Director of Research at Medical Management Centre, Karolinska Institutet, Stockholm, Sweden (jovret@aol.com).

research) and finishes by giving recommendations for improving research in this field.

Is my research scientific?

Nearly everyone says they "do research" these days, including journalists, patients, and consumers looking for a new car. What, then, makes research scientific? To strengthen the scientific value of your research, it is useful to consider the following study:

You carried out a project for your organization to find out how much and how effective quality improvement activities had been over the last 5 years. The aim was to get an overview and learn the lessons so as to help formulate a future programme. You collected various data about how many projects, the results of the projects and you interviewed 21 personnel about their involvement and views about the value of the activities. You did a report to management. Is there a publishable paper here, and how could you write it to get it published in a scientific journal.

Is this a valid scientific study? One way to answer is to consider theories of science, but a quick and practical way is "would it be published in a scientific peer reviewed journal"? Although there are other considerations apart from scientific value which decide publication, it is a useful "rule of thumb" and directs our attention to the criteria which reviewers use to assess scientific value and whether the study would meet these criteria.

Does the material meet basic scientific requirements?

A first step is to consider your data and/or ideas and ask whether it meets some basic requirements for scientific publication. These are:

- Are the methods for gathering data *valid* for investigating the study concept or question? (this also applies for doing a review of already-published research)?
- Are the data *reliably* and systematically gathered, following a defined method and would another researcher following the same method gather the same data and could they reproduce the study from the description?
- Are the methods for *analysis* valid for answering the study question and systematically and correctly applied?
- Do the *conclusions* follow from the findings and are all conclusions supported by data and or justified?
- Are the findings *generalisable* beyond the one site, case or situation?

- Does the study *relate the findings to previously published* research and show that it either fills an "important" gap in empirical knowledge, provides conflicting or supporting evidence, or tests a theory?

If the entire above are met, then the study makes a scientific contribution. But the contribution is only "significant" scientifically if the answer is "yes" to this question:

- Does the study provide an *explanation, understanding or prediction* which is valid?

These together are a tough set of requirements to meet. For some practice-based researchers, or scientific researcher who did not properly plan their study, the question at this point is "*could* the data be developed to meet these requirements?"

If the answer is, "more or less, but I am not sure *how well* it meets these requirements" then it is certainly worth going onto the next steps towards publishing, and using your doubts to inform your discussions of the limitations of the research (as discussed in a later chapter on how to publish your research).

If the collected data are not valid for the question, perhaps you can use the data for another question which is important. Sometimes data validity is poor because the sample was poor for some reason (biased or too small). One question is whether more data collection, without repeating the whole study, would strengthen the research and give data needed for a worthwhile publication?

As regards publication, if the answer is a clear "no", it does not rule out the paper for publication. Some journals are entirely or largely for and read by practitioners (clinicians or managers). These will publish papers which do not meet such high demands for methods rigour (e.g. a small sample, simple-before after study), or papers which do not give explanation or prediction.

The answer may be "no" because you have no or little empirical data but that you have an idea or a view about concepts. Most journals interpret their mission to publish "original research" in terms of the paper providing new and significant empirical findings (and meeting other criteria). Conceptual or ideas papers are always more difficult to publish, although some journals have special sections for "commentary papers" or "debate" papers. The answer is to find a journal which might publish what you have to say. Note the structure and style to these types of papers, and then use writing to clarify and focus your ideas and relate them to what has already been published (or show that there is little published and why this is an important topic, idea, or point).

Think about how the example study of quality activities in one organization meets these requirements and what would be needed to strengthen the research to make the study meet these requirements.

The purpose of the above was to show how to strengthen the scientific value of your research through asking if your data and analysis meet these criteria of scientific value. It also shows how important planning and designing the research is to ensuring these criteria are met — the single most common mistake is rushing to data gathering without working for many months planning and focusing the research, informed by a review of previous research.

This short chapter closes by summarizing these points and adding new ones to give recommendations and suggestions for research in what the paper referred to earlier as the "middle ground".

Seven recommendations for strengthening quality and safety improvement research:

— Funding and priorities for quality improvement research needs to include research into different quality intervention processes ("QIPs").
— There is a need to identify which QIPs most need evaluating in terms of whether the information from an evaluation can affect significant decisions being made by managers and quality specialists.
— More research of this type should be designed to answer quality practitioner's and manager's questions about which approaches are most effective, and about which "context conditions" are critical so as to allow transfer and replication or translation.
— Research should be encouraged which uses a variety of theory testing and theory building research methods to describe QIPs, study outcomes, understand how QIPs work to produce outcomes, and understand the assisting and hindering context factors.
— The quality of research into QIPs would be improved with better descriptions of the quality improvement process (i.e. the activities and what was actually done) and of the context of the process over time. Research needs to investigate the extent to which the programme was implemented: how "broadly" across all areas of the organisation and how "deeply" in each area. This should not rely only on retrospective reports by those responsible for the programme.
— Research should be encouraged where researchers collaborate with quality practitioner's to document projects which are significant but which would otherwise not be reported.
— An electronic internet data base of reports of quality improvement processes should be created to give practitioner's easy access to this evidence. This data base should include unpublished projects which were selected for presentation at quality conferences, and guidelines to help users to assess the different scientific status of the reports produced using different research methods.

Improving research into quality programmes

Research in this field could be improved by researchers paying attention to common failures of previous research:

- Implementation assessment failure: the study does not examine the extent to which the programme was actually carried out. Was the intervention implemented fully, in all areas and to the required "depth", and for how long?
- Pre-study theory failure: the study does not adequately review previous empirical or theoretical research to make explicit its theoretical framework, questions or hypotheses, not sufficiently distinguished between types of quality programmes.
- Outcome assessment failure: the study does not assess any outcomes or a sufficiently wide range of outcomes such as short and long term impact on the organisation, on patients and on resources consumed.
- Outcome attribution failure: the study does not establish whether the outcomes can unambiguously be attributed to the intervention, or whether something else caused the outcomes
- Explanation failure: there is no theory or model which explains how the intervention caused the outcomes and which factors and conditions were critical.
- Measurement variability: different researchers use very different data to describe or measure the quality programme process, structure and outcome. As a result it is difficult to use the results of one study to question or support another or to build up knowledge systematically.

Recommended bibliography

Berwick, D., 1996. Harvesting knowledge from improvement, *JAMA*, 275 (11), p. 877-878.

Berwick, D., 1998. Developing and testing changes in delivery of care. *Annals of Internal Medicine,* 128 (8), p. 651-656.

Berwick, D., 2005. Broadening the view of evidence-based medicine *Quality & Safety in Health Care,* 14, p. 315-316.

Davidoff, F. & Batalden, P., 2005. Toward stronger evidence on quality improvement: draft publication guidelines: the beginning of a consensus project. *Quality & Safety in Health Care,* 14, p. 319-325.

Grol, R., Baker, R. & Moss, F., ed., 2003. *Quality improvement research: understanding the science of change in health care.* London: BMJ Books.

Jordan, H., 1995. How rigorous need a quality improvement study be? *Joint Commission Journal of Quality Improvement,* 21, p. 683-691.

Nolan, K. M. & Nolan, T., 2003. *Learning from quality improvement in healthcare systems: the challenges of pragmatic science.* In Max, M. B. & Lynne, J., ed., 2003. *Clinical symptom research: challenges and opportunities.* Chapter 13 [internet]. Bethesda: National Institutes of Health. Interactive Textbook. (published 2003) Available at: http://symptomresearch.nih.gov/chapter_13/sec2/ckns2pg1.htm [accessed 20 October 2007].

Øvretveit, J. & Gustafson, D., 2002. Evaluation of quality improvement programmes. *British Medical Journal,* 11 (3), p. 270-275.

Øvretveit, J., 2002. *Action evaluation of health programmes and change.* Oxford: Radcliffe Medical Press.

Øvretveit, J., 1998. *Evaluating health interventions.* Buckingham: Open University Press.

Strengthening the Practical Value of Quality and Safety Improvement Research

JOHN ØVRETVEIT*

SUMMARY

This chapter considers ways to:

— strengthen the practical value of quality and safety improvement research;
— enable practical actors or decision makers to make more effective actions by combining research with other evidence.

It does so by considering which evidence actors need to motivate them to change what they do, and how much of this can be provided by quality improvement research.

This chapter concentrates on evaluation research into the effectiveness of quality and safety interventions, but many of the points apply to other types of quality and safety research. There are two proposals: first that decisions makers should not wait for evidence from controlled trials to decide whether to introduce a large-scale or complex quality programme or change. They do need information from research, but should not delay making changes for interventions for which there is some adequate evidence. The key issues are which evidence is sufficient to justify acting or not acting and which evidence do they need about implementation, costs and possible harm.

The second proposal is that researchers give decision makers with information about the cost of implementation and the risks of harm or waste, rather than only evidence of effectiveness. It proposes that the burden of proof required should be proportionate to the likely benefits and costs in terms of possible harm and cost and ease of implementation. It describes a decision process to

* Director of Research at Medical Management Centre, Karolinska Institutet, Stockholm, Sweden (jovret@aol.com).

assess these aspects where evidence is absent or weak. It concludes that decisions makers will always need to assess the evidence for their situation, but that researchers can improve the information they provide for practical action.

Introduction

This paper considers quality improvement research from the perspective of health service practitioner and policy user's perspective — termed "actors" in this paper. Quality improvement research (QIR) has a practical as well as a scientific aim — an earlier chapter considered how to strengthen the scientific value. For research to contribute more to quality improvement, researchers need to be aware of issues relevant to how their research can be enacted by leaders and managers working in health services. A key issue for these actors is, how strong is the evidence from one study and other research? Is there evidence enough for us to change what we do?

This paper considers the evidence which actors need to decide whether to change their service or policies, what to change and how. It does not assume decision making and action is wholly rational and evidence based. There are many factors which help and hinder evidence being used in decision making which will not be considered here. The paper does assume that some actors wish to use research to make more informed, and therefore more effective decisions, and it concentrates on which information they need from a quality improvement study.

If researchers are more aware of actors needs and design their research with this in mind, as well as the scientific questions, then their research is more likely to be used to make practical improvements. Researchers cannot blame actors for not using their research if it does not provide the information actors need to decide whether to change and what, and how, to change. There are also issues in how researchers or others present the information — there is an analogy here to the information which patients need to make decisions about treatments and the form in which they need it.

1. Actors needs from quality improvement research

This paper concentrates on quality improvement research into "interventions" or changes which are intended to improve care for patients. Interventions to improve quality are led by actors at different levels of the health system, to change "targets" at different levels of the health system (see next page).

There are limitations to the language and concepts of natural science and some medical research when applied to social change, and especially large scale social

LEAD ACTOR "IMPLEMENTING" THE "INTERVENTION"	"TARGET" OF CHANGE
Project team or clinical team	Individual worker behaviour or thinking, organisational procedures.
Leader or improvement group for an organisational unit (department, primary health care centre, nursing home, whole hospital).	All or one section of the organisation, organisational structure or procedures
Leader or improvement group for a health system (regional or national)	All or some health facilities or organisations in the area

programmes. The concepts of intervention and targets conveys a mechanistic action, through which "actors" do things to passive "targets" according to a planned timescale which is implemented in step. Many quality improvement changes are better conceptualised as an interaction between leaders, facilitators and clinicians and workers, using methods and negotiating change to everyday practice. All are potential actors, some more or less active or willing than others: rather than causal mechanisms, an understanding of conscious actors interpreting ideas, political processes, and conditions which help and hinder certain changes may be more relevant. The paper later considers more sophisticated concepts for describing quality improvement "interventions" to social systems and better terms such as "pathways of influences", but will follow for the moment the conventional term "intervention".

In summary, QIR studies many different types of intervention at different levels. Actors at one level create the conditions for, or direct action for quality improvement by actors in the level below. Actors at each level need different types of information about the interventions which they are interested in, or which they could implement within their area of responsibility.

2. Evidence to act

It is well documented that managers and policy makers tend not to use research in decision making and action. There are many explanations for this, but some are due to the limitations of research which does not give the information which decision makers need. This paper considers how much of this information can be provided by different research approaches, and how research can be modified to be more useful. It also considers is how much researchers can and should provide information which is more useful, or whether other ways of assisting decision making are needed.

Criticisms which actors have about much of quality improvement research are similar to concerns which clinical practitioners and patients have about evidence from

medical intervention research. This is that much of the research is inconclusive, can only be certain in relation to the sample studied which may not be representative. Actors are also unclear about at what point the strength of evidence justifies changing the traditional treatment or status quo. When is the evidence strong enough to action on?

The more we consider the user of the research and the situation in which decisions are made, the more we see how many considerations there are in their decision, and how research only provides limited information. In part the answers are decided by where the research stands on the spectrum of "purely scientific" vs. "purely practical": the latter is designed more with the practical users needs in mind, rather than only to test or build previous knowledge. In part the answer may be decided by the level of the decision maker. It is possible that the higher the level and the broader the scope of the programme, the less research will be relevant and the more political and other influences will dominate the decision. Whether this is due to the research difficulty of linking results to interventions, or is always the case, and how much this should or can change is open to debate, but governments in both the UK and Canada are emphasising evidence based policy making.

3. What is evidence?

The natural science paradigm applied in medical intervention research aims to create objective evidence of the effectiveness of the intervention. The intervention, patients, and situation is carefully controlled through a randomised controlled trial (RCTs) design so that other explanations for the outcomes can be excluded. This type of research is wrongly described as generalisable, but it is only generalisable to patients similar to those in the trial. Researchers readily accept that practitioners and patients need to assess the extent to which the trail outcomes are likely in the individual case, in terms of both patient and situational characteristics.

This type of research is viewed by some working in health care as the strongest type of evidence for decision making: evidence of effectiveness from a RCT of a controlled intervention on selected patients, or, even better, evidence from a systematic review of a number of different RCTs. It is not known how many working in health care view evidence only as evidence of effectiveness from a randomised controlled trial. Some do not view other findings from research as strong or useful evidence for decision-making. What is clear than opponents of quality interventions cite the lack of this "more certain" knowledge as a reason for not making changes. However, it needs to be noted that evidence based medicine is the "use of the best available evidence", and also there are broader conceptions of evidence, for example

"anything that establishes a fact or gives reasons for believing something" (Oxford American Dictionary).

When considering how to make QIR more useful to decision makers we need not only to consider the needs of different decision-makers at different levels and the different types of interventions, but also the conception of evidence held by them and researchers and the type of evidence which different types of research are capable of providing.

4. Research, guidance and decision-making for action

There are a number of situations improvement leaders could use research to make better decisions and actions which would lead improved outcomes for patients. This section of the paper considers three situations.

4.1. Should we carry out this improvement change

The first is deciding whether to carry out a particular change which has been recommended by higher levels. Research carried out and reported elsewhere can help decide whether to make the change, and local quality improvement research in the organisation can pilot the change and add useful information for this decision. Research which only considers effectiveness of the intervention for improving care is, however, limited for leaders to decide to carry out a particular change. Other information is needed about costs of the intervention and ease of implementation which also includes factors which help and hinder implementation. Also information is needed about harm or negative outcomes which could result for patients, personnel or the organisation (sometimes terms information for risk assessment). This is thus a combination of effectiveness-cost-ease and harm information (the ECEH) for a particular quality intervention. It describes different types of evidence which researchers who wish to make their research more useful would need to gather. A later section below considers research and other approaches which could provide these types of evidence.

4.2. Which improvement should we do?

A second situation is where decision makers need to compare possible interventions for their organisation to decide which is most likely to result in safer or higher quality care. For a particular problem, such as hospital acquired infection (HAI), which of

the different possible actions should be used? Further, whilst some interventions are mandated, improvement leaders still have a wide choice of different alternatives if they wish to improve quality. Which of the many problem-intervention combinations should they address: HAI, or whether to introduce a reporting and root cause analysis system, or a shift hand-over communications improvement etc? A hospital management team considering, for example the 30 recommended NQF interventions would need to prioritise which to work on first. Is there research which helps them make better informed decisions about which to implement?

Ideally research would provide them with, for each intervention, the ECEH information as an index which allows them to compare what would be best in their local situation. However, this is unrealistic. On the other hand, using only effectiveness information which comes from research elsewhere may actually lead to poorer decisions as it would not consider cost, ease and harm in their local situation.

One way forward for choosing which intervention is best for the local situation is to carry out a structured decision-process which combines different forms of evidence to calculate the best intervention. Bayesian decision conferencing techniques are useful for this purpose as they make it possible to combine effectiveness research evidence with participants' estimates about local cost, ease and possible harm for each alternative. The method then ECEH assessments for each interventions and compares them to show the best alternative. It then allows sensitivity analysis to assess whether the final result would change with extra evidence about one or other of the interventions (Phillips & Phillips, 1993).

4.3. How should we carry out the improvement?

A third situation is planning and carrying out the intervention, after deciding which intervention to use. The research which is helpful here to make the implementation effective is research on what helped and hindered implementation elsewhere, and local action evaluation or research about the process of implementation which is reported to decision makers in time for them to make adjustments.

5. To what extent do different research paradigms and models provide for these types of evidence

The paper proposed that research can be more useful for improvement leaders if it provides more information than on effectiveness alone. The following considers the extend to which different paradigms and types of research are able to provide this

information, and whether other "bridging methods" between the research and improvement leaders — such as a structured decision-process — are needed to enable them to make more use of the research.

Research into quality improvement interventions can be considered as three broad types: research which applies the RCT design or a quasi-experimental comparative design, research which uses a pragmatic testing before-after, or time series design, and social research models and designs.

The first type is experimental research using a natural science model. This can give evidence of effectiveness for "dose-type" quality interventions which can be standardised and controlled and where other controls are possible for the targets and context to rule-out confounders. The limitations of this research for improvement leaders were described above. This research takes time (typically 3 years) and many resources to do one RCT, yet and many RCTs are needed. For many quality and safety interventions RCTs are difficult or not feasible. One example is research into rapid response teams (RRTs): it has taken 10 years to generate about 2-3 good RCTs of this intervention and most are from academic medical centres.

The more we move from "treatment type" safety interventions (e.g. simple single controllable dose-type) towards multiple interventions ("bundles", where synergy may be important), and the higher the level and longer the timescale (e.g. a RRT, or a safety programme in a hospital), the more difficult it is to use RCT designs and control or compare the intervention. With these larger — and longer — scale interventions there are many outcomes. Many short and long term outcomes need to be valued from many stakeholder perspectives. The boundary between intervention and context is arbitrary and, wherever the boundary is drawn, there are context factors which interact with the intervention (helpers and hinders) as it evolves over time in a changing situation. The concept of intervention and confounders begins to be less useful, rather "actions-taken-to-implement" and assisting and hindering conditions become more relevant ideas. In summary, the concepts of evidence which assume an RCT model and a simple patient treatment need to be developed when we consider quality and safety interventions, especially those which are complex social interventions in complex changing social systems.

Before/after pragmatic quality improvement research is a second type of research which can be carried out more quickly and at a lower cost than RCTs or comparative experimental designs. This research can be carried out for some interventions on a pilot basis to give some evidence of effectiveness in local situations. However the results are less certain because of lack of controls, and often there is no study of costs, or of what helps and hinders implementation.

The third broad type is social research which studies the intervention in context without controls. This research gathers different stakeholders' perceptions of results

and also describes the evolving intervention and what helps and hinders implementation and development. This research is more like innovation research than intervention research as it studies how the intervention is adapted and develops in the local situation.

6. Conclusion

Research can never provide full information which improvement leaders need to act. However, more and different research can improve decision-making and action. Further, methods "bridging" research and practice which allow different types of evidence to be combined can also significantly improve decisions, and are especially needed where research is limited and action is required quickly.

"When do we know enough to act?" One answer is, when enough of the people who need to act are persuaded by different forms of information a) that the idea can be implemented in their setting, b) it is probable that certain beneficial results will occur c) the costs are bearable (and may produce X savings in year, d) and the harm is low. In practice, all these cannot be accurately assessed from research and a full rational comparison based on sound evidence cannot be made. Rather, the paper proposed:

a) that if leaders are going to make better actions they need to consider cost, ease of implementation, and harm, not just effectiveness;
b) we will never have complete evidence but we can make better use of what we do have;
c) many types of evidence can be used to assess these four aspects, including experiential knowledge and opinion (using Bayesian decision-analysis tools), and leaders need to make judgments about these four aspects for their local situation, not in general;
d) leaders can never be certain, and will always need to act with a large amount of faith and hope — to expect otherwise is unrealistic and uncourageous;
e) decisions should never be fully rational anyway — values and a non-rational preferences do and should be part of the assessment.

There is an ethical case for implementing interventions with low cost, which are low risk of harm, and easy to implement, and to do so sooner rather than later. In these cases the burden of proof should shift from advocates to "those against" — "those against" need to prove the harm of acting rather than advocates having to present evidence which is not possible for many safety interventions. At present the defendant is considered guilty before the trial, and the jury is rigged.

There are needless deaths, suffering and costs to inaction while decision-makers become "convinced". A good example is the WHO oral replacement therapy programme started in 1981 after "weak evidence" showed that the treatment saved lives and no alternative home based treatment was possible. RCTs were later done. Randomised controlled trials which showed that ORT was as effective as intravenous therapy. But the "early" use large programmes before RCTs led to three million fewer deaths from diarrhea annually (Victoria et al., 2000). There are many other examples from HIV/AIDS prevention and QI in developing countries.

7. References

NQF, 2003. *Safe practices for better healthcare* [internet]. Washington, D.C.: National Quality Forum. Available at http://www.qualityforum.org/publications/reports/safe_practices.asp [accessed 30 May 2005].

Phillips, L. & Phillips, M. C., 1993. Facilitated work groups: theory and practice. *Journal of Operational Research Society,* 44 (3), p. 533-549.

Victoria, C. G., Bryce, J., Fontaine, O. & Monash, R., 2000. Reducing deaths from diarrhea through oral rehydration therapy. *Bulletin of the World Health Organization,* 78 (10), p. 1246-1255.

Patient Safety Research: a Challenge for Public Health

Paulo Sousa*
Cláudia Furtado*
Vasco Reis*

SUMMARY

All healthcare services around the world occasionally and unintentionally harm patients whom they are seeking to help. In the last few years, several studies have estimated that around 4% to 17% of patients have experienced an adverse event, and that up to half of these incidents could have been prevented.

These incidents have several implications in different clinical areas and levels of care, and represent nowadays a public health problem. In recognition of this, patient safety has become a core issue in many modern healthcare systems and a fundamental part of improvement quality projects.

Despite a growing knowledge patient safety research lacks a systematic approach worldwide. Therefore it is crucial to define a patient safety strategy, establish priorities, capacitate researchers and engage all stakeholders, with the final aim of reducing the possibilities of harming patients.

It should be recognised that healthcare will always involve risk. However, analysing and tackling the root causes of incidents can reduce these risks in future. These could contribute for a health care delivery of excellence and based on the best evidence.

It is our belief that the patient safety issues will, in the near future, be one of the main areas for action, reflection and research in the public health domain.

* National School of Public Health, New University of Lisbon, Portugal (paulo.sousa@ensp.unl.pt).

Introduction

"Medicine used to be simple, ineffective and relatively safe. Now it is complex, effective but potentially dangerous"

Sir CYRIL CHANTLER

The purpose of this paper is to outline the importance and to summarise evidence about the extent and causes of problems on patient safety and quality in healthcare, as well as to describe international research and practices that were designed to identify and resolve these problems.

The quality of healthcare has traditionally been judged against the principles of safety, effectiveness, patient-centeredness, timeliness, efficiency and equity (IOM, 2001).

Health care quality can be viewed by different perspective, and thereby assume different definitions. Joint Commission on Accreditation of Healthcare Organizations (JCAHO) defines health care quality as the "the degree to which health services, with the current professional knowledge, increase the likelihood of desired health outcomes and decrease unwanted outcomes" (Batalden & Stoltz, 1993).

Patient safety is currently recognized as an extremely important component of health care quality. Nowadays there is a substantial body of evidence about the consequences of patient safety, or its absence, in health organizations, in its staff and especially in its patients/users. The absence of patient safety can result in: a) lost of confidence in health organizations and its professionals, deteriorating professional and patient's relationship; b) increase of economic and social costs; and c) reduction in the possibility of achieving expected outcomes (WHO, 2002; Arah & Klazinga, 2004; Wears, 2004). These consequences will also have a direct impact on the quality of health care delivered.

For the reasons mentioned above patient safety is highly ranked in the health policy agenda worldwide, particularly in Europe, North America and Australia. It has also been a core issue in the strategy of some health international organizations, as the World Alliance for Patient Safety, launched in 2004 by the World Health Organization, and the High Level Group of Health Services and Medical Care created by the Council of Europe. The last one, which was set in 2004, consists in a group of programs and incident reports systems with the participation of countries like Sweden, Netherlands, Ireland, United Kingdom and the Check Republic.

In 2005, took place the European Union Luxembourg Presidency Conference on Patient Safety, and the final meeting of HOPE Exchange Programme in Cardiff, sponsored by HOPE — European Hospitals and Healthcare Federation, with the title "Patient Safety: Learning, Sharing, Improving", among other examples.

More recently, in September 2007, the University College of London (Faculty of Public Health), the World Health Organization World Alliance for Patient Safety (www.who.int/patientsafety/en), the European Comission (www.ec.europa.eu), with the support of the Portuguese European Union Presidency, have organized the conference "Patient Safety Research: shaping the European agenda". This conference joined up researchers, policy makers, funding agencies and other stakeholders.

Patient safety is not, as some could expect, a straightforward concept, and consequently its absence is not easily identifiable, nor are its solutions. Indeed, patient safety approach is not simple mainly due to health organization's complexity, the sensibility of this subject and also to its multifactorial character.

Health organization's complexity

It is common to consider health organizations as systems of huge complexity essentially due to: a) its mission, in most countries moved by equity, accessibility, universal coverage principles, etc.; b) its activities (imperfect market, patient uncertainty); c) its user's characteristics (in general with low capacity to make informed decisions; with high expectations; physical and emotional vulnerable, etc.); and d) its staff, multidisciplinary in nature, with different professional lines, usually work loaded and urged to be updated with the scientific and technological evolution.

Sensitivity of the subject

Patient's safety is one of the most sensitive themes to approach in health care delivery. Despite the evident benefits to patients resulting from the modern system of health care delivery, there are obviously inevitable risks associated that may cause adverse events (Kohn, Corrigan & Donaldson, 2000; Berwick, 2004). Nevertheless patient's expectations are always high and health care incidents are difficult to accept.

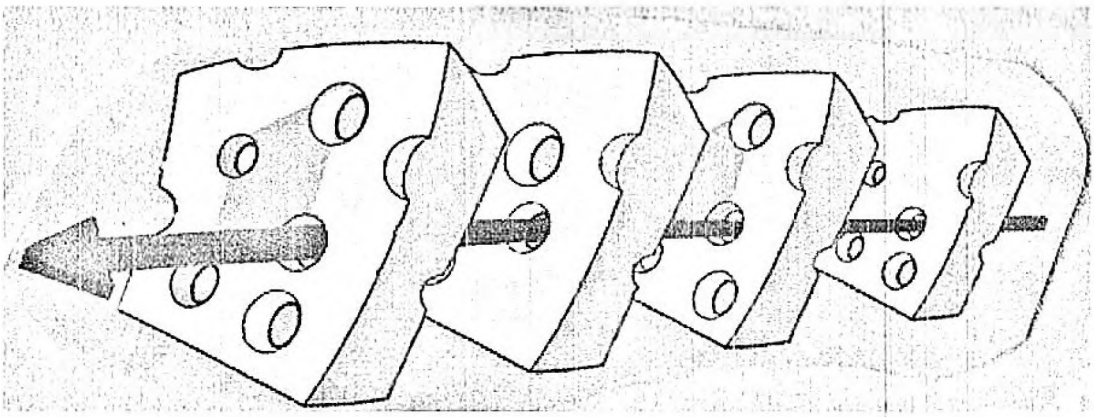
Clinical and technological advances, increase in life expectancy, widespread access to information, a change to a more responsible and demanding attitude — patient empowerment — have contributed to raise patient's expectations.

Multifactorial character underlying patient safety

It is well recognised that patient's lack of safety result from various combinations of individual, team, organisational and patient factors. James Reason (2002) has

explained this causal multiplicity with the "Swiss cheese" theory, in which holes are considered as gaps. In accordance to this theory for an adverse event or any patient's injury to occur it is necessary that the holes of different cheese layers get in line, as illustrated in Figure 1.

Figure 1
Image "Swiss cheese"



Some authors also explain this multifactorial character with the Donabedian's 'triad', based on structure, process and outcome (Donabedian, 1984; Department of Health, 2001; Reason, 2002). Structure refers to hospital buildings, staff, equipment, and resources. Process describes how structure is put into practice, such as specific therapies. Outcome refers to results of processes, for instance, results of a specific therapeutic intervention or patient satisfaction.

Due to these multiple characteristics, patient safety is frequently analysed in different perspectives, among which are the accreditation processes, risk management, clinical governance, medical error, adverse events, sentinel events, near miss, etc. But whatever approach is made it will be always indissociable from a health care quality perspective.

Why we should be concerned with patient safety?

Patient's safety has become in the last decade a core subject in health policy's agenda in different countries worldwide.

Although since 1950 some punctual work had be done, it was only in the beginning of the 1990s that this subject had a stronger development, with the publication of a Harvard Medical School study demanded by the Commissioner of Health of New York State, which showed that 3.7% patients were injured during inpatient stay (Brennan et al., 1991).

In the mid-1990s was published "To Err is Human" by the Institute of Medicine, which concluded that 48,000 to 98,000 Americans died in hospitals every year because of preventable medical errors (Kohn, Corrigan & Donaldson, 2000).

These reports raised awareness about adverse events issues of both academics and policy makers and prompted the realization of more in depth studies, which provided a more detailed picture of this issue.

In the USA, the publication of another important document "Crossing the Quality Chasm" (IOM, 2001), prompted new actions, including laws' approvals in 15 American states that turned mandatory the adverse event report.

In Denmark, the Danish Adverse Event study was published in 2001, based on a review of 1097 patient's records (Schioler et al., 2000). The study found that 9% of patients admitted to a Danish hospital were exposed to an adverse event and 40% of the adverse events were preventable. The adverse events resulted in an average of 7 days prolonged hospital stay, with inherent economic and social costs.

In 2001, a study by Vicent et al. (2001), estimated that around 10% of patients (900,000 using admission rates for 2002/3) admitted to English NHS hospitals had experienced a patient safety incident, and that more than 50% of these incidents could have been prevented. This study also estimated that 72 000 of these incidents may have contributed to patient's death.

These results are also in accordance with the main findings of Australia and New Zealand studies in the same period (Wilson et al., 1995; Ruciman et al., 2002).

In Sweden, this problem has also had an increasing attention from different local, regional and national authorities. The National Board of Health and Welfare (NBHW) is responsible of the systematization, analysis and definition of corrective or preventing strategies. The NBHW is also responsible for the publication of adverse event's reports, named Lex Maria, which are divulgated to the respective regions. With this approach it is expected to learn with the errors, and improve and correct the underlying causes (Øvretveit, 2003; NBHW, 2001).

There is also another structure in Sweden, the *Medical Responsibility Board,* which based on patients or its familiar's complaints can initiate a disciplinary process against the health professional. Patients can also have access to a patient insurance, which is activated when an adverse event or error, with damage consequences to the patient, occurs.

In addition to the consequences on patient's health status, adverse events also generate a significant financial and social burden, which should not be neglected. Studies in United Kingdom (UK) estimated that every year due to adverse events 2000 million pounds were associated with additional inpatient stay, and about 400 million pounds were paid in litigation processes (Vicent, 2000; NDHSA, 2005).

In USA, studies estimated that preventable adverse events costed between 17,000 and 29,000 million dollars, including medical expenses and indirect costs (Khon; Corrigan & Donaldson, 2000; Leape, 2002).

In addition to these costs, it should also be taken in consideration indirect costs caused by the lost of confidence and satisfaction with health care organizations.

Most of the studies (Leape et al., 1991; Wilson et al., 1995; Department of Health, 2000; Thomas & Brennan, 2000; Vicent, 2000; Leape, Berwick & Bates, 2002; Reason, 2004) appointed to an adverse event occurrence rate between 4% and 17% in total admissions, which *per si*, represents a serious public health problem. Governments recognized this problem and defined strategies to characterize the problem and identify ways of reducing or solving the incidence of these occurrences.

In fact there have been major safety developments in the last few years and these have occurred in a wide range of government and private healthcare organisations. Some of these key activities are listed below.

One of the measures adopted in many countries was the creation of specialized agencies, which were due to analyse the problem and propose measures to invert the situation. As examples we can mention the National Patient Safety Agency in United Kingdom, the Danish Society for Patient Safety and the Australian Patient Safety Foundation.

Another measure adopted in countries, like United Kingdom, Sweden, Norway, Czech Republic and the Netherlands, was the implementation of a national system of voluntary adverse event report (Scioler et al., 2000; Altman, Clancy & Blendon, 2004; NPSA, 2004; Lewis & Flecher, 2005).

In addition to the creation of specialized agencies and national systems of voluntary adverse event report, many governments made efforts to promote a learning culture in spite of a blaming culture. The emphasis is now on a root cause analysis and on the reinforcement of strong leadership and an organization's ability to listen to all members of the healthcare team.

The National Patient Safety Agency of UK has recognized the vital importance of involvement of stakeholders, and developed a strategic document named, *"Seven steps to patient safety a guide for NHS staff"*, which defined 7 steps that all NHS organizations should take to promote patient's safety and therefore improve health care quality (NPSA, 2004).

Each step provides a checklist that helps planning the activities and measure performance and effectiveness of all actions taken to promote patient's safety.

Health organizations to adopt and follow these steps need a strong leadership and team building as well as a high level of commitment among professionals and services/departments that are part of the health care network.

Table 1
Seven steps to promote patient safety

Step 1	*Build a safety culture,* create a culture that is open, fair, with no blame nor punishment
Step 2	*Lead and support your staff,* establish a clear and strong focus on patient safety throughout your organization.
Step 3	*Integrate your risk management activity,* develop systems and processes to manage your risks and identify and assess things that could go wrong.
Step 4	*Promote reporting,* ensure your staff can easily report incidents locally and nationally.
Step 5	*Involve and communicate with patients and the public,* develop ways to communicate openly with and listen to patients and their families.
Step 6	*Learn and share safety lessons,* encourage staff to use root cause analysis and learn how and why incidents happen.
Step 7	*Implement soluctions to prevent harm,* embed lessons through changes to practices, processes or systems.

Following these 7 steps does not prevent adverse events to occur, but it is assured that health care delivered is as safe as possible, and if an event not expected occurs the corrective measures will be taken in a short period of time and based in the best available evidence (NPSA, 2004).

Patient safety — a focus on research

Despite the global commitment to improve patient safety there is a need not only to develop research but also to describe how research can be applied into policies and/ or practices.

In fact, patient safety research is still at the beginning worldwide, but in most countries the research is still fragmented and undervalued. The reasons underlying are i) data for health research is inadequate, ii) paucity of researchers in the field, iii) allocation of research funding is still limited, and iv) lack of directions in research programmes (Lilford, 2002; WHO, 2002; WHO, 2006). Even in countries with large support research tend to be opportunistic rather than strategic or based on priority needs.

It is therefore essential to raise awareness and mobilize resources towards patient safety research, which can only be strengthened through a workforce capable of both

planning and undertaking research, and also of implementing results and doing follow up actions.

It has been recognized in patient safety discussions that to achieve this aim it is essential to create multidisciplinary teams and develop networks around patient safety research. (Meyers & Eisenberg, 2002; Lilford, 2002; Reinertsen, 2006).

Because their multidisciplinary nature, and their areas and methodologies of research, public health teams are in good position to leadership this approach. Public health researchers and schools can start the process of creating capacity for research, through educational courses directed to areas focusing patient safety and quality, and also of fostering networks in this subject (WHO, 2006; Øvretveit & Klazinga, 2007). Sharing knowledge can be an opportunity to support individual or groups of researchers and therefore improve communication and information flow among all researchers engaged in developing patient safety. It is also important to include patient's point of view when we draw a patient safety strategy.

However it must be highlighted that the development of patient safety research needs a clear political willingness in order to allocate research funding to patient safety (we recommend the reading of the EU strategic prorities for research for 2007-2013 — The seventh framework programme; and The Decision N.º 1350/2007/EC adopted jointly by the European Parliament and The Council)[1]. Therefore it must be evidenced, to governments and other stakeholders, that the development of research and strategies focused on patient safety is cost effective.

In fact, there are well identified areas, where the lack of safety compromises the quality of health care with huge damages to patients, health services and its staff, and consequently to governments (WHO, 2006; Øvretveit & Klazinga, 2007), with additional costs. It is also known that drug therapy; surgical procedures; falls; pressure ulcers and hospital-acquired infections account for most of the avoidable harm and are, for the reasons mentioned above, priority areas for research. These areas provide an opportunity to engage healthcare professionals and to assure funding for patient safety activities and for quality improvement as a whole (www.patientsafetyresearch.org).

[1] The Second Programme of Community Action in the field of health (2008-2013) came into force on 1st January 2008. One of the objectives to be pursued through the action is "To improve citizens' health security". This programme is intended to complement, support and add value to the policies of the Member States and contribute to increase solidarity and prosperity in the European Union by protecting and promoting human health and safety and by improving Public Health. The financial envelope for the programme is 321millions euros. The programme will be implemented by means of annual work plans which set out priority areas and the funding criteria.

Conclusions

Safety is in the core of patient care and is a decisive component of quality management. Improving patient safety has become a core issue in many modern healthcare systems and the consequences of its absence are, nowadays, an important public health issue on health policy's agenda worldwide.

Despite a growing knowledge both on patient safety and on strategies to improve it, which were presented in this chapter, there are still considerable commonalities in the challenges faced by most countries on patient safety approach.

For this reason, it is important to define a patient safety strategy, based on a clear identification and hierarchization of interventions and research areas, which can constitute a global framework.

To develop this field it is essential to capacitate researchers, create national and internationally networks of knowledge and experiences and obtain funding to support research and the implementation of research findings. It is important not only to develop research but to describe how research can be applied into policies and/or practices.

It should also be made an effort in order to include patient safety research in educational institutions curricula and develop higher educational degrees in this area, such as doctorates and masters.

To accomplish its mission a patient safety strategy must operate at multiple levels of the healthcare system, should be defined and implemented by multidisciplinary teams, and patients might and should have an important role in its definition. At the same time, it's crucial in patient safety strategy to select indicators that are able to monitor and evaluate the efficiency and effectiveness of the interventions that are the result of its implementation.

To conclude patient safety research, its implementation and evaluation, are essential to support health policies and effectively contribute to health system strengthening by providing evidence on corrective and preventive interventions, with the final aim to improve safety and quality of care.

It should be recognised that healthcare will always involve risk and a patient safety strategy will not prevent adverse events to occur. However, analysing and tackling the root causes of incidents will reduce these risks in future.

References

Altman, D. E., Clancy, C. & Blendon, R. J., 2004. Improving patient safety: five years after the IOM report. *New England Journal of Medicine.* 351 (20), p. 2041-2043.

Arah, O. A. & Klazinga, N. S., 2004. How safe is the safety paradigm? *Quality & Safety in Health Care.* 13 (3), p. 226-232.

Batalden, P. B. & Stoltz, P. K., 1993. A framework for the continnual improvement of health care. *Joint Commission Journal of Quality Improvement.* 19, p. 424-447.

Berwick, D., 2004. Lessons from developing nations on improving health care. *BMJ,* 328, p. 1124-1129.

Brennan, T.A., et al., 1991. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. *New England Journal of Medicine*, 324, p.370-377.

CORDIS, 2007. Seventh Research Framework Programme (FP7). Brussels: European Commission. [online] Available at: http://cordis.europa.eu/fp7/about_en.html [accessed 21 January 2007].

Department of Health, 2000. *An organization with memory: report of an expert group on learning from adverse events in the NHS.* London: Department of Health.

Department of Health, 2001. Building a safer NHS for patients. London: Department of Health.

Donabedian, A., 1984. *La calidad de la atención médica.* México, DF: La Prensa Mexicana.

Donaldson, L., 2004. Patient safety global momentum builds. *Quality & Safety in Health Care*, 13 (2), p. 80-86.

Global Forum for Health Research, 2006. *Monitoring financial flows for health research 2005.* Geneva: Global Forum for Health Research.

Kohn, L. T., Corrigan, J. M. & Donaldson, M. S., ed., 2000. *To err is human: building a safer health system.* Washington: National Academy Press.

Institute of Medicine, 2001. *Crossing the quality chasm: a new health system for the 21st century.* Washington, D.C.: National Academy Press.

Leap, L. L., et al., 1991. The nature of adverse events in hospitalized patients: results of the Harvard medical practice study II. *New England Journal of Medicine*, 324, p. 377-384.

Leape, L. L., 2002. Reporting of adverse events. *New England Journal of Medicine,* 347 (20), p. 1633-1638.

Leape, L. L., Berwick, D. M. & Bates, D. W., 2002. What practices will most improve safety? Evidence-based medicine meets patient safety. *JAMA*, 288 (4), p. 501-507.

Lewis, R. Q. & Fletcher, M., 2005. Implementing a national strategy for patient safety: lessons from the National Health Service in England. *Quality & Safety in Health Care*, 14, p. 135-139.

Lilford, R. J., 2002. Patient safety research: does it have legs? *Quality & Safety in Health Care*, 11, p. 113-115.

Meyers, G. S. & Eisenberg, J. M., 2002. The end of the beginning: the strategic approach to patient safety research. *Quality & Safety in Health Care*, 11, p. 2-4.

NBHW, 2001. *Overall quality indicators in health care and medical services.* Stockholm: Socialstyrelsen.

NDHSA, 2005. National strategy for quality improvement in health and social affairs 2005-2015. London: National Directorate for Health and Social Affairs.

National Health Care Quality Registries in Sweden, 1999. National Health Care Quality Registries in Sweden. Stockholm: Landsting Forbundet and Socialstyrelsen.

NPSA, 2004. Seven steps to patient safety: a guide for NHS staff. London: National Patient Safety Agency.

Øvretveit, J., 2003. *What are the best strategies for ensuring quality in hospitals?: WHO Regional Office for Europe's Health Evidence Network (HEN) Report.* Geneva: World Health Organization.

Øvretveit, J. & Klazinga, N., 2007. *Developing a health system quality and safety strategy: guidance for WHO Europe member states.* Copenhagen: World Health Organization.

Reason, J., 2002. *Human error.* Cambridge: Cambridge University Press.

Reason, J., 2001. Diagnosing vulnerable system·syndrome: an essential prerequesite to effective risk management. *Quality & Safety in Health Care*, 10, p. 21-25.

Reason, J., 2004. Beyond the organizational accident: the need for "error wisdom" on the frontline. *Quality & Safety in Health Care*, 13 (Suppl. II), p. 27-32.

Reinertsen, J. L., 2006. *Quality is now strategic.* In Society for Healthcare Strategy and Market Development, 2006. *Futurescan: healthcare trends and implications, 2006–2011.* Chicago: Health Administration Press.

Ruciman, R. B., et al., 1998. A classification for incidents and accidents in the healtcare system. *Journal of Quality in Clinical Practice*, 18 (3), p. 199-211.

Ruciman, W. B., 2002. Lessons from Australian Patient Safety Foundation: setting up a national patient safety surveillance system: is this the right model? *Quality & Safety in Health Care,* 11, p. 246-251.

Schioler, T., et al., 2000. Incidence of adverse events in hospitals: a retrospective study of medical records: Danish study of adverse events. *Scandinavian Journal of Public Health,* 32 (4), p. 275-282.

Shaw, C. & Kalo, I., 2002. *A background for national quality policies in health systems.* Copenhagen: World Health Organization European Region Office.

Thomas, J. S. & Brennan, T. A., 2000. Incidence and types of preventable adverse events in elderly patients: population based review of medical records. *BMJ,* 320, p. 741-744.

Vincent, C., 2000. How to investigate and analyse clinical incidents: clinical risk unit and association of litigation and risk management protocol. *BMJ*, 320, p. 777-781.

Vincent, C., Neale, G. & Woloshynowych, M., 2001. Adverse events in British hospitals: preliminary retrospective record review. *BMJ*, 322, p. 517-519.

Vincent, C. & Coulter, A., 2002. Patient safety: what about the patient? *Quality & Safety in Health Care*, 11, p. 76-80.

Wears, W. R., 2004. Organization and safety in health care. *Quality & Safety in Health Care*, 13 (Suppl. II), p. 10-11.

Wilson, R. M., et al., 1995. The quality in Australian healthcare study. *The Medical Journal of Australia*, 163, p. 517-29.

World Health Organization, 2002. *Patient safety study: rapid assessment methods for estimating hazards: report of a WHO working group meeting*. Geneva: WHO.

World Health Organization, 2002. *Quality of care: patient safety*. Geneva: WHO. (Fifty-Fifth World Health Assembly Provisional Agenda Item 13.9 March).

World Health Organization, 2006. *World alliance for patient safety: Forward Programme 2006-2007*. Geneva: WHO.

Quality Improvement and Chronic Health Conditions — Implications for Research

Duncan Neuhauser*
Mary Dolansky**
Mamta Singh***

Research that integrates quality improvement (QI) concepts facilitates the generation of new questions and expands the scope of health services research. The focus of health services research is often on disparities; some people get a desired intervention or outcome and some, who could benefit, do not. The addition of QI concepts moves health services research beyond this scope and demonstrates how these disparities can be reduced and improvements made by asking different questions such as "are interventions or outcomes better today than yesterday?" Asking questions in this way expands the traditional research questions generated through health service thinking such as "is X associated with a better outcome as compared to the usual way of doing it?" This integration of QI concepts can occur with research at the macro-system, micro-system, and patient levels. This paper describes an agenda to enhance macro-system, micro-system and patient level chronic condition research through the integration of QI strategies. But before doing so some background and definitions are in order.

Quality improvement theory is a set of principles that involve knowledge, skills, measurement, and implementation of change using a systems-approach. Chronic health conditions, such as hypertension, diabetes, asthma, insomnia and depression call for different forms of care than acute events like accidents or heart attacks. Chronic health conditions require daily lifestyle interventions such as to quit smoking, exercise, change dietary habits and adhere to a medication regime. These daily lifestyle changes require constant monitoring and daily tests of change. If the desired change did not occur today, then what is a better way of doing it to achieve the required change

* Division of Health Services Research, The Charles Elton Blanchard Department of Epidemiology and Biostatistics, Case School of Medicine, Case Western Reserve University, Cleveland. Co-Director of Health Systems Management Center.

** Multidisciplinary Clinical Research Training Program Case Western Reserve University/Cleveland Clinic Foundation.

*** Department of General Internal Medicine Case Western Reserve University — MetroHealth Medical Center.

tomorrow? To date, traditional research on interventions to improve the process of care and outcomes for chronic disease management have had limited success. In addition, many research studies do not get replicated and therefore miss the benefit of improvement in the intervention that is achieved through repeated tests of change. The QI *Breakthrough Series* strategies promote the implementation of several small cycles of change in order to improve outcomes.

QI strategies include customer and statistical mindedness, organizational behavior knowledge and systems theory. Using these strategies, experts in QI ask "Who do we serve? How do we know we are meeting their need?" and "How can we organize to do it better? Another QI model is the Plan-Do-Study-Act (PDSA) model. This model enhances traditional research by building in the testing of repetitive improvements. The PDSA model moves research to 'action' research and facilitates the translation of science to clinical practice (Clancy, Fraser & Palmer, 2006).

Along with the integration of the above QI strategies, eight QI steps are proposed to enhance traditional research to improve the processes and outcomes of chronic health condition care. The QI steps are:

1) Define and communicate a clear mission statement that endorses improvement of chronic health conditions.
2) Define a set of key processes/interventions that will have to perform well to achieve the mission.
3) Define the key quality characteristics to measure these processes.
4) Collect these data in "real time".
5) Then put someone in charge of watching these data that have the knowledge to understand this information.
6) Give this person the responsibility for improving this care.
7) Give this person the resources to make improvements including "carrots and sticks" to be used to make the change.
8) Plan, Do, Study and Act to either hold the gain or make a new change.

Research at the macro-system level

Quality improvement steps integrated into research at the macro-system can lead to innovative research questions. First and foremost, these research questions require organizational leadership that defines a clear mission to improve chronic health conditions. Examples include "We wish to be the healthiest province in Canada" or "We want to have the healthiest work place of any aluminium company". Links between organizations that have such a mission and outcomes can then be studied. The

second step, defining the key processes or interventions needed to be implemented to achieve the aim, is imperative Examples in this step in chronic health conditions are the integration of healthcare across services such as inpatient, outpatient and long-term care. The third step, identifying the key quality characteristics to measure these processes, can be studied by determining at what point in the service delivery trajectory can hemoglobin A1C tests be performed. In the past, this was difficult to achieve. However, with the steady growth of information technology and electronic medical records, we are on our way to being overwhelmed with comparative quality data. The problem will soon be which measures are important.

Examples of the next step, collecting data in "real time", could look something like "how many patients over the age of 65 are in our hospital *now* did not receive their influenza immunizations?" The researchers with QI knowledge understand the importance of using current data; not last year, not last month, but today. To implement this step, organizations will need to organize health care information systems differently. There is a start in this direction. For example in Sweden there are 80 disease registries that collect data from patients under care. The real time data are used to improve care at the current moment. Pivotal in macro-system research is the integration of QI steps 5-7. The successful management of chronic conditions requires that someone is in charge of continuously monitoring the data and that this person has the knowledge to understand the data and the responsibility and resources to improve the care. Research in this area is fertile ground for future research in chronic health condition management. Some movement in this direction has occurred by diabetes specialists at Group Health of Puget Sound in Seattle, Washington. The population based diabetes specialist may be responsible for diabetes care for 10,000 enrolled members even though he may only see a few of them himself.

The integration of the QI step, "act to hold the gain", is an untapped area of research. How do organizations continue improvement work, what is needed? How is a corporate culture organized to promote long lasting change? Chronic condition research at the macro-system level can be advanced by use of the QI "plan, do, study, act" as this model of improvement will facilitate replication of research and the application of research to clinical practice. The repeated tests of change in care systems will lead to successful management of chronic conditions.

Research at the micro-system level

The micro-systems level refers to the department or clinic within a larger macro-system (hospital system). Micro-systems include areas such as asthma clinics, surgery centers, or general internal medicine practices. It is at this level that process redesign

needs to occur. Here the important issue of the re-engineering of care is carried out at a system level. Now consider these same eight QI steps reviewed above. Research implications at the first step involve the leader of the unit and the organization's hierarchy. The leaders at the micro-system level carry out the mission of the organization while balancing expense expectations. The awareness of the influence of the leader on the outcomes is an important piece to research on chronic condition care. The 2nd and 3rd QI steps, define a set of key processes/interventions and outcomes, are important to the improvement of chronic conditions. For example, in an asthma clinic a research question would be, "What is the process of good asthma care?" Here are two possible definitions "Good asthma care is when the patient controls the asthma and the asthma does not control the patient." This goal suggests some key quality outcome measures. The patient who has to go to the emergency service for an asthma attack is an example of the asthma controlling the patient. The clinic director would take on the task of reducing these numbers. Another goal could be: "Our asthma care is so good that our patients brag about it to their friends". This goal suggests different key quality measures such as satisfaction.

The growth of evidence based guidelines help to define the processes and outcomes of care. We are beginning to see examples of care where a high proportion of patients receive care under these guidelines. The QI step of collection of data in 'real time' will facilitate traditional research in chronic care. Redesigning information systems to provide 'just in time' clinic performance data is needed. For example, at Henry Ford Health System in Detroit, an internist can turn on her computer, find out how many diabetic patients she cares for, what percent are not up to date with their hemoglobin A1C test *today*, who they are and their telephone number. She can compare her percentages with those of her colleagues in her group and with other groups in the Henry Ford System (Hebert & Neuhauser, 2004; Olsson, J. et al., 2005; Alemi, & Neuhauser, 2006).

She is a knowledgeable participant in the delivery of care, wants to be an excellent physician and the system will reward her for making improvements.

Years ago, the noted medical sociologist, Eliot Freidson, observed some clinics where there was no collective oversight of quality (Friedson, 1970). Patient and doctor would meet behind a closed door and they would be the only ones who knew what happened. This would be in contrast to a clinic where there was an agreed upon norm of peer review and continuous group learning. Traditional research that measures the 'continual learning' (PDSA cycles of a team) of a clinic (micro-system) would provide insight into high functioning quality chronic care.

Other examples exist on the application of QI principles to re-engineer chronic care at the micro-system level. For example the use of group patient office visits such as those being used by David Aron MD and his colleagues at the Cleveland Veterans

Administration Medical Center. Small groups of patients meet once to learn about self management of their diabetes. Another example of micro-system research enhanced by QI strategies is the use of community coaches to improve the process and outcome of care for people with chronic illness. Community coaches, know the community culture, are able to read and write, and often have health problems themselves. Another example is the use of home monitors using cell phone technology to transmit "real-time" data about the home bound patient's condition. Our local Visiting Nurse Association refers to this monitoring system as "Molly the Monitor" and their use has led to a reduction in costly nurse home visits. Other examples of re-engineering processes at the micro-system level to include the integration of web-based information, computerized reminders, use of automated health-teller machine (similar to bank tellers) and 'retail' healthcare to replace some office visits. The 'retail' movement has already occurred in retail stores and pharmacies across the United States as more and more 'convenient care clinics' (CCC) arise (Hansen-Turton et al., 2007).

Research at the patient-level

Traditional research at the patient-level stands to gain the most by the addition of QI strategies. At the patient-level, QI thinking directs us to help each patient define their aim of chronic care, what they can do to meet the aim, and understand ways to know they are meeting the aim. These goals can be met using the tools of QI such as statistical mindedness and system thinking. Patients can be taught how to monitor variations in their condition and what they can do to manage these variations themselves. Patients can keep control charts of their outcomes so that they can witness the variation. This personal control will motivate patients to take responsibility for the management of their care. Research on the integration of control charts for self-care is needed. For example, a panel of 20 patients with hypertension could volunteer to participate in a study. With the use of the PDSA model', and the QI tool of a control chart, patients would control their hypertension. This could be done with greater statistical power, faster and cheaper than traditional randomized clinical trials. In addition, self-care of chronic conditions can be enhanced by the integration of a QI 'systems' approach. The acknowledgement that patients with chronic conditions live in a system will empower the patient to move from self-blame to the consideration of changing the system around them to manage chronic conditions. Research that acknowledges this system approach may provide new ways to help patients manage their chronic conditions.

Consider the same eight steps described above as applied to traditional research on self management of chronic conditions. The first step, "define a clear mission statement

that endorses improvement." A common cited example of this is in diabetes care. Patients decide their own priorities and goals. The patient might say "I am more concerned with not sleeping well than with my glucose level." These priorities need to be taken into account. When this patient has solved his sleeping problem he may be ready to start on his glucose level. This prioritization is too often ignored in traditional research studies. The second step, define the processes that need to be changed to meet one's goal A patient may say, "if it is raining out, I do not want to go running." To use the QI step, the research would incorporate alternatives to the problem using QI tools of brainstorming. In addition, processes that incorporate the 'system' can be addressed. The QI concept that the system is perfectly designed to produce the out comes it gets, can be applied and used. Change the system and the processes and outcome will change. For example if the chronic condition indicates a need to loose weight, some system changes can be implemented. These include selling your car and using a bicycle, moving to a fifth floor apartment with out an elevator, getting a young dog which needs to be walked frequently, eliminate the sit down desk in your office and only have a stand up desk, get a stationary bicycle attached to a battery which is the sole power source of your television, when you get home eat as much celery as you possibly can, and say good by to your friends at the pub and find new dog walking friends. Do all those things and your weight will change and stabilize until you once again redesign the system of your life. The 3rd QI step is to define ways to measure key characteristics of these important processes over time, day by day. Tools generated from QI include the use of graph paper to plot data over time to learn about common and special cause variation. This graphed information can be discussed with nurse, physician, family and support groups. Such graphs have become a routine part of care at the Huron Hospital Diabetes Center (Neuhauser & Diaz, 2007) and for hypertension care at the Cleveland Clinic (Chris Hebert MD) (Hebert & Neuhauser, 2004). Such run charts provide real just in time data so that learning and behavior change can occur immediately. One diabetic patient showed a special cause spike in her blood sugar level and labeled it "Kit Kat Bar peak" as a result of eating chocolate bars. By putting the patient in charge of their data, they are empowered to become experts on causes of variation and what they can do about it. They can systematically carry out experiments or 'plan do study act' cycles to learn the causes of variation. The gold standard approach to this is the factorial design of experiments where changes are randomly introduced (see Olsson et al., 2005). This will rarely be used in practice but it does set the ideal standard. A research question is how simple and easy can these control chart methods be made so they can be used by every one. Another untapped area of research is the delegation of responsibility to the patient to manage their own life. Too often market driven health care promotes cures that do not require any effort on the part of the patient other than to have the surgery, take the pill or eat this food product. This is persuasive,

because it promises effortless relief but it creates a passive patient/consumer. Chronic conditions, by definition, can not be cured so easily and require the patient to be involved in their care — without them, no intervention will be successful. Another area of research is what resources are needed by patients to make lifetime lifestyle changes and what rewards would motivate patients? Resources include the patients 'team' (family, supportive "buddy" or a peer group, and physician). Team concepts are integral to QI methodology. Research studies that examine the integration of the team and ways to enhance the team approach are needed. In observing hundreds of students doing personal improvements over 90% make a real improvement in their life (Alemi & Neuhauser, 2006; Alemi et al., 2000). This success rate depends on many things including choosing their own priority for change, paying attention through daily measurement, carrying out several small plan do study act cycles and knowing one will present one's own results in the form of a story board to classmates in a supportive environment.

Conclusion

QI concepts can be integrated into research at the macro-system, micro-system, and patient levels. All this comes with a large research agenda in support of improvements and the diffusion of best practices. This integration and the QI processes will in turn create innovative research ideas and move the traditional health service research to the next level of science.

Reference list

Alemi, F. & Neuhauser, D., 2006. *A thinking person's diet and exercise program.* Victoria, British Columbia: Trafford Press.

Alemi, F., et al., 2000. Continuous self-improvement: systems thinking in a personal context. *Joint Commission Journal of Quality Improvement,* 26 (2), p. 74-86.

Clancy, C., Fraser, I. & Palmer, C., 2006. Turning research to ACTION through delivery systems. In 1st Annual Meeting, Washington, D.C., Jun 1, 2006.

Friedson, E., 1970. *Professional dominance.* New York: Atherton Press.

Hansen-Turton, T., et al., 2007. Convenient care clinics: the future of accessible health care. *Disease Management,* 10 (2), p. 61-73.

Hebert, C. & Neuhauser, D., 2004. Improving hypertension care with patient-generated run charts: physician, patient, and management perspectives. *Quality Management in Health Care,* 13 (3), p. 174-177.

Institute of Medicine. Committee on Health Professions Education Summit, 2003. *Health professions education: a bridge to quality.* Washington, D.C.: National Academy Press.

Landon, B. E., et al., 2007. Improving the management of chronic disease at community health centers. *New England Journal of Medicine,* 356 (9), p. 921-934.

Neuhauser, D. & Diaz, M., 2007. Quality improvement research: are randomized trials necessary? *Quality & Safety in Health Care,* 16 (1), p. 77-80.

Olsson, J., et al., 2005. The one-person randomized controlled trial. *Quality Management in Health Care,* 14 (4), p. 206-216.

PART 2
PLANNING QUALITY AND SAFETY IMPROVEMENT RESEARCH

Challenges in the Early Stages of Quality Improvement Research — Defining the Focus and Adjusting to Events in the Services Studied

ANNA ESSÉN*

SUMMARY

Phase of the research when this was written: A review of the literature and a theoretical paper was written after the initial phase of my data gathering and in the middle phase of my PhD education.
Methodological challenges: Constructing a theoretical problem and a theoretical framework to use.
Practical challenges: Dealing with time delays related to the implementation of the technology studied. Waiting for results to show.
Main lessons for other researchers: Position your paper within a discipline, research area and identify a target journal as early as possible. Talk about your paper with peers.

Introduction

My PhD thesis is about the use of information technology in home care for older people and examines whether and how standardization and customization can be combined. I studied a municipal elderly care organization that decided to invest in a *telemonitoring* technology in home care in 2003. The care managers wanted to learn about the possibilities and risks associated with the use of such new *telehealth* technology. I designed a comparative study, with an experiment group consisting of 23 single senior households who would get the new *telemonitoring* service and a control group where 23 seniors would get standard services. Choosing which monitoring technology to implement and finding municipal resources for purchasing this technology required some time. Almost two years passed before the technology was installed in the elderly care organization.

* University of Stockholm School of Business — Sweden (aes@fek.su.se).

A note at the end of this paper describes the methods for data gathering and gives one example of how I presented the data in the findings section of the paper "balancing standardization..." which I wrote in 2006, in the early-middle phase of my phd education.

The idea to write a paper at this time came from my supervisor, who encouraged me to submit a paper to "the academy of management meeting", a conference with peer-review proceedings. Due to the delays, I had just started gathering data about the care personnel's use of the new technology at this time. I did however have "old" empirical data gathered for a report that I had worked with before my PhD education. This previous report was concerned with to what extent there were variations in the behavior of care personnel. I decided to write a paper based on the empirical data already gathered as I would not have time to generate any new data; the deadline for submission was soon in time. The question was: what knowledge gap could I begin to fill with this data? What problem could I solve with this empirical evidence?

The challenge of addressing a relevant problem

The inspiration to write about standardization emerged when I read an article by Hanlon et al. (2005) in the management journal "Human Relations". I didn't agree with what I read. Somewhat simplified, the line of argument in Hanlon et al. (2005) was this: technology equals standardization. As a consequence, technology reduces the possibilities for care workers to provide individualized services. This argument was overly one-sided from my point of view. I had a feeling that the technology (telemonitoring) I was studying could contribute not only to a more standardized but also to a more customized care service provision.

Now, intuition, or "having a feeling" about a certain state of affairs does not suffice to make an academic argument. It struck me that I did not have much data to support my claim that technology can lead to standardization as well as customization. My interviews were performed at a time when the use of the new monitoring technology was in a nascent stage and the empirical data would show few radical effects. In general, it would be very difficult to measure and compare the degree of customization and standardization in the services provided to the experiment group and the control group respectively. I had some quotes supporting the idea that the technology could imply an increased standardization and customization but this was not enough for an article.

I returned to my interview transcripts and I found that they did provide many insights about how care work is a matter of combining dealing with expected problems and needs in pre-defined ways on the one hand, and dealing with context-

specific and unpredictable issues on the other hand. I concluded I would not write a paper focusing on the impact of technology on standardization and customization. Rather, I would illustrate how the organization of elderly care more generally relies on balance between standardization and customization. I would discuss technology as one element in this larger context.

The question was: was there any established theory that I could use to support my argument that care provision is a combination of standardization and customization? What conceptualization could I use to clarify this claim and to structure my empirical findings?

The first round

There is a vast customization literature, which suggests that organizations can produce customized products by means of standardization. I spent weeks on reading this literature, realizing that it mainly deals with how manufacturing firms can combine modules to create individualized products. Many mass customization scholars stated that there was a need for more research about mass customization in services. OK, I thought, here is a gap. I could contribute to the mass customization literature by elucidating how mass customization works in care services settings. I could also discuss the limits of mass customization theory in this setting where some services are completely customized. I wrote a draft.

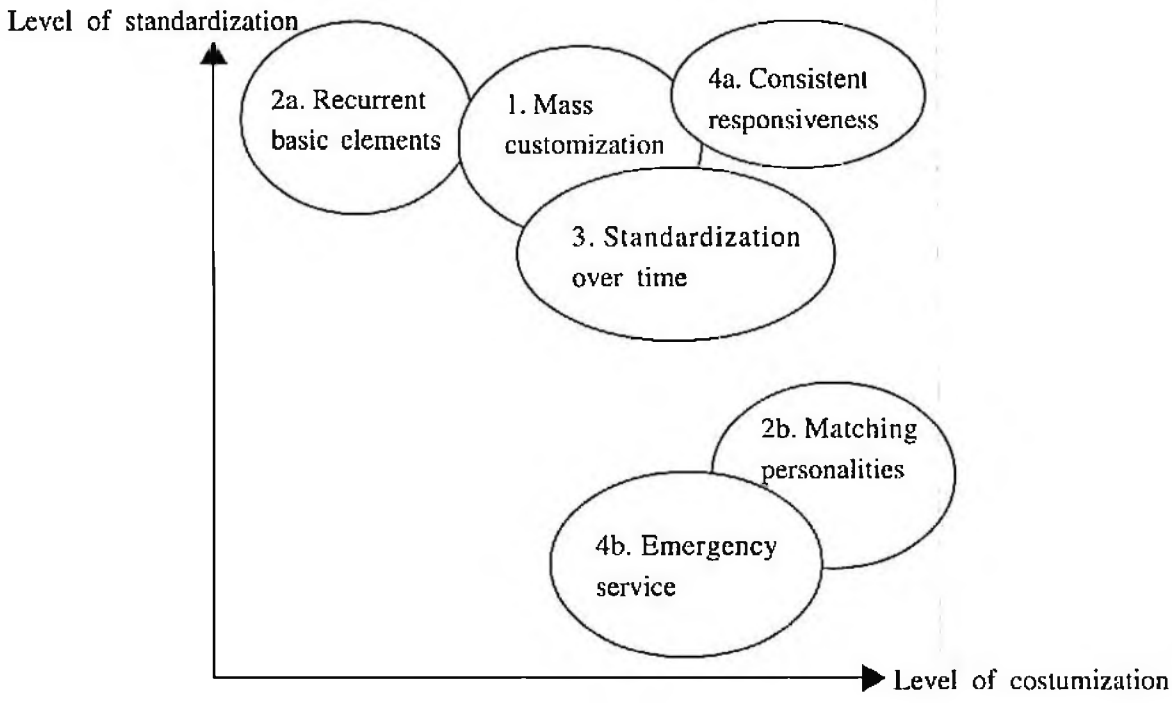
However, a meeting with my supervisors made me painfully aware of that "contributing to the mass customization literature" was not a good idea, as this concept had buzzword connotations. OK. I put the draft in the trash. I started reading about earlier works focusing on either standardization or customization in the service marketing literature. This literature mostly contrasted standardization and customization but provided some useful definitions. Finally, I bumped into a few healthcare articles concerned with how care organizations can *combine* standardization and customization in their service delivery (e.g. Bohmer, 2005). Wow! I had "colleagues" who where also interested in this issue! My "colleagues" did not argue that they had "solved" the issue, but rather stated that more research was needed about this topic. Fortunately (J). I decided to draw on Bohmer (2005) and others, but to add something beyond their argument I needed to provide some conceptual model of *how* care providers can combine customization and standardization. I needed to provide some theoretical explanation for my observations. I happened to recall an article about strategies by Mintzberg (1983), which, I thought, did shed light on how the care providers managed to combine standardization and customization (see Figure 1). Further, I tried to inductively "innovate" a new kind of visualization of services: a

"both-and" model (see Figure 2) which would represent an alternative to the dominant view of standardization as customization two end poles on a continuum. Yes. This must be a theoretical contribution I thought.

Figure 1
Assumptions underlying the Standardization-Customization continuum in the extant literature

	STANDARDIZATION ◄	► CUSTOMIZATION
Approach to variability	Reduce variability	Respond to variability
Needs	Homogenous	Heterogeneous
Assumption	One best way of satisfying needs	No single best way
Process	Predefined	Adjusted to needs and context
Knowledge	Explicit, possible to transfer	Tacit, experience-based
Low goal-attainment if	Deviation from standard	Insufficiently adjusting process
Management style	Supervising that specification is met	Delegated decision making
Output	Uniform output	Heterogeneous outputs
Associated advantages	Economies of scale, predictability	Customer satisfaction
Associated disadvantages	Resistance from employees	High cost, lack of consistency

Figure 2
The suggested integrated standardization-customization model



I finished the new version of the paper under much stress and sent the paper (the version presented in this anthology) to "the academy of management meeting". It was accepted! This was important as it gave me inspiration and energy. I received feedback from two reviewers. One was very positive and merely asked me to add some facts to the method section. The other was very critical, underlining that the Mintzberg model should be eliminated as it confused rather than added to my analysis and that my empirical data did not support the model in Figure 2. I realized more work would be needed to make a theoretical contribution on basis of my empirical data. I presented the paper at the conference but after this, I put the paper aside as I had to deal with other more urgent tasks.

A few months later I went to Boston on a scholarship and met with Bohmer (as mentioned above, he wrote one of the papers I drew on). He told me about the ordeal he'd gone through to publishing his paper (which I built on), and the various bad versions that had preceded the published version. He had experienced difficulties in choosing between management/business journals and healthcare journals, ending up with a healthcare journal. Realizing that he, a researcher at Harvard, also had problems with publishing was actually nice. It taught me that positioning is always difficult. That first drafts are often lousy. That publishing is a matter of rewriting; about starting all over again. And again. And again.

The second, third and fourth round

A few months ago in May 2007, I decided to make an effort to improve the paper. I wanted to make it publishable and possible to include in my thesis. I read the Hanlon et al. (2005) article that initially triggered the idea to write the paper anew. I noted that Hanlon et al. (2005) used routines rather than standardization as their main concept. I realized that routines may be a better concept than standardization to discuss. Using routines instead of standardization as search term, I found numerous new articles. Many healthcare scholars argued that routines are detrimental to the individualization of care services and thereby to the quality of care. On the other hand, there were scholars arguing that routines improve healthcare quality by reducing avoidable mistakes. In short, there was a debate between routine advocates and routine critics. I thought they were both wrong and right, both too one-sided.

I mentioned my interest in routines to my supervisor, who handed me an article from in a business journal by Feldman & Pentland (2003). It was excellent! This article provided a conceptual model of routines as consisting of several dimensions, a model recognizing that a routine can incorporate pre-defined (standardized) and emergent (customized) elements. This model seemed attuned with my observations.

Indeed, this model supported the argument I wanted to make and I felt I could use it to structure my empirical findings.

I rewrote the paper. This meant eliminating a lot of references and adding new ones. The new aim was to reveal and criticize the assumptions underlying routine critics and routine advocates (in the healthcare literature) and to provide an alternative framework that would allow for a less one-sided and more fruitful view of routines. I was rather harsh against the "previous literature", emphasizing how wrong they were to make my contribution clearer. I asked my supervisor to read the new version of the paper. My supervisor was not satisfied. The model I was applying was already "discovered" by scholars in my field, he said. Even if the healthcare authors I was "fighting" against may not have read the article by Feldman & Pentland (2003) I could not claim I was making a contribution simply by applying this model. He told me to, instead of engaging in polemics with healthcare scholars, I should — in a modest way — criticize those "in my own field", i.e. the management literature. Instead of simply applying the model by Feldman & Pentland (2003), I should find some weakness in their model. May be I could add something new to their model based on my empirical data?

I found it difficult to criticize the Feldman & Pentland (2003) model as I thought it was great. But perhaps someone else had identified a weakness in this model? I found a few papers referring to this model, and in the end of the papers there were usually demands for more research about certain aspects of the model. I also found a review where the author mentioned this model and he argued that the internal relations of this model were little elucidated. I felt I had data to shed light on these aspects. Hence I rewrote the paper again. The aim of the new paper was to, departing from the Feldman & Pentland (2003) model, unpack the routine and explore its internal dynamics. This delimitation enabled me to make greater use of my rich data. I could really elaborate on this smaller area, with more nuance (compared to the previous versions).

I asked for feedback again. My supervisor thought the paper was good but that I could narrow the purpose further. He also suggested I improve my problematization of the previous literature. He gave me an article about problem formulation by Locke & Golden-Biddle (1997), which was very helpful. He further brought to the fore that my quotes actually highlighted how care personnel's background influence how they execute routines and how this adds variety to the care services provided. It may sound trivial but the Feldman & Pentland (2003) model did not make this explicit. I rewrote the paper. I added to the new paper that its contribution was to explore the internal dynamics of the routine and more specifically to highlight how individual's background plays a role in the execution of routines, adding this aspect to the Feldman model.

I decided to submit the paper to human relations. I deemed it necessary to take on a more modest approach, not criticizing other publications in this journal, not

arguing that they were one-side and wrong, while Bohmer, Feldman & Pentland and I was right. Rather, I acknowledged that previous works provide important insights (including Hanlon et al., 2005), maintaining that my paper simply elucidated another perspective and thereby could complement the dominant view in the extant literature.

I am still waiting for an answer as regards my submission to Human Relations. Even if my manuscript will be rejected, I however expect some valuable comments concerning how I can improve it. I will most likely rewrite the paper several additional times.

Lessons learned

What I learned from this cyclic process was that problem formulation is not easy. What is a relevant problem depends on whom you ask. One way of "finding" a gap in the literature is to read what kind of research review-authors and other scholars ask for. However, it is important to refer to articles in journals within your specific discipline and research area. What's more, it is crucial to consider what empirical data you have/will be able to gather. Addressing some problems might require large-scale studies and study periods spanning over several years.

My journey was triggered by my conviction that some of the arguments flourishing in the management debate were "wrong" — this was a weakness in the literature that I wanted to address. It could be a good starting point to find someone you disagree with. It is however not enough to simply criticize or disagree with previous scholars. You have to come up with a better alternative. One possibility here is to inductively innovate a completely new conceptual solution to the problem, based on your empirical data. I however found it easier to rely on a more established but rather abstract theory, modify this and to propose the modified theory as "my" alternative. I used my empirical data to illustrate the usefulness of the conceptual idea I proposed.

I also learned from this process that research is about rewriting papers. We are research student to get better, right? Hence, we should be grateful for critique. This is easier said than done. But hang in there.

A note on the methods used for data gathering and analysis

Research setting

The public home-help services setting is interesting in relation to the aim of the research, as public home-help providers need to guarantee a consistent delivery of a

minimal standard of service which is regulated by the law, ensure a fair distribution of services and adherence to financial restrictions while also delivering services that respond to the needs of every individual. The research reports empirical data about two Swedish home-help providers. Provider A is privately owned and the largest private producer of publicly financed home-help services in Sweden. Provider B is publicly owned and serves Heby, a rural municipality in Sweden.

Data collection and analysis

The study is based on participant observations (McGall & Simmons, 1969), performed during a large number of informal and formal meetings at provider A and B during May-September 2003 and during 2004-2005. Field notes were taken when relevant. Further, 37 more focused, in-depth interviews (McCracken, 1988) with home-help managers, home-helpers at provider A and B, and aid-coordinators employed by the municipality of Heby have been conducted. Interviews were performed face-to-face (30), via phone (7) and on two occasions: during May-September 2003 and during December 2005. Interviews started with the researcher asking informants to describe the way they deliver services at a rather general level.

These descriptions unfolded many details worth further probing, such as: how do you act when unexpected events occur in this situation? The interviewer proceeded more explicitly focused on the themes in the theoretical framework, asking about e.g. variability in needs encountered, approach to variability, the degree of predefined/ emergent dimensions, knowledge required to execute services, managerial style and strategy, and the benefits and risks associated with the standardized/customized practices in use. Questions were asked in an open-ended fashion and were not specified in detail prior to the interviews, allowing the interviewer to word questions spontaneously (Patton, 2004). More structured methods would leave little room for unexpected issues to emerge. F-2-F interviews lasted for 90-120 minutes, phone interviews about 30 minutes. Answers were audio-recorded and transcribed.

The content of interview transcripts and field notes was coded as related to standardization and customization respectively based on the themes in the theoretical framework (Figure 1). Empirical data on strategies were further coded on basis of the Mintzberg and Walters (1985) typology. Hence, directed content analysis (Hsieh & Shannon, 2005) was performed. The author was however also open for unexpected issues to emerge (Patton, 2004). The findings presented represent a broader set of frequently mentioned examples encountered during interviews/observations. The quotes selected represent the dominant view expressed by informants. The author performed, transcribed and analyzed the interviews.

An example of the empirical findings and how they were presented in a final paper

This section of the thesis presents empirical illustrations of how standardization and customization are combined and integrated in the delivery of home-help services. The empirical examples are related to a few major dimensions of home-help provision.

Combining standardization and customization

In Sweden, the delivery of home-help services is publicly financed and regulated by law. The social service law declares the legal right of every individual to receive support in order to maintain a decent standard of living, ensuring the dignity of each individual. The law also asserts that this implies different levels of support services, depending on the situation of each individual. Hence, attaining the goal as articulated in the law in a consistent manner requires that seniors be provided with different types and amount of home-help services. As public resources are scarce, fairness is a major concern. Seniors in need of support and care in their home send formal, standardized applications to the municipal gatekeepers: the aid-coordinators. The aid-coordinators make the decisions about how much and what kind of service each senior is to be granted, i.e. how much service the municipality will finance.

Handling applications and creating a service plan

Interviews made it clear that the municipal aid coordinators perform an investigation consisting of several pre-determined steps before deciding who is granted municipal home-help service and how much. The final decision is outlined in a *service plan,* which includes descriptions of what assistive technologies the senior is granted, and specifications such as "help during meals, 20 minutes, every morning", "help with shower 10 minutes, twice a week" etc. When creating the service plan, the aid coordinators choose among a finite set of pre-defined home-help services that are considered vital for a decent standard of living. This degree of standardization and centralization reduces the internal variability stemming from personnel. *"The detailed service plans constrain the freedom of aid-coordinators or home-helpers to deliver services according to their own principles... individual decisions would vary highly depending on the daily mood of the worker, the relationship between the worker and the senior, and the senior's talent in convincing the worker about his/her needs...* (aid-

coordinator). *"Cute seniors would perhaps be granted more help, leaving fewer resources for other seniors"* (manager).

Of course, seniors have heterogeneous needs. The standardized process of handling applications results in heterogeneous output in terms of customized service plans, "*although the service plans have the same format and are created in the same way, their content varies*" (aid-coordinator). The service plans are adjusted to the unique situation of each senior as they consist of different combinations of standard components.

The intention behind the detailed and standardized service plans is to ensure a fair distribution of home-help service in order to secure each individual's legal right to support. However, the service plan does not specify *how* the services are to be performed. Further, the service plans do not take emergencies into account.

References

Bohmer, R. M. J., 2005. Medicine's service challenge: blending custom and standard care. *Health Care Management Review,* 30 (4), p. 322-330.

Halnon, G., et al., 2005. Knowledge, technology and nursing: the case of NHS Direct. *Human relations,* 58 (2), p. 147-171.

Hsieh, H-F. & Shannon, S.E., 2005. Three approaches to qualitative content analysis. *Qualitative Health Research,* 15, p. 1277-1288.

Locke, K. & Golden-Biddle, K., 1997. Constructing opportunities for contribution: structuring intertextual coherence and "problematizing" in organizational studies. *The Academy of Management Journal,* 40 (5), p. 1023-1062.

McCracken, G., 1988. The long interview. Thousand Oaks, CA: Sage (Qualitative Research Methods Series; 13).

McGall, G. J. & Simmons, J. L., ed., 1969. Issues in participant observation: a text and reader: 1-5. Reading, MA: Addison-Wesley.

Patton, M. Q., 2004. Qualitative research & evaluation methods. 4th ed. London: SAGE.

Plan for an Evaluation of Implementation of Total Quality Management Programmes in Saudi Arabian Medical Laboratories

AKRAM ALOQBI*

Introduction

Rapid development in Saudi Arabia has led its government to use modern management methods such as Total Quality Management (TQM). There is a considerable literature suggesting benefits to implementing TQM in hospitals, but there are few empirical studies describing programmes in the health care sector and especially in non-western countries. There are even fewer studies of TQM in medical laboratories, which is the focus of this study.

The Saudi government has spent billions of dollars to create an infrastructure to guarantee adequate and appropriate Health care for all citizens. Modern hospitals have been built and foreign expertise has been recruited. Indeed, Saudi Arabia allocates around 6.5% of its gross national product (GNP) to Health care expenditures (MOH, 2001). Similar to UK, Saudi Arabia provides free health care services, and no payment is required at the point where the care is provided. But unlike the UK, Saudi Arabia does not levy income tax. Since the annual population growth rate is very high and nearly half of the population is under the age of 15, the cost of providing health care services, and maintaining modern hospitals, will grow. This places additional pressures on health care.

The result is rising criticism of Ministry of Health (MOH) services by the public. Indeed, the MOH is less and less able to provide effective and responsive services to meet the increasing demand by residents who consider the provision of health care services as a right rather than a privilege.

Patients who perceive the MOH services as being of poor quality may increase their use of private providers. Therefore, patients' views of services offered are becoming increasingly important. It is necessary to assess and compare the quality of Clinical Laboratories provided by both MOH and private hospital or providers. As

* Doctoral Researcher at Royal Holloway, University of London, UK and officer in the Ministry of Health of the Kingdom of Saudi Arabia (akramaloqbi@yahoo.com).

Saudi Arabia moves toward to the privatisation and restructuring of its health care system, issues of quality have become of increasing concern and patient perceptions of quality will be factors in providers' survival and successes.

Assessment of quality has not been applied or studied widely in the clinical laboratories sector. Only a few studies compare the laboratories quality of different hospitals or providers. This is the first study that involves the comparison of the quality of clinical laboratories in public (MOH) and private hospitals and private providers. Further, this study attempts to use self assessment by managers and staff, and assessment by patients.

What is quality?

Many authors have defined and discussed quality and TQM from different perspectives — these are not reviewed in this paper (Zabada et al., 1998, Donabedian, 1989, Palmer, 1997, Haddad et al., 1998, Nwabueze and Kanji, 1997, Deming, 1986, Juran, 1989, Crosby, 1979, Kurz, 1995, Short and Rahim, 1995, Anderson, 1992, Geber, 1992, Bergman, 1994, Yasin, 2002, Hansson, 2000, Staniszewska and Henderson, 2004, Kerssens et al., 2004, Ummel, 1991, Sewell, 1997 and Rahman and Bullock, 2005).

Management approaches that are seen as essentially western in origin are viewed with some suspicion in many Middle Eastern and Islamic countries, and this represents an obstacle to their acceptance (Al-Zamany et al., 2002). Many researchers suggest that the differences between the Islamic values and Western values represent the main barrier to implementing western management theories and the main reason for the failure of such implementations (Bjerke et al., 1993; Al-Meer, 1999).

Wong (1998) stressed that many of the TQM programmes implemented in the developing countries fail due to the lack of a real understanding of the principles. Al-Khalifa (2000) suggests that the general way to move forward in developing countries in terms of quality practices is to create a driving force, which is usually associated with pressure from customers or an initiative from the owner or the managing director.

Al-Zamany et al. (2002) examined the cultural acceptability of quality management exemplified by the European Business Excellence Model (EFQM) in the context of Yemen. The study also investigated the barriers to the implementation of such a model. Al-Zamany et al. (2002) also indicated that there was no mismatch between the activities proposed by the model's criteria and the Islamic culture, since all respondents perceived them as generic terms that reflect good management practices.

One obstacle that the organisations in this study encountered was the resistance to changing the culture. Hence considerable work needs to be done to set priorities for implementation and to facilitate gradual change which will be accepted more readily.

Saudi Arabia is a country of rich culture and enormous diversity. Saudi Arabia health care organisations rely heavily on foreign labour and expatriate technical and medical expertise from the USA, Europe, and other developed countries. TQM implementation in this context therefore has to take into account 'cultural diversity', as well as cultural differences.

Although general TQM research in the Saudi Arabia has increased in the past years, it is still very limited in the health care arena. Most studies investigated general TQM issues. There is no study which explored the evaluation of TQM programme in specific department such as clinical laboratory in hospitals or independent provider in Saudi Arabia. The literature states that to implement the TQM philosophy, revolutionary changes in values, methods, and attitudes must be made. The effects and acceptances of such changes in Saudi Arabian clinical laboratory arena have yet to be discovered.

Lagrosen (2000) completed a study in a hospital in Motala, Sweden, the maternity clinic of which has received the Swedish Quality Award for the health sector. This hospital was trying to implement TQM. The purpose was to assess the effects of using TQM, thereby judging whether this use could be valid, and to find some success factors for implementing TQM in this kind of organisation. Certain quality dimensions have been defined. Based on them, the results of quality have been assessed and found to be positive. The major positive effects were better evaluations, increased ability to implement changes, and increased creativity. Only two negative effects were found: a temporarily increased workload, and envy from the other clinics. Since the positive effects greatly outweigh the negative ones, this gives indications that the use of TQM in hospital could be valid. Furthermore, some success factors for implementing TQM were found. The foremost of these include sufficient information, commitment by management, and evaluations of the operations.

Øvretveit (1997) studied the experience of nine public hospitals in different European countries, which describe themselves as implementing TQM programmes. One of the aims of the research was to evaluate the hospitals' progress in implementing TQM, and compare their actions to a model definition of the aims and sub-aims of TQM using quality experts such as Deming and Juran. Their findings indicate:

- Quality was not viewed as an integral part of daily work: The main reason given was lack of time, but the underlying reason was that the quality methods and approaches used in hospital were mostly concerned with documentation, or activities that were not directed at improvements to clinical care.

- Emphasis on results: There appeared to be a lack of 'performance orientation' to quality activities, and no clear results were expected from groups, or targets set for them.
- Management component of TQM: Only two hospitals put effort into ensuring that managers were developed for their new role, and these hospitals were trying to ensure that quality was a joint management and professional activity. Other hospitals found that their quality programme highlighted weaknesses in hospital management as a whole Quality structure. All hospitals studied had recognised the need for and worked out ways of employing quality expertise; they also found ways to integrate this expertise into the organisations' management process.
- Measurement: The hospitals employed different types of quality measures. These included staff satisfaction surveys, quality award assessments, patient complaints' surveys, professional quality measures, cost of quality measures, and specific quality projects.
- Scientific methods for systematic improvement: The hospitals studied used different methods to assure and improve quality, but there was little awareness of a disciplined scientific approach to quality improvement. Too much time was spent on documentation, planning, policy making, and formulating and checking standards, but there was little scientific basis for many of the standards and methods being implemented.
- Quality methods for the quality programme itself: None of the hospitals studied applied quality principles and methods in a reflective way when they established and implemented their quality programme.
- Training in Quality Methods and theory: All hospitals had spent some time and money on training programmes.
- Physician contribution to the quality programme: All hospitals found that physician leadership and involvement in quality activities is both essential and difficult to achieve.

There are few case studies in the literature available on quality in developing countries (Djerdjour & Patel, 2000). Although in the Middle East, TQM in health care is still young, there are some studies published which report their TQM experience. Benjamin and Seaman (1998) studied the implementation of TQM in a primary health project sponsored by the Ministry of Health in Bahrain. Their findings confirmed that TQM improves health care organisations. They found that participation, decisions must be based on relevant data, strategic commitment of top management must underpin improvement efforts, and threats to power and status can derail change efforts.

In Kuwait, Adrees (1996) studied the perceptions and expectations of TQM in the Health care sector using SERVQUAL measurement. The results indicate there is a negative relationship between patient expectation and senior management expectations. Also, the findings showed a stronger negative relationship between the patient expectation and the actual performance for the services received. This conveys that quality is not up to the patient's standard.

In the Kingdom of Saudi Arabia, Saeed (1994) measured the degree of effectiveness in quality programmes in hospitals in the Ministry of Health. His sample was made up of the nursing staff of six public hospitals. Saeed's (1994) research concluded that one-third of the sample believed that TQM programmes were non-effective. Alsaloom (1997) studied the TQM system as applied in Al-Amal Hospital in the KSA. The findings illustrate there are many problems in trying to apply the TQM system, namely: Failure to follow one specified work plan, lack of regular assessment of the overall, quality control system, the absence of incentives' regulations to prompt staff to improve their performance and a need for a careful establishment for setting-up performance criteria and to involve staff decision-making.

Evaluation theory

Health evaluation has developed from different traditional disciplines and backgrounds such as medical research, epidemiology, statistics and social science theory and methods. Thus, these diversities of subjects, perspectives backgrounds lead to disagreement among evaluators regards evaluation theory and practise for example the role of evaluator; which values should be represented in the evaluation; the questions that evaluator should ask; which is the best methods in the limited time and resources; ensure the utilisation by the evaluator; what elements do or should influence the previous options with regard to role, values, questions, methods and utilisation (Øvretveit, 1998).

However, there are some points of agreement among theorists such as evaluators have to be more active to ensure that their results are acted on; evaluated are usually unwelcome; any single evaluation have limitations; time and resources constraints for the evaluations make it for difficult transaction (Øvretveit, 1998).

Evaluating quality

Quality evaluation often examines outcome for example patient satisfaction, service process which assist the provider to the aspect that led to quality outcome and

what change might be need, third type is an experimental evaluation to find out whether the specific features of the service really lead to high or low quality outputs, this can be done by external evaluators or by internal one (Øvretveit, 2002).

Øvretveit (1998) provide the following questions to clarify which type of evaluation is wanted:

- What is to be evaluated? Patient quality, carer quality, professional quality, management quality inputs, process or outcome one feature or part of a service or the complete service.
- Who makes the evaluation? External quality inspectors, provider for an internal self-review, quasi-independent developmental body, external specialists or academic.
- When? Regularly, special studies, in response to a crisis or a complaint.
- Why? For accountability, to protect patient, to provide information to patient or purchasers, to help continual self-improvement or as an integral part of quality assurance.
- For whom? For taxpayers, patient or patients associations, government, owners, managers, health personal or scientific reasons.
- How? Using ready-made standards and systems, deriving and using users' and stakeholders' criteria, deriving and using professionals' and managers' criteria or using evaluators' criteria.
- Which design? Patients' perceptions following the service, expectations compared to experiences, trends over time for one service, service quality comparisons.

Why evaluate quality?

There are many reasons to assess quality for example:

We evaluate quality for patients to be able make up to date choices with regards which service to use, or whether to undertake a treatment. To guarantee that the quality judgment of people who do not have power or voice are recognised. For example Mamdani and Bangser (2004) provide a literature review concerned the poor people's experiences of health services in Tanzania, and they pointed out the key barriers that the poor face in accessing quality health care, this study shows that the health care quality should be available to people whether poor or rich. So that professionals can develop their practice and observe the effect of service change on their quality of practice. So that managers

can measure value for money, insure that all patients' interests are served and raise control over professionals. So that government can protect the public. To contribute to scientific knowledge about the reason of high and low quality outcomes from the systems of care (Øvretveit, 1998).

Evaluation of health care quality faces many challenges. Lawton (1998) summarised these difficulties in:

- Complexity of quickly collecting and interpreting a hug source of data.
- Defining the objectives of complex service where multiple objectives conflict.
- Lack of correlation between overall organizational objectives and specific objectives.
- Lack of relevant and measurable targets for final output and outcomes.
- Lack of resources to build data.
- Staff resistance to data collection.
- Lack of staff evaluation training.
- Cost of performance evaluation.
- Lack of interest.

Evaluating a quality programme

Three common types of quality initiatives are audit, accreditation and quality programmes. Quality programmes for example were introduced to health care organisation to train staff to use quality methods, evaluation the programme itself is a part of this programme, for example national or regional TQM in health care services in UK and Makkah Region Quality Programme (MRQP) in Saudi Arabia. However, to evaluate such programme:

- First, the programme and its history should be described and that depend upon evaluation users' criteria and questions. Common question are, for example, what should we do to raise personnel's motivation to work on quality improvement?
- Second, study the service providers who are the target of the programme and ask them about their opinion of the programme, its progress and impact.
- Third, compare the service quality plan and objectives against what had really been done in the service.
- Fourth, compare what the service has done to a prescriptive model of what service should do to practise a successful quality programme.

- Fifth, measure the quality of the service directly before the programme and at different intervals throughout the quality programme.
- Sixth, use framework of main choices which all service encounter when pursuing a quality programme and compare it to those of other services (Øvretveit, 1998).

Self-assessment

Self-assessment is often used when health care organisations aim to improve and measuring the culture of quality. In recent years, self-assessment has become an essential management technique for continuously improving the whole system performance with a rapid increase in the number of organisations which have started, or are planning to start, self-assessment activities.

Self-assessment is a term that, to an increasing extent, is found in quality management literature, where it is often defined as "a process of evaluating an organisation against a model based on TQM". National and international institutes such as the EFQM in Europe and NIST in the USA have played key roles in the dissemination of self-assessment.

Research questions

The review of the literature together with an assessment of the needs of decision makers in Saudi Arabia at this time led to the formulation of these questions for the research:

- How far is Saudi Arabian Medical Laboratories from being excellent organisations? What are the current situations of Saudi Arabian Medical Laboratories?
- What are the critical success factors in TQM implementation in Saudi Arabian Medical laboratories?
- What are the inhibiting factors in TQM implementation in Saudi Arabian Medical Laboratories?
- What are the common areas of strength and key areas for improvement?
- How can Saudi Arabian Medical Laboratories improve and develop their services?
- How do top managers and all level of employees perceive TQM at the Saudi Arabian Medical Laboratories?

Research methodology (overview)

The research objective in this study is to evaluate and understand the approach of Total Quality Management (TQM) implementation in Medical Laboratories in the Kingdom of Saudi Arabia (KSA), and to recommend a proposed model for effective adoption and implementation. This research is concerned with assessing the level of TQM perception and understanding in the KSA Medical Laboratories. Also, it is concerned with studying the approach completeness, degree of maturity, and effectiveness of TQM programmes in the Medical Laboratories in Saudi Arabia.

The choice of methodology depends on the purpose, the objectives, the process of investigation, and the desired outcome. In this study, the plan is for the researcher to gather data to answer these questions by using questionnaires and interviews to collect the data from Saudi Arabian Clinical Laboratories. I believe that the mix of theory-then-research-then-theory strategy would be useful here because this is exploratory research. The general objectives of exploratory research are to gain insights and ideas. It is appropriate for problems on which very little is known and because if lack of knowledge at the beginning of the research. Exploratory research is also characterised by its flexibility with respect to methods used for gaining insight and developing hypotheses. Case study methods will be used because it is an empirical inquiry that investigates a contemporary phenomenon within its real life context, particularly when boundaries between phenomenon and context are not clearly evident. Quantitative and qualitative data collection methods allow the use of triangulation to increase the strength of the evidence.

In the first phase of the research, a survey will be used to establish levels of understanding and perception of TQM of the different laboratories in the KSA to measure the degree of "TQM maturity". In the second phase, a case study method will be used to assess the approach adopted in the TQM implementation process in certain KSA laboratories. In addition, closer investigation will be made of the feasibility and inhibiting factors of TQM programmes in the selected laboratories.

Stage one — survey

Phase one consists of the perception and understanding survey. The objective of this survey is to assess the level of perception and understanding of TQM principles among a sample of hospitals, in the KSA, which include two categories TQM, implemented hospital and TQM Non-implemented hospital. Each category will include samples of CEOs, laboratories managers, senior department, quality professionals, staff and patients. Afterwards, a comparison is made of the same category of laboratories and between different categories.

The survey method is the most widely used method in social science research. There are four types of surveys. They are self-administered surveys, mail surveys, telephone surveys, and face-to-face interview surveys. This study will use a self-administered questionnaire with personal follow-up to increase response rate, and will be used with personnel at all organisational levels to obtain clusters of opinions. The questionnaire method is appropriate also because the aim is not to establish direct correlations or causality.

The next step is to develop the survey instrument plan a system for recording answers. Both of these steps have been made by using a questionnaire from Ramirez and Looney (1993). This questionnaire was adopted to measure the degree of understanding and perception of TQM in the laboratories in the KSA. It was adopted for three reasons. First, the questionnaire is simple to use in recording answers and has simple measurement scales. Second, its validity has been established in TQM research. Lastly, the wide usage of the questionnaire gave the researcher the opportunity to compare the results of the survey to different geographical areas worldwide to establish any cultural influences.

According to Youssef and Zairi (1995), the list of 22 factors generated by Ramirez and Looney (1993) proved to be a very useful vehicle for checking applicability, order of criticality, and relevance of TQM in a much wider context. The Youssef and Zairi (1995) study deals with the empirical analysis of the TQM critical factors that were used in the Ramirez and Looney (1993) questionnaire in different regions of the world (UK, Middle East, Malaysia, and Singapore). Their project was an attempt to verify the applicability of a list of 22 critical factors based on the Malcolm Baldrige National Quality award criteria and the teachings of three TQM gurus to organisations operating on a global basis.

The next step in the Design and Planning Phase is to pilot test the instrument, the pilot test takes more time and effort, but it is likely to produce reliable measures. Also, it helps to find during the planning stage alternative explanations or threats to internal validity and how to avoid them. The last step in this phase of survey research is to draw the sample. Sampling is described as the process of systematically selecting cases for inclusion in a research project. By studying the sample and understanding the characteristics of the sample subjects, the properties can be generalised to the population elements. Therefore, sampling considerations are important in ensuring the validity of the research.

The target sample for the survey in this phase of the research will represent in the larger hospitals in the Saudi Arabia (more than 200). This is because they have longer and deeper experience with TQM. This comes from the extra governmental funding and the quality laws enforced upon larger hospitals. Statistics show 70 hospitals of this size and a sample will be drawn from these. There are also four types of hospitals

in the Saudi Arabia; they are public hospitals, private hospitals, specialist hospitals, and armed forces hospitals. The representation of different types is necessary in this study.

At this point, there is a need to determine the number of questionnaires to be distributed. According to Neuman (1997), a researcher's decision about the best sample size depends on three things:

1. The degree of accuracy required.
2. The degree of variability or diversity in the population.
3. The number of different variables examined simultaneously in data analysis.

Researchers disagree about what constitutes an adequate and respectable response rate most researchers consider anything below 50% to be poor and over 90% as excellent.

The last step in the data collection phase is to organise the data. After receiving the completed questionnaires, the answers of each respondent will be organised and coded into the Statistical Package for Social Sciences (SPSS) for PC, and analyse using the techniques available in the SPSS package.

Stage two — case study

While the survey method in this study will be used to answer the who, what, where, how many, and how much questions, the case study method would be used to answer the how and why questions. The aim is to examine in detail the approach to TQM implementation and its degree of effectiveness in a sample of Laboratories, and to compare and contrast the existing applications of TQM in Saudi Arabian Laboratories against criteria of more successful applications, taking into consideration the cultural differences. The case study approach offers advantages for the aims of the research:

- TQM implementation is not widespread in the Saudi Arabia.
- Access is not a problem and information can be obtained because the researcher has connections in the Ministry of Health.
- The researcher is a Saudi male (a local). This can help in reaching an understanding in problems of culture, for example communications and teamwork in the country.
- The research is exploratory in nature and the phenomenon under investigation (TQM implementation) is quite complex and does not have uniformity. The

case study method is an appropriate method for studying complex phenomena such as TQM implementation.

The aim was to assess the comprehensiveness of the approach to TQM implementation. Therefore, this research will use the 'Baldrige National Quality Program: Health Care Criteria for Performance Excellence'. This allows assessment of the degree of systemisation and highlights the critical and inhibiting factors. Also, to document types of benefits achieved against the degree of maturity of TQM, four factors were concentrated on: culture, patient satisfaction, cost, and top management commitment. This allowed greater understanding of the level of quality in the case studied Laboratories, and preparing for future model.

The unit of analysis in the case study phase is the hospital. Eight hospitals will be selected by the researcher to be studied. Four of them are TQM implemented hospitals and four of them are TQM non-implemented hospital in Saudi Arabia. The researcher will select the case studies hospitals from the sample in the previous Phase. The selected hospitals will be covered four different types of hospitals:

- A specialist hospital.
- An armed forces hospital.
- A public hospital.
- A private hospital.

The interview, participant and direct observation, and documentation are common types of data collection in qualitative research. These data collection methods will be used as essential sources of information for the case studies in this research. Interviews in case studies can be open-ended, focused, or a formal survey. The most common interview used in case studies is the open-ended. In an open-ended interview, the interviewer can ask the respondents for the facts as well as for their opinions about events. The nature of the Baldrige assessment that will be used in the case study phase to satisfy the research objectives suggested that an interview approach should be used. Open-ended interviews will be used in these case studies to understand and assess the process of TQM implementation. This type of interview was selected because the interviewer could use the Baldrige assessment as the guideline and allow room for interviewee opinions, thoughts, and any extra information.

Face-to-face interviews will be chosen because they are intimate, and the interviewer interacts directly and develops rapport with the interviewee. Also, during the interview, questions and answers can be further clarified. Some disadvantages associated with face-to-face interviews are that they are time-consuming and expensive, the interviewer may influence the response of the interviewee, the

interviewee can of course tell lies or respond in a socially acceptable way to make a good impression or to satisfy good self-image or to please the interviewer as the think. Also, the data involved are more difficult to summarise and analyse.

The questions in the interview should be understandable to the respondent, otherwise they lose their effectiveness. The rationale behind the interview is to allow the subjects to describe in their own words their particular experience and identify their attitudes. The only way to find out about the subjects' beliefs, attitudes and perceptions is to ask them directly. When the interview responses are put together with the interviewer observation, then a better understanding of the matter at hand can be reached. By making a field visit to the case study site to do the interviews, the researcher will use the opportunity for direct observation. The researcher will observe the systemisation of the Laboratory, conditions of buildings, workspace, sidewalk activities, patient's movement, nurse's politeness, etc. Keeping in mind that this research is exploratory in nature, the researcher decided to use purposive sampling in the case studies. Six interviews will be conducted for each hospital. The selected sample of interviewees included the people who are thought to be having and could provide the most information about the laboratory. The interviews will take place at the Laboratory, and all will be conducted in the offices of the interviewee. This provides the interviewee with a comfortable and relaxing environment. Appointments will be set before each visit to take into account the time of the interview because these interviewees are very busy people. All through the interview, the researcher will be asked for documented evidence of what will be said. This is to ensure that what will be said is true.

Data analysis

The coding will be by categorising the data. The data will be grouped together under the headings of the Baldrige assessment. These headings include: leadership, strategic planning, focus on patients and other customers, information and analysis, staff-focus, process management, and organisational performance results. Categorisations of data will allow interpretation which may be difficult to achieve if cases will be presented randomly. The data will read and re-read to enable concepts to emerge. Data analysis will be carried out to reduce, organise, and give meaning to the data. It is the process of bringing order to the data and organising the data available into patterns, categories and basic descriptive units.

Interpretation involves connecting meaning and significance to the analysis explaining descriptive patterns, and looking for relationship and linkages among descriptive dimensions.

A review of research generated the following hypothesis to explore and which will allow the findings to be related to others research:

1) Driving Forces that promote TQM implementation in the medical laboratories are:
   - Management commitment to quality
   - Improving quality and productivity connection
   - Increasing Patient orientation
   - Developing employee's involvement
   - Achieving a positive change
   - Improving communication between management & staff

2) Restraining Forces that inhibit TQM implementation in the medical laboratories are:
   - Lack of QM understanding
   - Lack of training
   - Lack of Strategic Quality Planning
   - Inadequate knowledge base
   - Organisational resistance to change
   - Lack of competent management
   - Low Morale
   - Organisational culture

3) TQM Critical Success Factors in Saudi Arabian Medical Laboratories are:
   - Education & Training
   - Quality Data & Reporting
   - Management Commitment
   - Recognition that people are the most vital and valuable
   - Customer Satisfaction Orientation
   - Role of Quality Department
   - Communication to Improve Quality
   - Continuous Improvement

References

Al-Zamany, Y., Dulaimi, M. F., Hoddell, S. E. J. & Savage, B. M., 2002. The cultural acceptability of the EBEM in Yemen. *Managerial Auditing Journal,* 17, p. 568-575.

Djerdjour, M. & Patel, R., 2000. Implementation of quality programmes in developing countries: a Fiji Islands case study. *Total Quality Management & Business Excellence,* 11, p. 25-44.

Donabedian, A.,1989. Institutional and professional responsibilities in quality assurance. *Quality Assurance in Health Care,* 1, p. 3-11.

Eiriz, V. & Figueiredo, J. A., 2005. quality evaluation in health care services based on customer-provider relationships. *International Journal of Health Care Quality Assurance,* 18, p. 404-412.

Hansson, J., 2000. Quality in health care medical or managerial? *Managing Service Quality,* 10, p. 78-81.

Kerssens, J., et al., 2004. Comparison of patient evaluations of health care quality in relation to WHO measures of achievement in 12 European countries. *Bulletin of the World Health Organization,* 82, p. 106-114.

Lagrosen, S., 2000. Born with quality: TQM in a maternity clinic. *International Journal of Public Sector Management,* 13, p. 467-475.

Lawton, A., 1998. *Ethical Management for Public Services.* Oxford: Oxford University Press.

Mamdani, M. & Bangser, M., 2004. Poor people's experiences of health services in Tanzania: a literature review. *Reproductive Health Matters,* 12, p. 138-153.

McLoughlim, V., Leatherman, S., Flether, M. & Wyn, J., 2001. Improving performance using indicators: recent experiences in the United States, the United Kingdom and Australia. *International Journal for Quality in Health Care,* 13, p. 455-462.

Nwabueze, U. & Kanji, G., 1997. The implementation of total quality management in the NHS: how to avoid failure. *Total Quality Management,* 8, p. 265-280.

Øvretveit, J., 1997. A comparison of hospital quality programmes: lessons for other services. *International Journal of Service Industry Management,* 8, p. 220-235.

Øvretveit, J., 2002. *Action evaluation of health programmes and changes: a handbook for a user-focused approach.* Oxon: Radcliff Medical Press.

Palmberg, K. & Garvare, R., 2006. Sustained quality management: how to receive the Swedish quality award twice. *International Journal of Quality & Reliability Management,* 23, p. 42-59.

Palmer, R. H., 1997. Using clinical performance measures to dative quality improvement. *Total Quality Management,* 8, p. 305-311.

Ramirez, C. & Looney, T., 1993. Baldrige Award winners identify the essentials of a successful quality process. *Quality Digest,* 14, p. 38-40.

Saeed, K. S., 1994. The extent of effectiveness in quality programmes in the Ministry of Health in K.S.A. *Arab Journal of Administrative Sciences,* 2, p. 5-37.

Short, P. & Rahim, M., 1995. Total Quality Management in hospital. *Total Quality Management,* 6, p. 255-263.

Staniszwska, S. & Henderson, L., 2004. Patient evaluations of their health care: the expression of negative evaluation and the role of adaptive strategies. *Patient Education and Counselling,* 55, p. 185-192.

Yasin, M. M., Zimmerer, L. W., Miller, P. & Zimmerer, T. W., 2002. An empirical investigation of the effectiveness of contemporary managerial philosophies in a hospital operational setting. *International Journal of Health Care Quality Assurance,* 15, p. 268-276.

Youssef, M. & Looney, T., 1995, Benchmarking critical factors for TQM: Part 2: empirical result from different regions in the world. Benchmarking for Quality Management & Techology, 2 (2), p. 3-19.

Zabada, C., Rivers, P. & Munchus, G., 1998. Obstacles to the application of Total Quality Management in health care organisations. *Total Quality Management*, 9, p. 57-66.

PART 3
INITIAL DATA GATHERING

The Impact of a National Breakthrough Collaborative to Improve Care for Depression

GERDIEN FRANX*

SUMMARY

This paper describes ongoing quality improvement research in depression care in the Netherlands. The focus is the evaluation of an improvement collaborative which uses the Breakthrough method developed by the Institute for Health Care Improvement in Boston. The aim of the Collaborative is to implement evidence-based practices in primary and secondary care for people suffering from a depression. We present the context of our research, the breakthrough method as an implementation strategy, our research method and the difficulties encountered so far during our research activities that started in December 2006 and will end in July 2009.

Introduction

In the Netherlands there are 300,000 new cases of depression per year, and around 750,000 persons suffer from the disease (Meijer et al., 2006). For a newly developed depression, the prognosis is good; 50% recover within three months. But if the duration of exceeds six months the prognosis deteriorates progressively and with a duration of 12 months long-term chronicity has already been reached (Spijker, 2001). Depression has serious consequences — it is associated with important limitations in social and occupational functioning. The consequences are compounded and extended by the fact that depression has a high tendency towards relapse, recurrence and chronicity.

Effective treatments are available for all categories of depression (Bijl et al., 2003; Ormel et al., 2003) and are recommended in national guidelines. Two

* Trimbos-Instituut, the National Institute for Mental Health and Addiction Utrecht, Netherlands (gfranx@trimbos.nl).

national clinical depression guidelines have been released over the past few years in the Netherlands — one multidisciplinary guideline adopted by the professional organizations of social workers, nurses, psychologists, psychiatrists, psychotherapists and pharmacologists, and one adopted by the national association of general practitioners.

Several studies have shown that health professionals do not follow the current guidelines (Tiemeier et al., 2002; Ormel et al., 2003; Spies et al., 2004). In a pre-study problems in Dutch mental health care that needed to be resolved, were identified (Henkelman & Franx, 2004). One of the major problems was related to the use of antidepressants. Antidepressants were prescribed by general practitioners to 68% of their patients with depressive symptoms, regardless of symptom severity. Guidelines recommend counselling and other less intensive treatments in mild cases (Spijker et al., 2003; Braspenning et al., 2004; Spies et al., 2004). Up to 30% of patients stop taking antidepressants within the first six to eight weeks, 15-45% stops psychotherapy treatment preliminary (Ormel et al., 2003).

Quality collaborative as a strategy to implement evidence based practices

Quality improvement collaboratives are increasingly used in the Netherlands to achieve rapid improvements in all sectors of health care. The Ministry of Health and Welfare has financed several quality improvement programs including some in mental health. In December 2004, the Trimbos-Institute, a national institute for mental health and addiction, initiated the first national improvement project on these problems.

The depression collaborative was the first of its kind in mental health, so far only the general hospitals had been participating in breakthrough collaboratives ran by the Dutch Institute for Healthcare Improvement (CBO) (Van Splunteren et al., 2003). The outcomes are promising: general practitioners managed to reduced the unnecessary prescription rates of antidepressant from 60% to 11% (Franx et al., 2006). More practices got interested and in December 2006 a second depression collaborative started to spread the positive results from the first pilot project. This second depression collaborative was set up to be evaluated by a research project that we will describe in the remaining part of this chapter.

A key element of the Breakthrough method is that, instead of focusing on changing provider behaviour, an improvement collaborative encourages organizations to target their system of care as well. Other basic principles of the breakthrough method are: the use of guidelines, multidisciplinary collaborative care,

continuous feedback loops, expert opinion leaders and change agents. Breakthrough collaboratives can take different forms, many local variations on the original model exist (Øvretveit et al., 2002). The second depression collaborative has the following characteristics:

- 25 multidisciplinary teams involving general practitioners, psychiatric nurses, psychiatrists, psychologists, psychotherapists, social workers pharmacists, physiotherapists, a facilitator.
- A six expert national team as members of the collaborative, with similar professional backgrounds.
- A national project team of quality improvement experts.
- A model for good depression care, derived from clinical guidelines and available evidence, developed by the expert team and translated to a set of SMART goals and quality indicators and a stepped care treatment model (Meeuwissen et al., 2004).
- A model for change, the Nolan model, a framework for designing and implementing change following a plan-do-study-act cycle.
- Five national conference days with workshops and discussions.
- A format in Excel for data gathering and data analysis.
- A virtual network environment for exchange of best-practices and online discussions between experts and team members.

Research method

A research team, a collaboration of researchers from the Centre for Quality of Care Radboud University Nijmegen, the psychiatric ward of the Academic Medical Centre in Amsterdam and the Trimbos-Instituut, submitted a research proposal at the Netherlands Organisation for Health Research and Development, ZonMw (www.zonmw.nl). Our proposal was accepted for funding and received €250,000. An additional fund supplied another €100,000.

Our study is an evaluation of an implementation strategy, the Breakthrough Depression Collaborative, as described above. This Collaborative, a large scale implementation program aiming to implement national depression guidelines in primary and specialty mental health care, is considered as the intervention to be evaluated in our study. Coming from different regions and settings in the Netherlands, the teams of the Collaborative improve their practices to reach specific goals, set by a national expert team.

Objectives

The objective of the study was to evaluate the effectiveness, costs and feasibility of the Depression Collaborative as a strategy to implement depression guidelines in mental health. The information generated by this study will serve policy makers and funding bodies who are actually considering to allocate money to quality improvement projects like the Collaborative. By measuring outcomes as well as focusing on the implementation process, the study provides insight in the black box of implementation in multidisciplinary settings. This can result in new ideas about how to improve implementation strategies and processes.

Research questions

With this study we aim to answer the following questions:

1. Does a Depression Breakthrough Collaborative lead to better adherence to guidelines with better outcomes for patients compared to care as usual?
2. Does implementing guidelines with the Breakthrough Method lead to more efficient health care compared to care as usual?
3. What are the implementation activities and experiences of the improvement teams and what barriers and facilitators for successful implementation can be identified?

Design

The design is a quasi-experimental, controlled before and after study including a process evaluation and a cost-effectiveness estimation. The effectiveness of the Breakthrough method as an implementation strategy is assessed by collecting and analyzing data about care processes of participating teams and outcomes at the patient level and by comparing them to outcomes of patients who receive care as usual within a reference group.

Study population and inclusion

The intervention group consists of mental health workers participating in the Depression Collaborative and their patients diagnosed with depression. The

professionals in the intervention group are: general practitioners, psychologists, psychotherapists, psychiatrists, social workers and specialised mental health nurses. Pharmacists may be included in the breakthrough team, as well as staff members involved in quality improvement in their own organisation. In the reference group participating professionals are: general practitioners, psychologists, psychotherapist, psychiatrists, and social workers. All Breakthrough teams, selected to take part in the Collaborative from April until June 2006, were asked to take part in this study. Patients in the intervention group are recruited by the professionals of the improvement teams. Information and informed consent is organised by the research team. Only adult patients diagnosed with depression are included.

The reference group is a subgroup of professionals and patients participating in a different study, the Netherlands Study of Depression and Anxiety (NESDA). This is a cohort study about the long term prognosis of anxiety and depression in different settings and regions in the Netherlands. In proposing a group comparison between Collaborative patients and NESDA controls, more conclusive information about the effectiveness and efficiency of Breakthrough Collaboratives is gathered and generalization to other programs becomes possible. Evaluations without controls can only give suggestive but no conclusive findings and can be misinterpreted by policy makers or media (Øvretveit, 2002).

Data gathering

Data of the intervention group are gathered by the Trimbos-Instituut. Reference group data at the patient level, also covering a large amount of biological data, are collected by the NESDA research team. Outcomes in the intervention group are collected on the professional (a reduction of antidepressants prescription), organisational (reduction of the waiting time) and on the patient level (reduction in depressive symptoms and an improvement in disability status).

Selected measurements instruments are in line with the ongoing NESDA research instruments. Patient outcomes data are gathered at baseline and after one year. The process evaluation is performed within the intervention group only, meaning the teams participating in the Depression Breakthrough Collaborative Process evaluation measurements.

The processes evaluation has two goals. The first is to examine whether the Breakthrough method has been executed as planned, and to what degree the teams have been exposed to the method. The second goal is to show what happens during the implementation and to understand the differences in success or lack of success between the different teams of the Collaborative. Also we want to determine the experiences of

the participants with the Breakthrough method, in order to optimize it for further use in the mental health sector in the Netherlands. Derived from these goals, the main features to be measured are: 1) the exposure of the team members to the Breakthrough method, 2) the implementation strategies performed at the local level, 3) the experiences of participants and 4) the crucial success and fail factors. Measurements consist of an analysis of team documents and reports, an interview with each participating Collaborative team, and a questionnaire to measure the team climate.

Efficiency will be evaluated in an economic evaluation of the intervention group, looking at implementation costs, health care costs and to economic consequences of productivity losses. The design of the economic evaluation is a cost-effectiveness study, which takes a societal perspective and a time horizon of one year. The outcomes measured are patients' depression severity score and patients' disability score at one year. Direct healthcare costs include consultations with various care providers, use of medication, and separate diagnostic tests. These volumes will be extracted from administrative systems and from patient self-reports in a validated questionnaire. Indirect healthcare costs include patients' time and money spent to travel and meet with care providers, which will be estimated on the basis of their mail addresses. Healthcare costs also include implementation costs, that is, health professionals' time spent on meetings or reading papers as part of the improvement strategy. These volumes will be recorded with structured self-report checklists, which are completed by local coordinators. Non-healthcare costs include productivity losses (absence from work) and informal carers' time. All volumes will be valued with average national prices in a chosen index year, using the best available guidance.

Power

It is expected that the Breakthrough strategy will primarily lead to fewer patients in primary care receiving medication for depression (decrease 65% to 50%).

The study is powered to detect these differences with a minimum of 30 teams (15 intervention and 15 controls), each providing at least 25 patients (total n = 625 patients) (alpha = 0.05, beta = 0.20, ICC = 0.05). Estimated number of patients is based on experiences in a previous pilot Collaborative.

Analysis

To answer research question 1, about the impact of the Collaborative on adherence to guidelines and outcomes for patients, a comparison will be made of post-intervention

scores on primary and secondary outcome measures, controlled for known prognostic factors (if available) such as baseline data on these measures. Regression models will be constructed for each outcome measure, which include group (intervention, control), measurement moment (before, after), interaction of group and moment (= effect of breakthrough), known prognostic factors. A random effects (multilevel) approach is used to account for clustering of data within teams. Additional explorative analyses will be performed to explore the influence of other factors on the interventions, such as the factors measured in the process evaluation.

To answer research question 2, whether the Breakthrough Method leads to more efficient health care, the analysis is focused on constructing an incremental cost-effectiveness ratio for the time horizon of one year, using the observed costs and outcomes. Discounting is not applicable. A cost-effectiveness plane will be used to express the random variation, using bootstrap re-sampling if appropriate, and one-way sensitivity analyses are planned to examine the impact of crucial parameters.

To answer research question 3, about the implementation activities and experiences of the improvement teams and the barriers and facilitators for successful implementation, descriptive figures and qualitative data will be provided, referring to the group participating in Breakthrough Collaborative. Data about the exposure of the Breakthrough method will be described. Data about implementation strategies will be compared to effective strategies from the literature (Grol et al., 2005). Data about the experiences of participants and their ideas about crucial success and fail factors will be coded and categorized, structured and interpreted. The interpretations will be compared to existing Collaborative and implementation literature findings, shown to colleague researchers and discussed with the respondents to enhance their validity.

Discussions

The quasi experimental design we have chosen, includes a comparison between patient data gathered from our Collaborative's teams and data coming from a different cohort study. There was no possibility for a true RCT since randomizing quality improvement teams into two groups, one receiving a Collaborative intervention and one who doesn't, is hardly an option. Therefore, as a second best, we chose a design of an intervention group being compared to a control group from a different study. Incorporating a control group in our study had the advantage of being able to draw more robust conclusions about the influence of the Quality Collaborative on the changes in the quality of care and patient outcomes. This type of design is generally considered to be of a better quality than less rigorous methods, such as observational

designs. Also we considered, that in proposing the controlled study design we would have better chances of receiving the funding needed. This expectation is based on the guidelines of the funding body ZonMw and on previous experiences with proposals describing observational study designs.

The disadvantage of our design was that a lot of energy is spent in making sure that both groups are comparable. This implied thorough patient data gathering, with questionnaires and patient interviews that consume a large part of the project's budget. The consequence of this focus on patient data in our research is that less attention could be spent on the gathering of other types of data, like process data and qualitative data about the black box of the quality improvement processes and experiences of those involved in quality improvement. These other types of data are very relevant to quality improvement research in general, where the primary outcomes are changes in professional behaviour and care processes.

Participation in our study was on a voluntary basis. We hoped to include all 18 eligible participating Collaborative teams, but only 10 have agreed to participate. Some of the collaborative teams refused to take part in the research activities, on top of the efforts spent on the routine breakthrough collaborative's. Reasons to refuse study participation were a lack of time for extra data gathering and the fact that the ethics committee overlooking our research demanded informed consent by patients. Inclusion in our evaluation research would demand from the team members to inform their patients about their quality improvement activities and ask them informed consent, something that was not needed by joining the Quality Collaborative in itself. As a consequence of this suboptimal inclusion, our study risks to be underpowered. Moreover, we wonder if the included teams will generate the number of patients needed in the study. We rely on general practitioners to inform their patients of the study and give them the informed consent form.

References

Bijl, R. V. & Ravelli, A., 2000. Current and residual functional disability associated with psychopathology: findings from the Netherlands mental health Survey and Incidence Study (NEMESIS). *Psychological Medicine,* 30, p. 57-668.

Braspenning, J. C. C., Schellevis, F. G. & Grol, R. P. T. M., ed., 2004. *Tweede nationale studie naar ziekten en verrichtingen in de huisartspraktijk. Kwaliteit huisartsenzorg belicht.* Utrecht/Nijmegen: Nivel/WOK.

Franx, G., Spijker, J., Huyser, J. & Doelder, P. de, 2006. Daling in depressie: doorbraakmethode brengt afname in overbehandeling teweeg. *Medisch Contact,* 61, p. 1592-1595.

Grol, R., Wensing, M. & Eccles, M., 2005. Improving patient care: the implementation of change in clinical practice. Amsterdam: Elsevier.

Henkelman, L. & Franx, G., 2004. *De bruikbaarheid van de Doorbraakmethode voor de verspreiding van best-practices in de hulpverlening aan mensen met een depressie: rapportage voorstudie.* Utrecht: Trimbos-instituut.

Landelijke Stuurgroep Multidisciplinaire Richtlijnontwikkeling in de GGZ, 2005. *Multidisciplinaire Richtlijn Depressie: richtlijn voor de diagnostiek en behandeling van volwassen cliënten met een depressie.* Utrecht: Trimbos-instituut.

Marwijk, van H. W. J., et al., 2003. NHG-Standaard Depressieve stoornis eerste herziening. *Huisarts en Wetenschap,* 46 (11), p. 614-623.

Meeuwissen, J. A. C. & Donker, M. C. H., 2004. Minder is meer: stepped care in de geestelijke gezondheidszorg. *Maandblad Geestelijke Volksgezondheid,* 11, p. 2004-2015.

Meijer, S. A., Smit, F., Schoemaker, C. G. & Cuijpers, P., 2006. Gezond verstand: evidence-based preventie van psychische stoornissen. Bilthoven: RIVM.

Øvretveit, J., et al., 2002. Quality collaboratives: lessons from research. *Quality and Safety in Health Care,* 11, p. 345-351.

Øvretveit, J., 2002. *Evaluating health interventions: an introduction to evaluation of health treatments, services, policies and organizational interventions.* Maidenhead: Open University Press.

Ormel, J., Bartel, M., & Nolen, W., 2003. Onderbehandeling bij depressie: oorzaken en aanbevelingen. *Nederlands Tijdschrift voor Geneeskunde,* 147 (21), p. 1005-1009.

Tiemeier, H., et al., 2002. Guideline adherence rates and interprofessional variation in a vignette study of depression. *Quality and Safety in Health Care,* 11, p. 214-218.

Spies, T., Mokkink, H. G. A., De Vries, Robbé P., Grol, R. P. T. M., 2004. Huisarts kiest vaak voor antidepressiva onafhankelijk van de ernst van de depressie. *Huisarts en Wetenschap,* 47, p. 364-367.

Splunteren. P. van, et al., 2003. *Doorbreken met resultaten: verbetering van de patiëntenzorg met de Doorbraakmethode.* Assen: Koninklijke Van Gorcum.

Spijker, J., Bijl, R. V., De Graaf, R. & Nolen, W. A., 2001. Care utilisation and outcome of DSM-III-R major depression in the general population: results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta Psychiatrica Scandinavia,* 104, p. 19-24.

Spijker, J., et al., 2003. Psychiatrische morbiditeit onder volwassenen in Nederland: het NEMESIS-onderzoek. IV: duur van depressieve episoden. *Nederlands Tijdschrift voor Geneeskunde,* 147 (41), p. 2025-2028.

Evaluating the Implementation and Effects of a Multilevel Quality Collaborative in Hospital Care

MICHEL DÜCKERS*
CORDULA WAGNER*
PETER GROENEWEGEN*

ABSTRACT

Phase of the research when this was written: In the middle of the process. Before the data collection of the final part of the study, but after reviewing the literature and collection and analysis first data.
Main research question: How does the participation by hospitals in a multilevel quality collaborative (MQC) result in enhanced quality of care and the development of an organizational infrastructure for improvement, stimulating the adoption and sustainable spread of best practices?
General design and sample: Monitoring of process and effects of implementation and spread of breakthrough projects by multidisciplinary hospital teams joining six quality improvement collaboratives (QIC's). Data are gathered among actors within hospitals (e.g. executives, programme coordinators, managers, doctors and project leaders of improvement teams) as well as external change agents responsible for supporting hospital actors. Data from a national survey are used to compare the state and effects of the improvement infrastructure in the intervention group (the MQC group) and a control group (non-MQC).
Main findings (expected): The study will clarify how leadership and support by the parties involved relate to the implementation, spread and sustainability of innovations within hospitals. Furthermore, it will illustrate how the intended improvement infrastructure connects national health care targets to care delivery by individual professionals.
Methodological challenges: 1) Analysis require a certain amount of statistical power. The limited number of hospitals and project teams combined with self report bias (too positive) and non response reduces the sample size and

* NIVEL — Netherlands Institute for Health Services Research, Utrecht, The Netherlands (m.duckers@nivel.nl).

availability and reliability of necessary data. 2) Developing an instrument to monitor the degree to which relevant conditions are met during the implementation of breakthrough projects in a QIC. 3) Conceptualization of the emerging organizational improvement infrastructure is needed to formulate an evaluation strategy in which intervention and control group are compared. 4) Controlling for hospital characteristics while comparing, all the more because the 24 MQC hospitals are not randomly selected. 5) Aggregating data at unit (micro) level to institutional (meso) level.

Practical challenges: Many actors and conditions play a role in implementing the MQC. The setting of the evaluation study is highly politicized. Evaluation inevitably means influencing the implementation, a process based on communicating success stories, as it also generates less positive insights.

Main lessons for other researchers: Besides general advice, a number of tips and recommendations is listed for conducting an independent evaluation study in a political environment.

Evaluating the implementation and effects of a multilevel quality collaborative in hospital care

> *"Conformity can be costly in a world of uncertainty which requires innovative institutional creation because no one can know the right path to survival."*
>
> DOUGLASS C. NORTH

A programme to improve hospital care and stimulate organizational development

This chapter was written as the research reached the stage where it was halfway completed. Our study involves an evaluation of a sector wide quality improvement and dissemination programme for hospitals in the Netherlands. The *Better Faster* programme is designed to stimulate quality improvement, the spread of breakthrough projects and systematic performance management within participating hospitals (view Box 1 for additional information on the context where it is implemented).

We will start our contribution by paying attention to the structure and other features of the improvement programme. Next, the chapter addresses the research objective and questions. Also it covers the study design, data collection, analyses and some of the challenges we were (and are) confronted with, both methodological and practical. Finally, based on our experiences so far, we will summarize some main lessons for other researchers.

BOX 1

CHARACTERISTICS OF THE DUTCH HOSPITAL SECTOR

Categories of hospitals: Dutch hospitals can be divided into categories. *General hospitals* concentrate on treatment, nursing and the education of doctors and nurses.

Top clinical hospitals: also provide medical training and highly specialized care (e.g. heart and neurosurgery, IVF) that requires expensive and specialized instruments.

University hospitals: deliver patient care, conduct scientific research and education for medical faculties and develop new medical technologies and techniques.

Number of hospitals: In 2006 there were 141 hospital locations, organized in 93 hospital organizations (mostly general, about 20 top clinical and eight university hospitals) (Roeding, 2006).

Trends: The number of hospitals has decreased in the last few years. In response to merging insurance companies and (regulated) competition with other hospitals and service providers, it is expected that hospitals are going to improve their bargaining strengths (scale extension) and quality of care delivery.

Number of medical specialists: 14,283 (Capacity Plan, 2005).

Partnerships Most medical specialists work in a hospital setting. The majority is self-employed, working with other specialists of the same speciality in a partnership.

Medical specialist and hospital; an integrated enterprise: Medical specialists and 'their' hospitals operate as one entity. This follows from the *Integration Act.* The law has integrated the claims of insured parties for specialist care and hospitals services.

Study subject: a multilevel quality collaborative

Better Faster was introduced in November 2003. The programme is a logical next step in the implementation of quality management systems in Dutch health care institutions. Since 1996 health care organizations are bound by law to provide effective, efficient and patient oriented care (Care Institutions Quality Act, 1996). Despite the obligation to develop a quality system to improve and assure the quality of care, not enough progress was made with the construction of quality management

BOX 2

RECOMMENDATIONS FOR THE DEVELOPMENT OF QUALITY MANAGEMENT SYSTEMS

In 2000 the Netherlands Health Care Inspectorate stated that more transparency was necessary to improve the external accountability concerning quality of care, and the internal management of health care processes. Four recommendations were made:

- those (parts of the) sectors that are behind on schedule should develop quality systems by using the experience and knowledge of others;
- coordination between different quality systems is important;
- additional influence must be allocated to patients and insurance companies;
- more attention for illuminating the risks the quality system tries to tackle and the outcomes of care delivery is desirable.

systems (Sluijs et al., 2000) and the spread of knowledge of best practices (Box 2 gives an overview of the recommendations).

Subsequently, the Ministry of Health launched an improvement programme, resting on three pillars:

1. Creating awareness by having authoritative experts from other fields of service delivery communicate appealing points of reference on relevant themes (e.g. safety, logistics and accountability).
2. Construction of a national set of performance indicators for safer and better hospital care.
3. A national action programme to stimulate transparency, efficiency and quality of care by implementing breakthrough projects in a selected group of hospitals.

Mission and time path of the third pillar programme

Although Better Faster is a mix of the three pillars, we are focusing primarily on the third one. The Ministry requested a number of parties, afterwards assembled in a consortium, to design and implement the third pillar. Their efforts resulted in a programme with the following mission: *Realizing a substantial and appealing*

performance improvement in 20% of Dutch hospitals on the areas of patient logistics and patient safety. Simultaneously, a 'flywheel' is established within participating hospitals, aimed at internal spread of results and newly developed competencies (Schellekens et al., 2003).

Each participating hospital implements two series of breakthrough projects in order to start the 'flywheel'. This flywheel is a metaphor for the infrastructure enabling further improvement and dissemination. Figure 1 shows the time path of the programme. 24 hospitals are divided in three groups of equal size. Each group joins the programme for a period of two years.

Figure 1
The time path of the third pillar programme

2004	2005	2006	2007	2008
Group 1:	Series I	Series II		
Group 2:		Series I	Series II	
Group 3:			Series I	Series II

General features of breakthrough collaboratives

The blueprint of the third pillar resembles what is known in the literature as a quality improvement collaborative (QIC) (Leatherman, 2002; Mittman, 2004). A collaborative brings together groups of practitioners from different healthcare organizations to work in a structured way to improve one aspect of the quality of their service. It involves them in a series of meetings to learn about best practices in the area chosen, about quality methods and change ideas, and to share their experiences of making changes in their own local setting. (Øvretveit et al., 2002). To improve performance, medical professionals use 'breakthrough' methods. The implementation is based on repeated application of the Nolan model. Professionals run improvement cycles (plan-do-study-act) and answer three questions: 1) what are we trying to accomplish?, 2) how will we know that a change is an improvement? and 3) what change can we make that will result in an improvement?' (Berwick, 2003; Langley et al., 1996).

The implementation is supported by external change agents. These individuals — change experts and experienced consultants — connect the innovations to the receptive context of the adopting hospital units. They organize a series of meetings where hospital teams receive instructions. The teams are responsible for implementing/ reinventing the innovation so that it fits their needs. They are supposed to systematically measure their outcomes, test several interventions, work under time

pressure and compete with other teams (Berwick, 1998; Øvretveit et al., 2002; Van Splunteren et al., 2003).

The third pillar approach

Seven of the third pillar programme targets are covered by breakthrough projects. Table I includes the programme targets and duration of the six breakthrough projects. In each programme hospital two series of projects are implemented by multidisciplinary project teams.

Table I
Breakthrough projects per quality area and their programme targets

QUALITY AREA	BREAKTHROUGH PROJECT	PROGRAMME TARGETS	PROJECT DURATION
Patient logistics	Working without waiting lists (WWW)	1. Access time for clinical consultation is less than a week	One year
	Operating theatre (OT)	2. Increasing the productivity of operating theatres by 30%	Two years
	Process redesign (PRD)	3. Decreasing the total duration of diagnostics and treatment by 40-90%	Two years
		4. Reducing length of stay by 30%	Two years
Patient safety	Medication safety (MS)	1. Decreasing the number of medication errors by 50%	One year
	Pressure ulcers (PW)	2. The percentage of pressure ulcers is lower than 5%	One year
	Postoperative wound infections (POWI)	3. Decreasing postoperative wound infections by 50%	Two years

In the first year the teams run plan-do-study-act-cycles (PDSA) to reduce infections, pressure ulcers, access time for clinical consultation et cetera. Series II is about coordinated spread of the pilots of the first year (working methods, competencies and results) over new units. Both series of improvement projects are required for activating the flywheel. More concrete: while participating in the programme and implementing projects the hospitals are expected to develop an infrastructure with indicators, accountability and feedback loops that will help them to start similar trajectories in the future, easier and without programme support. Hospitals have to provide structures, procedures and facilities that raise and maintain systems enabling the organization to control the quality of processes and outcomes.

The collaborative approach, nevertheless, stretches beyond the breakthrough topics. Besides a fourth safety target — implementing a system of blame free reporting — the collaborating hospitals commit themselves to substantially upgrade the position of patients in care delivery processes (patient participation) and the professional quality of medical specialists. These targets affect the safety culture and learning climate in the hospitals, the professionalism of the people working there, and the needs and experiences of patients.

Why is it a multilevel programme?

According to Ferlie & Shortell (2001) a single level approach for implementing improvement projects has its shortcomings. Instead, one should pay attention to the entire organization as a multilayered system. The consortium seems aware of this notion and constructed the third pillar as a combined bottom up-top down approach by adding a 'leadership network'. We have seen how the time path of each group of hospitals begins with breakthrough QIC's. Moreover, strategic managers of every hospital in each group take part in a recurring executive network, part of the sub programme Leadership & Organizational Development (L&O). L&O encompasses a collaborative platform where hospital directors can share knowledge and experiences related to change processes in their institutions. The sub programme has three targets, focussing on the flywheel:

- creating an infrastructure for improvement;
- performance management as a normal part of strategy and administration;
- leadership in innovation and improvement.

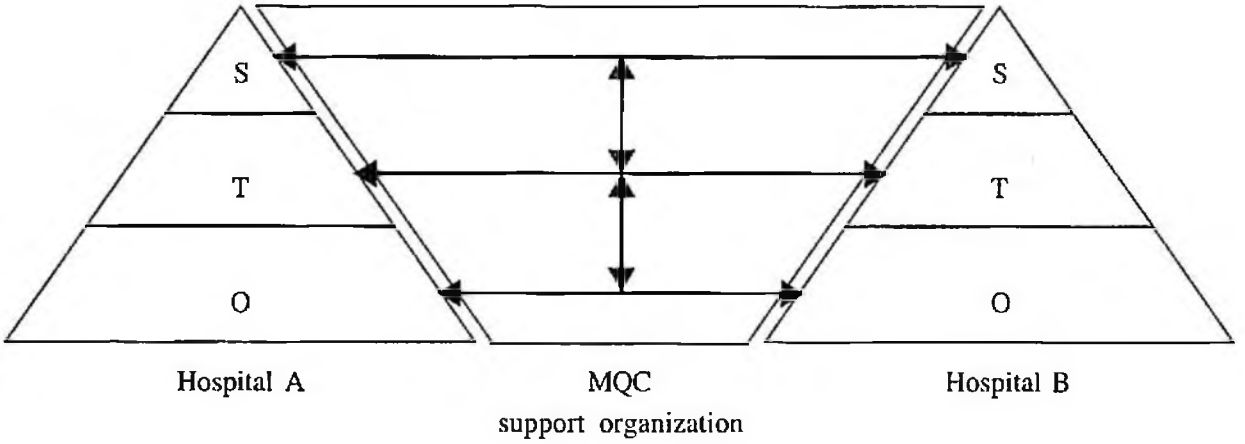
Taking this into account, we consider it appropriate to call the third pillar programme a multilevel quality collaborative (MQC). L&O serves as an umbrella thanks to:

- the explicit goal of building an infrastructure for improvement by the integrated implementation and spread of interventions, competencies and results;
- the presence of techniques and interventions to be implemented at every organizational level (Consortium, 2004).

In Figure 2 the MQC support organization (the Consortium's external change agency) is positioned between the hospitals. The support organization (the grey

shape) is connected to actors in the hospitals at strategic (S), tactical (T) and operational level (O). For each group of hospitals the consortium has two years to exchange knowledge and information needed for running both series. Thus, the MQC serves as a temporary network of horizontal and vertical communication lines. A temporary network with the purpose of stimulating organizational development and the broad implementation of evidence based medicine and effective quality improvement methods.

Figure 2
The MQC as an integrated approach for upgrading hospital care



The MQC-implementation, takes place in a complex environment comprised by exogenous conditions. It includes:

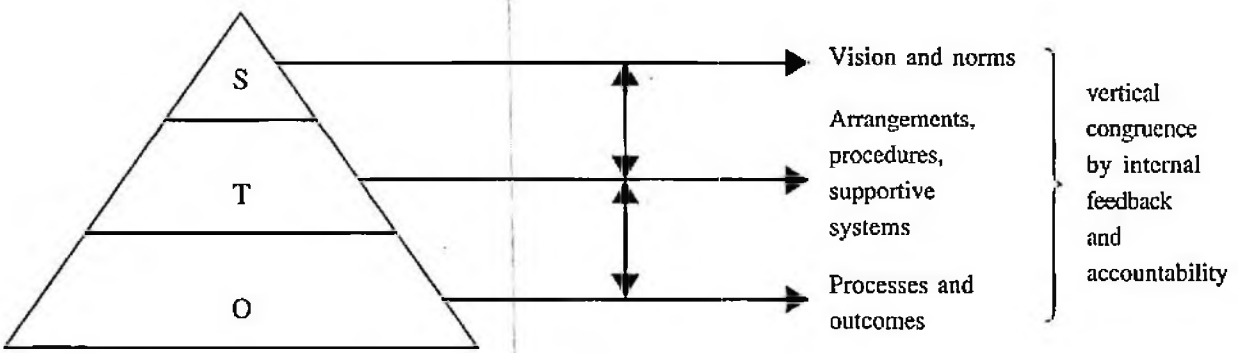
- two other Better Faster pillars (promising best practices and awareness by transparency);
- gradual introduction of an institutional coordination model of regulated competition (the quasi market; Bartlett & Le Grand, 1993);
- a change in the financial system bringing standardized output pricing;
- new roles for insurance companies and patients or their representatives in demanding high quality and low prices.

Why the improvement infrastructure is a quality control mechanism

The policy makers and the consortium assume that Better Faster and the element of competition generate an effective mechanism for channelling the behaviour of actors in the Dutch hospital sector at a macro, meso and micro level.

At *macro* level, the competition between hospitals is fed by the discrepancy between feasible transparent norms and the performance of a specific organization or organizational unit. This 'quality chasm' is to be crossed by applying working methods, procedures, techniques and conditions that seem promising in achieving the norms (i.e. the breakthrough projects). The national strategy can be summarized as: by creating competition between hospitals (quasi-markets) and formulating feasible norms (quality parameters and standardized prices), the hospitals are triggered, even pressured, to be fully aware of the status quo at meso or micro level (the chasm) and will do their best in optimizing processes and outcomes (creating or adopting new methods, interventions etc.). The implicit assumption of policy makers is that the presence of the right institutional conditions at macro level will guide the behaviour and choices of directors and top management of hospitals (*meso level*) and professionals working in the hospital units (*micro level*). The three L&O targets — infrastructure for improvement, performance management and leadership — aim at congruence between vision (S), supportive measures (T) and implementation (O) by systematic feedback between the levels (Figure 3).

Figure 3
Vertical congruence



Consequently, by developing such a performance management based improvement infrastructure, in fact a quality control mechanism is initiated, resembling Wagner's description of a quality management system at its highest stage (Box 3). The system that helps the programme hospitals to implement new innovations in a similar way, also facilitates the translation of national norms to practical changes in processes and outcomes on the floor. Where the MQC is a temporary structure, the organizational improvement infrastructure annex control mechanism is meant as a lasting configuration. It serves as a permanent macro-micro bridge.

BOX 3

THE HIGHEST DEVELOPMENTAL STAGE OF QUALITY MANAGEMENT

Wagner et al. (2006) divide quality management activities into four stages of development. Stage 0 is the lowest stage. Organizations of this category are characterized by: presence of a mission statement and annual quality report, encouragement of professional development through HRM, practical guidelines for medical treatment, quality improvement by peer review and care plans. Patient involvement is low. In stage 1 and 2 there is an increase in quality management activities. Stage 3 is characterized by: availability of a quality action plan and quality manual, training based on quality policy and systematic feedback, practical guidelines for the routing of patients and critical incidents, internal audit, satisfaction research and participation of patients in committees and improvement projects.

Study objective and research questions

As soon as the quotation at the beginning of the chapter is interpreted in the light of the adoption and spread of innovations, it contains an ambiguous, even paradoxical warning. To become successful the MQC-programme requires a substantial amount of medical professionals or teams to adopt promising working methods to achieve logistical and safety targets. The contradiction is that the programme strives at generating innovation and learning processes by encouraging conformity to a standard approach.

Indeed, the policy makers and programme architects expect this standard — based on improvement cycles (breakthrough methods) and performance management in an integrated organizational quality model — to be 'the right path'. Nonetheless, until today the evidence on the effects of breakthrough series or QIC's is rather limited (Landon et al., 2004; Mittman, 2004; Øvretveit & Gustafson, 2002; Øvretveit, 2003). And, to complicate matters, the evidence on the effectiveness of the envisioned improvement infrastructure also has its lacunas. There is, for example, limited evidence on the performance effects of applying integrated models (like Malcolm Baldrige Quality Award and European Foundation for Quality Management) (Minkman et al., 2007).

Our goal is to explore what lessons can be learned from the hospital experiment in the Netherlands. The research strategy outlined in this chapter aims

at describing and explaining the implementation and effects in a sector that, as we have seen, is characterized by important institutional changes. The main research question is: How does the participation by hospitals in a multilevel quality collaborative result in enhanced quality of care and the development of an organizational infrastructure for improvement, stimulating the adoption and sustainable spread of best practices?

With the purpose of examining the MQC-implementation, its effects on the quality of care, and the merits of the improvement infrastructure as a means to an end for hospital governance, we translated the main research question into five sub questions:

1. How does the design of the MQC, as presented in the vision documents and action plans of the consortium, correspond to the determinants of success known from literature?
2. To what extent is the success of the implementation influenced by the availability of the required conditions during the implementation process of the breakthrough QIC's?
3. What is the operational state of the improvement infrastructure and the spread of the QIC-projects in MQC-hospitals by the end of the second year and how can this state be explained?
4. How did the improvement infrastructure develop within Dutch hospitals between the beginning and end of the programme, and does the development differ between MQC and non-MQC hospitals?
5. What can we say about the effectiveness of improvement infrastructure and projects at micro and meso level and what is the relation between perceived and actual quality effects at both levels?

General design and sample

According to Greenhalgh et al., 2005 applied science into the process of dissemination, implementation and routinisation should be:

- *theory driven:* it should aim to explore an explicit hypothesized link between determinants of a particular problem, the specific mechanism of the programme and the expected changes in the original situation;
- *process — rather than 'package' — oriented:* best practice is a process, not an intervention package. Research questions should be framed with a view to illuminating this process e.g. 'what features or conditions account for the

success of project X in this context and the failure of a comparable project in a different context?';

- *participatory:* in process evaluations not the researcher but the practitioner frames the problem, makes the manipulations and interprets the data, while the researcher observes;
- *collaborative and coordinated:* it should aim to prioritise and study key research questions in a variety of settings, rather than small isolated teams 'doing their own thing'. In this way, the impact of place, setting and context can be systematically studied;
- *addressed using common definitions, measures and tools:* it should adopt standardized approaches to measuring key variables to enable valid comparisons across studies;
- *multidisciplinary and multi-method:* it should recognise the inherent limitations of experimental approaches for researching open systems, and embrace a broad range of research methods with the emphasis on interpretive approaches;
- *meticulously detailed:* it should document extensively the unique aspects of different programmes and their respective contexts and settings to allow for meaningful comparisons across programmes (to interpret idiosyncratic findings and test rival hypothesis about mechanisms);
- *ecological:* it should recognise the critical reciprocal interaction between the programme that is the explicit focus of research and the wider setting in which the programme takes place. The latter provides a dynamic, shifting baseline against which programme related activities occur.

When designing the study, we tried to keep these recommended characteristics in mind. Hence, evaluating the programme means: testing theory by applying a variety of methods — qualitative and quantitative — to illuminate the implementation process and success, and (exogenous) influences of distinctive features of the Dutch hospital setting. Furthermore, testing needs to be done in such a way that the study gives enough space to field actors to alter the implementation course without jeopardizing the possibilities to replicate the study by other researchers.

Eventually we came up with a design in which a) the implementation processes and effects of the six breakthrough QIC's are monitored; and b) attention is given to the development and functioning of the improvement infrastructure.

Part a) is carried out following a non experimental design. Unfortunately there is no possibility to compare implementation processes between an intervention group and a comparison group. The problem is that our subject — implementation and conditions — does not exist outside the intervention group. An extra complication is that the availability of effect measures at micro level also are a

result of the intervention. Outcome data on the achievement of programme targets (Table 1) can only be obtained as evaluation input, when the teams have monitored their project progress using performance indicators (intervention) for quite a while.

The sample size of part a is restricted by the total number of hospitals and teams receiving support from the external change agency during the first series (approximately ten teams per hospital in each series).

For part b, the organizational infrastructure development path, we can use a quasi-experimental design. Our strategy is to compare system characteristics of the intervention group to the situation in contrasting groups comprised of non-MQC hospitals. In the Netherlands there are about 100 hospitals. 24 of them participate in the programme, the others can be utilized for comparing purposes.

Data collection and analysis methods chosen to answer the sub questions

The *first* sub question addresses the relation between the causal assumptions of the MQC-policy theory and determinants known from the literature. We consider an answer to this question desirable as it a priori helps to identify the strengths and weaknesses of the programme and provides a touchstone for the rest of the study. Further data collection is shaped by the results of this ex ante evaluation.

To answer the *second* sub question a measuring instrument is needed. We must measure the extent to which conditions are met during the implementation of the breakthrough collaboratives. Since we had no knowledge of an instrument suitable for measuring the conditions for successful implementation of breakthrough QIC's (sub question two), we included a procedure in our research strategy to develop and test a new instrument, to be filled out by the project leaders of the multidisciplinary teams. This COPI-QIC (COnditions for Project based Implementation by multidisciplinary teams in a Quality Improvement Collaborative) includes questions about leadership of the strategic management, support provided by external change agents and hospital organization, the project's value, its complexity, team organization and other items. When filled out by the project leaders of two groups of eight hospitals (N = 148) at the end of the first series, data are analysed by running factor and reliability analyses. Additionally, a combination of interviews with hospital executives and programme coordinators of each hospital, questionnaires filled out by consultants of the consortium and the database of the consortium — with project indicator outcomes — is used to analyse QIC-implementation in programme hospitals. By using the COPI-QIC questionnaire in 16 hospitals, we expect to collect enough data for applying structural equation

modelling or multilevel analyses on teams of 16 hospitals. This should enable us to learn more on the relation between conditions at team and hospital level with respect to outcome data (performance indicators) and perceived effects on a number of quality dimensions (e.g. satisfaction of staff and patients, costs, productivity, staff motivation, clarity on the division of tasks).

Sub question number *three* covers the organizational infrastructure for improvement. In a previous section we pointed out that the hospitals will have to come up with structures, procedures and facilities that raise and maintain a system enabling the organization to control quality of processes and outcomes. Besides describing the system and its outcomes ('what happened'), we are looking for explanations ('how did it happen'). The descriptive component focuses at the strategies and choices of the actors within, the presence of clear norms, instruments, support and procedures, along with the actual use of output data for feedback, problem solving and learning purposes. These issues are examined by a combination of quantitative and qualitative techniques. Data are collected from the programme coordinators of each hospital, managers and medical specialists (interviews and questionnaires). Based on collected data the state of the improvement infrastructures and the contribution of leadership is modelled (using its functions for performance management; Leggatt & Dwyer, 2003). In our search for explanations we emphasize the role of leadership (and leadership climate; view Chen & Bliese, 2002). Leadership is presented as a condition sine qua non by the consortium for the success of the MQC-programme (Consortium, 2004).

Sub question *four* urges us to place our descriptive 'question three material' in a more historical perspective. In theory we can track the historical development of the improvement infrastructure in Dutch hospitals from 1995 until now. Longitudinal survey data on the quality management in hospitals is available from measurements in 1995, 2000, 2005 and (planned for) 2008 (Sluijs et al., 2007, also see Wagner et al., 1999). By exploring and connecting different datasets we can learn more on the quality management of Dutch hospitals since 1995. Depending on the number of matches we have an opportunity to test individual hospitals or cross sectional groups (MQC or not) whether:

- the intended progress in their quality management system did occur;
- changes took place because of their participation in the MQC or something else;
- (elements of) the improvement infrastructure is (are) positively related to quality aspects.

This bring us to our *final* sub question. Again, we are looking for a pattern in the relation between conditions at team level and the system at hospital level on the one

hand, outcome data (performance indicators) and a number of perceived quality effects. A variety of data sources is at our disposal. At hospital (meso) level we plan to analyze relations between the self reported quality aspects and the state of the improvement infrastructure. Additionally, to investigate relations between the state of the improvement infrastructure and outcome data from the national set of performance indicators (second pillar).

A similar exercise awaits us at unit (micro) level when exploring relations between conditions for successful implementation and quality effects. We can use self reported quality aspects identical to the ones measured at hospital level to start with. Secondly, the Consortium's database contains project outcomes of each MQC-hospital. As far as we can tell, combining the database and the COPI-QIC measurements may resolve in new insights concerning the congruence between perceived and actual quality effects of MQC-participation.

An additional question

One of the things we did so far is conceptualize the MQC as an intervention to establish an organizational improvement infrastructure/quality control mechanism at hospital level. This control mechanism serves as a macro-micro bridge: an institutional mechanism to canalize the behaviour of individual actors and bring it in harmony with the expectations of health care authorities, interest groups and representative organizations. We also stated that its intended positive effects have to do with quality goals and organizational development. Up until now we did not discuss possible negative unintended effects. At least one relevant and theory driven aspect of our study object remained unaddressed.[1] Berwick (2003) agrees that current strategies for developing the healthcare workforce are based on outmoded theories of control and standardisation of work. The MQC and the control

[1] Two other aspects are unmistakeably relevant, but cannot be answered based on the study material. The first one is a potential *conflict between allocation models*. The pillars of Better Faster are brought as a stimulus for hospital actors to share knowledge of best practice. One may wonder how knowledge sharing works in a setting with increased competition (quasi-markets) on quality parameters. Cooperation depends on trust, competition threatens this trust. Without intermediation there is a chance that only one (or none) of the motives of the models will be realized (Svensson et al., 2005). Secondly, there is a risk of a *performance paradox*. In reality the organizations that are effective in measuring performance indicators are not automatically the most effective organizations (Van Thiel & Leeuw, 2003). Performance auditing is likely to result in strategic behaviour (window dressing, fraud) and decreased reliability and validity of judgement. Other effects are bureaucratization, tunnel vision, sub optimization of processes and outcomes, as well as isomorphism.

mechanism fit within this pattern. Weggeman (1992) stresses, nonetheless, that professionals cannot be managed by rules, procedures and information systems. Professionals resist the standardization of their skills (their division into simply executed steps) because that makes them programmable by the 'techno structure' of the organization; it destroys the basis of their *autonomy* and drives the structure to the machine bureaucratic form (Mintzberg, 1979). The methods for answering sub question four should give us ammunition to test the hypothesis that the L&O-path leads to a decrease in the discretionary space of nurses and doctors. We should, for example, be able to study changes in the presence of management control instruments. An increase of management control automatically means a decrease of staff autonomy.

Methodological challenges

In this section we will discuss some of the challenges we are confronted with in this study. A major methodological challenge has been (and remains) our *statistical power*. Assessing relations between structure, processes and outcomes requires analyses that depend on the availability of sufficient data. This is the case with most of our analyses, irrespective of whether these involve confirmatory or explanatory factor analyses, logistic regression, multilevel or structural equation modelling. We can only run them if we have enough statistical power. In practice we have seen that our limited sample size is reduced by non responding project leaders, managers and medical specialists. Furthermore, the first year database of the first eight hospitals was filled with monitoring data from only half the project teams (Dückers et al., 2006). Literature shows that this is not an uncommon phenomenon in QIC data collection (Cretin et al., 2004). However, combined with non response to questionnaires, this complicates the matching of project results from the database to process variables from the teams gathered via the COPI-QIC. Moreover, tracking individual hospitals in the longitudinal hospital survey data is jeopardized also by the sum of non responses over the years and hospital mergers. The group gets smaller at every next measuring moment. Still, our goal is to learn more on the relations between the dependent variables (perceived and actual successes or quality aspects at meso and micro level) and the independent variables (QIC-implementation conditions and improvement infrastructure components) of the dissemination programme. Our challenge is that — besides regular problems concerning self reported data, (overestimation and overrepresentation of success) — the study sample gets filtered more and more as the process continues and cases fall out. All we can do is making the measuring instruments as user-friendly as possible. The next phase is to explore

the collected data thoroughly, look for promising patterns and experiment in extracting the most reliable and valuable lessons.

A second methodological challenge is *measuring the conditions for successful implementation* by improvement teams in QIC's. For this purpose the COPI-QIC was developed and tested, fortunately, with success. The testing procedure did cost an extra year because after the first series of the first group we only had 54 proper cases. By adding data from the second group, a year later, the sample size grew to 101 cases.

The third challenge had to do with *conceptualizing the improvement infrastructure*. Despite the variety of development ambitions and targets, the programme makers never defined the final organizational features in detail. We pointed out how L&O focuses at creating an infrastructure for improvement, founded on performance management and supportive leadership. It is mentioned more than once that, ultimately, the organization should be capable of adopting, spreading and sustaining breakthrough methods and results. Therefore, the challenge is to come up with some sort of explanatory model. Our current model is based on determinants found in the *ex ante* evaluation (first sub question) and a qualitative study among the programme coordinators of the first group (N = 8) at the end of the second year. By assessing the approach followed by the hospitals in the two years in which they participated an answer could be given to the question how hospitals deal with issues of internal spread and sustainability (Dückers & Wagner, 2007).

Our fourth challenge has to do with the quasi-experimental hospital study. Comparisons ask for *reference groups*. An extra difficulty we have to take into account, is that the programme hospitals are not randomly selected by the consortium. They went through an intake procedure in which the consortium investigated the candidates' readiness. Luckily, survey data allows us to control for the size (number of employees, beds or adherence) and type of organizations (general, top clinical or university hospitals; box 1). Furthermore, we are capable of controlling for the different (integrated) quality models and certificates. This has advantages and disadvantages. We need to explore the occurrence of these other models and certificates in our data. Next, after studying their dimensions and the areas they affect, it should be possible to distribute hospitals over different comparison groups.

Finally, a fifth methodological challenge plays a role in our study, as well as on a longer term. There is a *distance between mechanisms and outcomes*. Better Faster aims at internal spread of methods and results of effective pilots. So if everything goes according to plan, eventually, the programme hospitals will sort the intended results on quality and safety throughout the organization. Nonetheless, the matter is that according to research hospitals participating in quality improvement

programmes are 'not more likely to show improvement on quality indicators than hospitals that do not participate' (Snyder & Anderson, 2005). In general, when translating process and outcome data of micro events to an aggregated institutional level, the contribution of individual projects gets lost in translation. To us, this means that we must be cautious in interpreting meso phenomena as consequences of micro improvement actions. Moreover, there is a time lag between 1) the start of the programme, 2) the future point where hospital wide spread is realized and 3) the moment when macro datasets on hospital performance (the second pillar) are available. It will take years before the indicators scores of MQC and non-MQC hospitals can be examined at national level.

Practical challenges

In a practical sense the MQC-evaluation demands from us as researchers that we keep track of the overall programme progress e.g. all the relevant changes in the approach and organization of the consortium and other actors. Nonetheless, the real challenge has not primarily to do with time consuming activities such as reading status reports, joining meetings and coordinating research activities. The real challenge lies in the communication of our findings. In a programme where success and failure depends on the acts and neglects of the implementing actors, 'naming' can easily result in 'blaming'. Learning requires psychological safety and constructive feedback, while blaming brings the opposite. This is one side of the coin. The other side is more fundamental. The setting of the evaluation study is highly politicized. By our independent position we are bound to serve a public interest and, thus, to open communication of our study results. Any media attention attracted by positive or negative findings influences the implementation and spread within the hospitals. To maximize the chances of success, the consortium is given the task to spread promising success stories (Dückers et al., 2005). Our dilemma is that we may potentially interfere in the implementation process, that we influence the planned dissemination and that our professional (participatory!) relation with the involved parties gets affected by it. Many actors and conditions play a role in the MQC-implementation. High expectations are raised, many interests are at stake.

Lessons for other researchers

The research strategy presented in this chapter is a product of several choices. Some of the choices made are inherent to typical characteristics of the Dutch hospital

sector. In the strategy outlined, a focus is placed on the implementation and quality effects of the programme. The lessons from this study are of potential value to politicians and policy makers. Looking back, we can also summarize some lessons for colleagues who conduct similar studies in present or future. The first lesson is a general one: take the recommendations of Greenhalgh et al., into account when determining the final study design. The others are helpful when operating in a political arena:

— illuminate the assumptions behind the policy and programme theory and use the literature to check (postulated) mechanisms for flaws;
— study all relevant actors, their interests, tasks and responsibilities;
— make clear agreements with contractors and stakeholders on your independence and the shape and topics of your publications;
— analyses often lead to general conclusions, try to be as specific as possible: 'what works under which conditions';
— when reporting results: be honest, transparent and at the same time constructive (your aim is to promote learning, not blaming);
— and last but not least: make sure that your work can withstand the unavoidable scrutiny; when your findings are unwelcome, your methods will be the first victim.

Acknowledgements

This study was funded by ZonMw, the Netherlands Organization for Health Research and Development.

References I found most helpful and inspiring

Greenhalgh, T., et al., 2005. *Diffusion of innovations in health service organizations: a systematic literature review.* London: Blackwell Publishing.

Leggatt, S. & Dwyer, J., 2003. Factors supporting high performance in health care organizations. Melbourne: National Institute of Clinical Studies.

Øvretveit, J., et al., 2002. Quality collaboratives: lessons from research. *Quality and Safety in Health Care,* 11 (4), p. 345-351.

Wagner, C., Gulácsi, L., Takacs, E. & Outinen, M., 2006. The implementation of quality management systems in hospitals: a comparison between three countries. *BMC Health Services Research,* 6, p. 50.

Appendix listing of resources which would help other Q&SI researchers

Bartlett, W. & LeGrand, J., 1993. *Quasi markets and social policy.* London: Macmillan.

Berwick, D. M., 1998. Developing and testing changes in delivery of care. *Annals of Internal Medicine,* 128, p. 651-656.

Berwick, D.M., 2003. Improvement, trust, and the healthcare workforce. *Quality and Safety in Health Care,* 12 (6), p. 448-452.

Bloemen, Y. H. P., 2006. Do health insurers stimulate the spread of innovations between hospitals?: a study of the attitude and role of insurance companies in the new Dutch health care system. Tilburg: University of Tilburg (master thesis) (in Dutch).

Chen, G., Bliese, P. D., 2002. The role of different levels of leadership in predicting self and collective efficacy: evidence for discontinuity. *Journal of Applied Psychology,* 87, p. 549-556.

Conference on Quality of Care, Leidschendam, 2000. The Hague, The Netherlands: Health Care Inspectorate (in Dutch).

Consortium Sneller Beter pijler 3, 2004. National Strategy Better Faster Pillar 3. Utrecht/Rotterdam: Consortium Sneller Beter (in Dutch).

Cretin, S., Shortell, S. M. & Keeler, E. B., 2004. An evaluation of collaborative interventions to improve chronic illness care: framework and study design. *Evaluation Review*, 28 (1), p. 28-51.

Dückers, M. L. A., Wagner, C. & Groenewegen, P. P., 2005. Determinants of a knowledge based sector wide improvement programme in Dutch hospital care. *Acta Hospitalia*, 45 (3), p. 37-54.

Dückers, M., De Bruijn, M. & Wagner, C., 2006. Evaluation Better Faster Pillar 3: implementation of improvement projects in the first hospital group. Utrecht: NIVEL (in Dutch).

Dückers, M. & Wagner, C., 2007. Evaluation Better Faster Pillar 3: the internal flywheel; the situation in the first hospital group after two years of programme participation. Utrecht: NIVEL (in Dutch).

Ferlie, E. B., Shortell, S. M., 2001. Improving the quality of health care in the United Kingdom and the United States: a framework for change. *The Milbank Quarterly,* 79 (2), p. 281-315.

Harteloh, P., Casparie, A., 1994. *Quality of care.* In A health care approach to an industrial approach. Maarssen: Elsevier/De Tijdstroom (in Dutch).

Hoogervorst, J. F., 2003. *Better faster.* Letter send to the House of Representatives of the Dutch Parliament in November 2003 by Minister Hoogervorst. Reference no. CZ/B-2427994 (in Dutch).

Landon, B. E., et al., 2004. Effects of a quality improvement collaborative on the outcome of care of patients with HIV infection: the EQHIV Study. *Annals of Internal Medicine*, 140, p. 887-896.

Langley, G. L., et al., 1996. *The improvement guide: a practical approach to enhancing organizational performance.* San Francisco: Jossey-Bass.

Leatherman, S., 2002. Optimizing quality collaboratives. *Quality and Safety in Health Care,* 11, p. 307.

Minkman, A., Ahaus, K. & Huijsman, R., 2007. Performance improvement based on integrated quality management models: what evidence do we have?: a systematic literature review. *International Journal for Quality in Health Care* 19, p. 90-104.

Mintzberg, H., 1979. *The structuring of organizations: a synthesis of the research.* New York: Prentice-Hall.

Mittman, B. S., 2004. Creating the evidence base for quality improvement collaboratives. *Annals of Internal Medicine*, 140 (11), p. 897-901.

Øvretveit, J. & Gustafson, D., 2002. Evaluation of quality improvement programmes. *Quality and Safety in Health Care,* 11, p. 270-275.

Øvretveit, J., 2003. *What are the best strategies for ensuring quality of hospitals: Health Evidence Network.* Copenhagen: World Health Organization Europe.

Rogers, E. M., 1995. *Diffusion of innovations.* New York: Free Press.

Schellekens, W., Rouppe, van der Voort M. & Van Splunteren, P., 2003. A stone in the pond: stimulating hospitals to implement proven improvements. *Medisch Contact*, 58 (35), p. 1302-1304 (in Dutch).

Sluijs, E. M. & Wagner, C., 2000. Quality systems in health care institutions: the situation in 2000. Utrecht: NIVEL (in Dutch).

Sluijs, E. M., Keijser, A., Wagner C., (2007). Quality Systems in Health Care Institutions: the Situation in 2005 (in Dutch). NIVEL: Utrecht.

Snyder, C., Anderson, G., 2005. Do quality improvement organizations improve the quality of hospital care for medicare beneficiaries? *The Journal of the American Medical Association,* 293, p. 2900-2907.

Splunteren, P. van, et al., 2003. Effective breakthrough: improving patient care with the breakthrough method. Assen: Van Gorcum (in Dutch).

Svensson, J., Trommel, W. & Lantink, T., 2005. Avoiding a messy mix: hybrid organization of labour reintegration. Enschede: University of Twente. Paper written to be presented at the International Research Symposium on Public Management IX.

Van Thiel, S., Leeuw, F. L., 2003. The performance paradox in the public sector. *Beleidswetenschap*, 17 (2), p. 123-143 (in Dutch).

Van der Grinten, T. E. D., 2006. Caring for policy: on dependencies and changing relations in health care governance. Rotterdam: Erasmus University (In Dutch).

Wagner, C., 1999. Implementation and effectiveness of quality systems in health care organizations. Utrecht: NIVEL.

Wagner, C., De Bakker, D. H. & Groenewegen, P. P., 1999. A measuring instrument for evaluation of quality systems. *International Journal for Quality in Health Care,* 11, p. 119-130.

Weggeman, M., 1992. Managing professionals. Deventer: Kluwer Bedrijfswetenschappen (in Dutch).

PART 4
AFTER DATA GATHERING AND DOING THE ANALYSIS

Assessing Adverse Events in Brazilian Hospitals: a Pilot Study

Walter Vieira Mendes*
Claudia Travassos**
Mônica Martins*

SUMMARY

Objective: To evaluate the feasibility of the assessment tools used by Canadian Adverse Events Study (CAES) and adapted to the Brazilian context, to identify Adverse Events (AE) in Brazilian hospitals based on retrospective chart review.
Methods: Retrospective review of a random sample of charts of adult patients admitted in 2003 at a teaching hospital in Rio de Janeiro. Patients under 18 years old, psychiatric patients and patients whose length of stay was inferior to 24 hours were excluded, obstetric cases were included. Trained nurse reviewers selected cases with potential adverse events based on screening criteria. Screened charts were evaluated by physicians based on a structured implicit evaluation to identify AE occurrence, moment, place, origin and whether it was preventable or not.
Results: The incidence of AE was 10.1%. When obstetric cases were excluded, the AE incidence reached 12.7%. 69% of the cases with AE were considered preventable. The patient's ward was the most frequent place of AE occurrence (53.7%) and the most frequent origin of AE was medical procedure (29.1%).
Conclusions: The pilot-study indicated the feasibility of using the adapted CAES assessment tools to measure the frequency of AEs in Brazilian hospitals The incidence of AE at the studied hospital is similar to that one found in international studies. However, the preventable AE proportion was much higher in this study.
Key Words: Adverse events in hospitals; health services evaluation; patient safety; quality in health care.

* Escola Nacional de Saúde Pública/Fundação Oswaldo Cruz (ENSP/FIOCRUZ) (wmendes@ensp.fiocruz.br).

** Instituto de Comunicação e Informação Científica/ Fundação Oswaldo Cruz (ICICT/FIOCRUZ).

Introduction

Adverse events (AE) for this study was defined as an unintended injury or harm resulting in death, disability or dysfunction temporary or permanent or prolonged hospital stay that arise from health care. Health care includes the actions of each individual member of the hospital staff as well as the multiple care processes. AE include both errors of omission and errors of commission of the staff directly or indirectly involved with patient care. Although not the first study in this field, the Harvard Medical Practice Study (HMPS) (Kohn et al., 2000), conducted in New York City in 1984, drew attention to the occurrence of AE in hospitals, as a severe and hardly known problem. After that, several other studies were published using the same methodology — retrospective chart reviews — in the US (Thomas et al., 2000), Canada (Baker et al., 2004), Denmark (Schioler et al., 2002), France (Michel et al., 2004), Australia (Wilson et al., 1995), New Zealand (Davis et al., 2001), UK (Vincent et al., 2001) and, more recently, Spain (Aranaz, 2006).

A literature review (Mendes et al., 2005) based on the Medline, using the key-term *"adverse events"*, identified 17,295 publications in the last 50 years (August 2004). A new search was using the same method but in March 2007 identified 26,187 publications — an increase of 34% in three years, and showing growing interest on the subject. Because of concern about patient safety, the World Health Organization (WHO) created, in the year of 2004, *The World Alliance for Patient Safety.* The goal was the development of global policies to improve the quality of the care provided in the healthcare organizations (WHO, 2006).

In Brazil, the research on AEs has focused on the frequency of AEs associated to specific causes, e.g. medication use, surgical issues, anesthetics procedures, invasive non-surgical procedures, or the care provided by the healthcare staff (Mendes, 2005). Yet it is important to know the overall incidence of AEs in hospital organizations, in order to understand the size of the problem in Brazilian hospitals, to give incentive and guide the development of policies with to improve quality. The objectives of this study were to evaluate the feasibility of the assessment tools used by Canadian Adverse Events Study (CAES), and adapted to the Brazilian context, to identify AEs in Brazilian hospitals based on retrospective chart review.

Methods

This pilot-study has assessed the incidence of AEs and their possible causes in one Brazilian hospital. The study design was a retrospective chart review, based on the assessment tools developed by the CAES (Baker et al., 2004). By request, the CAES

team authorized us to use their assessment instruments. The pilot-study was performed in a public teaching hospital in the city of Rio de Janeiro, which provides acute care, emergency care and obstetric care. This hospital was selected for its voluntary willingness to cooperate and the comparatively better quality of their patient charts.

The study population was patients admitted in the year 2003 (13,980). A random sample of patients was selected. The sample frame excluded patients under 18 years old, patients that stayed in the hospital less than 24 hours and the cases with a most responsible diagnosis related to psychiatry. Unlike the Canadian study, obstetric cases were included in the sample due to the high rate of maternal deaths still existing in Brazil (73 maternal deaths for 100,000 live births in 2003). It must be noted that the great majority of deliveries occur in hospitals (97%). The parameters used to define the size of the sample were based on the Canadian study: the potential proportion of AE of 50% and the expected incidence of AE of 10% (maximum absolute error — 3%) with significance level of 5%. A loss rate of 10% was estimated. As the list of admissions used to select the sample did not contain detailed data, the rate of ineligible patients was estimated to be 20%. The final sample size was 553 patients with 385 considered eligible for the study.

The assessment of AEs involved 2 phases: phase 1 is an explicit review by nurses to screen for potential adverse events (pAEs) and phase 2 is an implicit structured review, by physicians, to identify AEs. The explicit review is based upon screening criteria. The presence of at least one screening criterion indicates the record for the second stage phase review. Standard records were specially designed for the training of the reviewers. Reviewers were released to the field only after reaching at least 80% of agreement with standard records. All physicians and nurses had over twenty-year experience. The physicians had clinical backgrounds.

The nursing review form (phase 1 review form) and the physician review form (phase 2 review form) developed by CAES were translated and adapted to the reality of the Brazilian hospitals. First, the forms (phase 1 and 2) were translated from English to Portuguese by two different translators, followed by a comparison of the two versions. Disagreements between translators were then discussed based on the accuracy and understanding of each term as well as to adjustments to Brazilian language and health care context. Second, an expert panel decided by consensus on the best translation of key terminology. The expert panel also decided on the list of screening criteria, based on CAES criteria. Software was developed for data collection, the forms pre-tested and phase 1 and phase 2 forms were back-translated.

The expert panel excluded 2 screening criteria, added one and modified 5. Excluded criteria were those meant to screen for unplanned admissions one year after or before the index admission (admission under study). These were judged inappropriate in the Brazilian context, because of lack of systematic documentation in the patient records

about details of previous admissions to the same hospital or to other organizations. The added criterion refers to elevations of the creatinine level during hospital stay. This criterion was included to identify those patients that developed acute renal insufficiency during admission. The other 5 criteria were modified in regard to their writing for accuracy of understanding.

In the assessment of AEs, the physician reviewers first identify the presence of unintended injuries. Then, they analyzed injuries for any association with temporary or permanent disability and/or prolongation of hospital stay or death. Finally, the physician reviewers, using a 6-point scale determine if the injury was caused by the care provided to the patient. An injury is classified as an AE when it is rated as 4 or more. The preventability of the AE is also judged according to a 6-point scale. An AE is classified as preventable when rated as 4 or more — Box 1.

The physician reviewers assessed if the length of hospital stay was to any extent due to the AE; to what extent it was associated to the AE and the number of days of hospital stay that were associated to the AE. The timing of occurrence of the injury related to the EA and the timing of detection of the AE are also assessed, as well as the site of occurrence. The origin of the AE — assessment of the patient, surgical procedures, orthopedic care (fractures), anesthesia, obstetrics, clinical procedures, hospital systems and medication — is also identified. Data on demographic characteristics of the patients is also obtained — sex (male or female), age group (10 to 20 years; 21 to 30 years; 31 to 40 years; 41 to 50 years; 51 to 60 years; 61 to 70 years; and 70 years and more) and color and race (White; Black; Brown; Yellow).

BOX 1

SCALES AND INSTRUCTIONS GIVEN TO PHYSICIAN REVIEWERS TO JUDGE CAUSATION AND PREVENTABILITY OF ADVERSE EVENTS

Causation

After due consideration of the clinical details of the patient's management, irrespective of preventability, and your response to the questions above, what level of confidence do you have that the health care management caused the injury? (choose one)?

- Virtually no evidence of management causation?
- Slight to modest evidence of management causation?
- Management causation not likely (less than 50/50, but "close call")?
- Management causation more likely (more than 50/50, but "close call")?
- Moderate to strong evidence of management causation?
- Virtually certain evidence of management causation.

Preventability

Rate, on a 6-point scale, your confidence in the evidence for preventability of the adverse event:

- Virtually no evidence of preventability?
- Slight to modest evidence of preventability?
- Preventability not quite likely (less than 50/50, but "close call")?
- Preventability more than likely (more than 50/50, but "close call")?
- Strong evidence of preventability?
- Virtually certain evidence of preventability.

Agreement in screening criteria between nurse reviewers was measured using simple agreement at a significance level of 5%. At each 10 case reviewed, the next was also reviewed by another reviewer, previously assigned as his or hers pair for comparison purpose. Cases for testing inter-raters agreement were automatically selected by the Software.

A database in MS-Access is generated by the Software. This database is exported to the format MS-Excel and data was analyzed using SPSS®13.0.

Results

In the pilot study, 385 patients were considered eligible cases. Of these, 252 (65.5%; CI 95%: 60.7-70.2) were non-obstetric patients, and 113 (34.5%; CI 95%: 29.8-39.3) were obstetric patients.

Of the total of cases, 269 (70%) were female. Even after exclusion of the obstetric cases, women prevailed — 136 (54%). The most frequent age group was 21 to 30 years old (26.5%). When excluded the obstetric cases, the predominant age group turned to be 41 to 50 years (20.2%). Regarding color/race, white patients prevailed (48% and 50%) — Table 1. The inpatient mortality rate was 10.3% and there wasn't any case of maternal death.

Of all cases, 167 (43.4%) were selected as pEA, *i.e.*, patients with at least one screening criterion identified by the nurse reviewer. Excluded the obstetric cases, the number of pEAs was 129 (55.4%).

Of all cases, 39 had at least one AE, which corresponds to an AE incidence of 10.12% — Table 2. Some patients had more than one AE. The number of AEs identified in these patients was 52. Among patients with EA, 27 (69.2%) were classified as having had a preventable AE. Excluded the obstetric cases, the AE incidence was 12.7%. Of these, 68.7% were classified as preventable AE — Table 2.

Table 1
Demographic characteristics of patients, total and non-obstetric cases, by sex, age group, and color/race, 2003

DEMOGRAPHIC CHARACTERISTICS		TOTAL CASES% (N)	CI 95%	NON-OBSTETRIC CASES% (N)	CI 95%
Sex	Female	69.9 (269)	65.3-74,5	54.0 (136)	49.0-58.9
	Male	30.1 (116)	25.5-34.7	46.0 (116)	41.1-51.0
Total		100 (385)	–	100 (252)	–
Age group	10-20 years	8.8 (34)	6.0-11.7	3.6 (9)	1.7-5.4
	21-30	26.5 (102)	22.1-30.9	10.7 (27)	7.6-13.8
	31-40	17.1 (66)	13.4-20.9	15.5 (39)	11.9-19.1
	41-50	14.8 (57)	11.3-18.4	20.2 (51)	16.2-24.3
	51-60	8.8 (34)	6.0-11.7	13.5 (34)	10.1-16.9
	61-70	11.7 (45)	8.5-14.9	17.9 (45)	14.0-21.7
	70 and more	12.2 (47)	8.9-15.5	18.7 (47)	14.8-22.5
Total		100 (385)	–	100 (252)	–
Color/race	White	47.9 (183)	42.9-52.9	49.8 (124)	44.8-54.8
	Black	20.9 (80)	16.9-25.0	22.9 (57)	18.7-27.1
	Brown	23.0 (88)	18.8-27.2	20.5 (51)	16.5-24.5
	Yellow	2.1 (8)	0.7-3.5	2.0 (5)	0.6-3.4
	Not-informed	6.8 (26)	3.6-8.4	6.0 (15)	2.7-7
Total		100 (385)	–	100 (252)	–

Table 2
Incidence of Adverse Events (AE) per 100 patients and the proportion of preventable AE, 2003

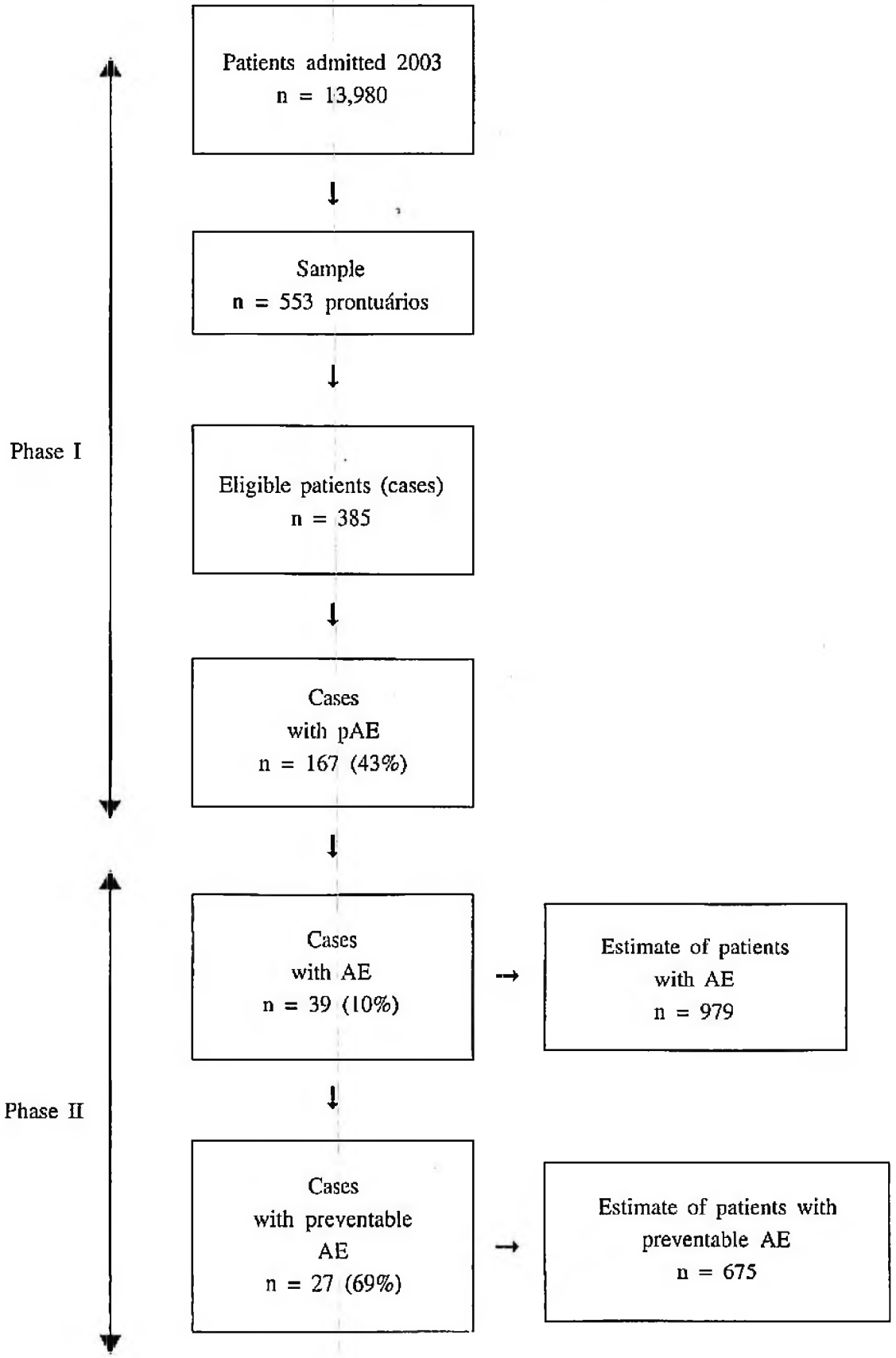
ADVERSE EVENTS	INCIDENCE	CI 95%	PROPORTION OF PREVENTABLE AE% (N)
At least 1 AE — all cases	10.1 (39)	7.1-13.1	69.2 (27)
At least 1 AE — excluded obstetric cases	12.7 (32)	9.4-16	68.7 (22)

It's important to point out that 7 obstetric patients (5.3%) experienced one or more AE, and that 70% of those were classified as preventable. There wasn't any maternal death reported.

The number of additional days in the hospital that was attributed to the occurrence of an AE was 244. Extrapolation of these results indicates 979 patients with AEs in this hospital in 2003. Of these, 675 are cases with preventable AE — Figure 1.

The most frequent origin of AEs was medical procedures (29.1% of the total of cases and 38.1% after exclusion of obstetric cases), i.e., non-surgical invasive

Figure 1
Estimate of incidence and preventability of AE for patients admitted to the hospital in the year studied



procedures such as: vesical catheterism, traqueal intubation, deep vein punction, endoscopy, broncoscopy, contrasted exams. Second in frequency were surgical procedures (25.5%) — Table 3. The proportion of error due to omission was 44.2% and from error in the care provided (action) was 55.8%. Excluded the obstetric cases, 42.9% resulted from errors of omission and 57.1% from action.

The higher frequency of AEs was observed in the wards (53.7%), followed by the operating theatre (27.8%) and ICU (11.1%). The results were similar for obstetric and non-obstetric cases — Table 4. The most frequent timing of occurrence and detection of AE (96.2%) was during admission.

Table 3
Origin of adverse events (AE) for total and non-obstetric cases

ORIGIN OF ADVERSE EVEN	TOTAL CASES% (N)	CI 95%	NON-OBSTETRIC CASES% (N)	CI 95%
Medical procedures	29.1 (16)	24.6-33.6	38.1 (16)	33.2-42.9
Surgical	25.5 (14)	21.1-29.8	26.2 (11)	21.8-30,6
Diagnosis	16.4 (9)	12.7-20.1	19.0 (8)	15.1-23
Obstetric	12.7 (7)	9.4-16.1	0 (0)	–
Medication	5.5 (3)	3.6-7.4	7.1 (3)	4.6-9.7
Fractures	3.6 (2)	1.8-5.5	0 (0)	–
Others	3.6 (2)	1.8-5.5	4.8 (2)	2.6-6.9
Anesthetic	1.8 (1)	0.5-3.2	2.4 (1)	0.9-3.9
System events	1.8 (1)	0.5-3.2	2.4 (1)	0.9-3.9
Total	100* (55)	–	100 (42)	–

* In three AEs, more than one origin was identified.

Table 4
Site of adverse events (AE) for total and non-obstetric cases

	SITE OF AE	TOTAL CASES% (N)	CI 95%	NON-OBSTETRIC CASES% (N)	CI 95%
Inside the Hospital	Delivery room	1.9 (1)	0.5-3,2	0 (0)	–
	Room or ward	53.7 (29)	48.7-58,7	60.5 (26)	55.6-65.3
	Emergency room	1.9 (1)	0.5-3,2	2.3 (1)	1.1-3.6
	Operating theatre	27.8 (15)	23.3-32,3	23.3 (10)	19-27.5
	ICU	11.1 (6)	8-14.3	11.6 (5)	8.4-14.8
	Procedures room	1.9 (1)	0.5-3.2	2.3 (1)	0.8-3.8
	Service area	1.9 (1)	0.5-3.2	0 (0)	–
Out of the hospital		0 (0)	–	0 (0)	–
Total		100*(54)	–	100 (43)	–

* In two AEs, more than one site was identified.

The most frequent screening criteria selected at phase 1 review were:

- *Criterion 19* — "Any unwanted events not mentioned above" — 74 (24.3%);
- *Criterion 3* — "Occurrence of injury or harm to patient during hospitalization (including any harm, lesion, or trauma occurring during index hospitalization)" — 37 (12.1%).
- For screening *criterion 17* — "Documentation or correspondence indicating litigation, whether merely intent to sue or actual lawsuit", no case was selected.
- For screening *criterion 18* — "Starting with normal creatinine at admission, did the level double during the hospital stay?", suggested by the experts panel for the Brazilian study, 8 (2.6%) of cases were selected — Table 5 (see next page).

Two of the CAES list of screening criteria not approved by the experts panel:

- *Criterion 1* — "Unplanned hospitalization (including readmission) as a result of any healthcare provided during the 3 months prior to the index admission".
- *Criterion 2* — "Unplanned admission to any hospital during the 3 months following discharge from the index hospitalization" — virtually did not modify the incidence of AE, which changed from 10.1% to 9.9%, and did not modify the proportion of preventable AE. Taken just the criterion "Starting from a normal creatinine in admission, was the value duplicated during hospital stay?" — There wasn't any change either in the incidence of AEs or in the proportion of preventable AEs.

Agreement between nurse reviewers in assessing potential AE was good, although better between the nurses A and B than between the nurses C and D — Table 6.

Table 6
Simple concordance between nurse reviewers in assessing potential adverse events

REVIEWERS	NUMBER OF ASSESSED RECORDS	SIMPLE CONCORDANCE	CI 95%
A and B	16	87.5%	85.8-89.2
C and D	17	76,5%	74.2-78.8

Discussion

The pilot-study indicated the feasibility of using the adapted CAES assessment tool to measure the frequency of AEs in Brazilian hospitals. The AEs incidence was of 10.1%. Of these, 69.2% were represented by preventable AEs. In excluded obstetric cases, the AEs incidence was 12.7%. The incidence of 10% found in this study is

Table 5
Potential Adverse Events (pAEs) per 100 patients by screening criterion

SCREENING CRITERIA FOR POTENTIAL ADVERSE EVENTS	% (N)	CI 95%
1. Unplanned hospitalization (including readmission) as a result of any healthcare provided during the 3 months prior to the index admission	11.5 (35)	8.3-14,7
2. Unplanned admission to any hospital during the 3 months following discharge from the index hospitalization	3.9 (12)	2.0-5,9
3. Occurrence of injury or harm to patient during hospitalization (including any harm, lesion, or trauma occurring during index hospitalization)	12.1 (37)	8.9-15.4
4. Adverse reaction to medication	1.6 (5)	0.4-2.9
5. Unplanned transfer to intensive or semi-intensive care unit	2.0 (6)	0.6-3.4
6. Unplanned transfer from or to another acute care hospital (excluding transfers for specialized exams, procedures, or care not available in the original hospital)	1.6 (5)	0.4-2.9
7. Unplanned return to surgery room	2.3 (7)	0.8-3.8
8. Unplanned removal, lesion, or repair of an organ or structure during surgery, invasive procedure, or vaginal delivery	1.3 (4)	0.2-2.4
9. Other unexpected complications during reference hospitalization which are not a normal development of the patient's disease or an expected result of the treatment	10.8 (33)	7.7-13.9
10. Development of a neurological alteration absent at admission, but present at the time of discharge from the index hospitalization (includes neurological alterations related to procedures, treatments, or investigations)	2.0 (6)	0.6-3.4
11. Death	8.9 (27)	6.0-11.7
12. Inappropriate hospital discharge/scheduling of inadequate discharge from index hospitalization (excludes unauthorized discharge)	9.0 (3)	1.3-4.6
13. Reversed cardio-respiratory arrest	2.3 (7)	0.8-3.8
14. Injury related to abortion or labor and delivery	1.0 (3)	0-2
15. Hospital infection/septicemia (excludes infections/septicemia occurring fewer than 72 hours after admission)	6.6 (20)	4.1-9
16. Dissatisfaction with care received as documented on medical chart, or evidence of complaint lodged (includes documents, documented complaint, conflicts between patient/family and healthcare professionals, and unauthorized discharge)	2.3 (7)	0.8-3.8
17. Documentation or correspondence indicating litigation, whether merely intent to sue or actual lawsuit	0 (0)	–
18. Starting with normal creatinine at admission, did the level double during The hospital stay?	2.6 (8)	1.0-4.2
19. Any unwanted events not mentioned above	24.3 (74)	20.0-28.5
Total	100 (305)	–

similar to that reported in other studies. The studies that focused in quality improvement of health care in Brazil has revealed an incidence similar to the incidence in the Brazilian study — New Zealand (11.3%), Australia (16.6%), England (10.8%), Denmark (9.0%), France (14.5%), Spain (10.12%) and Canada (7.5%).

However, the proportion of preventable AEs observed in the Brazilian research was higher than that observed in the other studies — Brazil (69.2%), New Zealand (61.6%), Australia (50%), England (52%), Denmark (40.4%), France (27.6%), Spain (42.8%) and Canada (37%). This result possibly suggests that patient safety problems are more frequent in Brazilian hospitals when compared with other more developed countries. Studies conducted in the United States of America with medical-legal focus obtained a lower incidence — California (4.6%), New York (3.7%) and Utah-Colorado (2.9).

In the Brazilian study, the expert panel excluded two criteria and included one. However, these changes did not have any impact on the incidence of AE or on the proportion of preventable AE. This result point out the fact that it is more important to focus on the comprehensiveness of the group of screening criteria — criteria are not mutually excluded — than to focus in the inclusion/exclusion of a single criterion. However, the unspecific criterion (number 19 — Any unwanted events not mentioned above), was the one with the highest frequency in the pilot-study (24.3%). Possible causes are due to recurrent problem of lack of medication in the hospital and other facts related to hospital's structure/system. In the case of Brazil, it might be necessary to create criteria for systems' failures, due to their high incidence.

The finding of AE in 7 cases (5.3%) with a responsible diagnosis related to obstetrics, being 5 of those preventable AEs has shown how the relevance of including the obstetric admissions in the studies on AE for Brazilian hospitals.

Some methodological limitations are worth mention. The validity of the screening process and the quality of data in the patients' records were not systematically assessed. Quality problems in one or another might cause sub-estimation of the AEs incidence. It's also important to point out that the validity of the implicit evaluation depends upon the experience and knowledge of the physician reviewer.

The pilot-study analyzed one single hospital, which represents a limited sample of the universe of public hospitals of the city of Rio de Janeiro. Moreover, the criterion used for choosing this hospital — good quality of their patient's records — selects the best hospitals. In general, hospitals that manage to have good quality patient's records deliver good quality of care. On the other hand, one has to note that the hospital under study may have a higher incidence of AE, due to the fact of being a teaching hospital. The reported incidence of AEs in teaching hospitals in other studies is higher than in other kinds of hospitals (Baker et al., 2004). Finally, it must be highlighted that the high volume of obstetric cases contributed to decrease the sample

of non-obstetric cases, a fact to be considered in future studies. But, we recommend the inclusion of obstetric in countries with low quality maternal care.

The pilot-study indicates that it is feasible to use retrospective chart review to assess AEs in Brazilian teaching hospitals. It also shows the need to assess the magnitude of patient safety problems in Brazilian hospitals in general. The tools adopted in this study can provide the basis to the development of monitoring tools and can be applied in association with methods directed to evaluate specific-origin AEs.

References

Aranaz, J. M., 2006. Estudio nacional sobre los efectos adversos ligados a la hospitalizacion: ENEAS 2005: informe Febrero 2006. Madrid: Ministerio de Sanidad y Consumo.

Baker, G. R., et al., 2004. The Canadian adverse events study: the incidence of adverse events among hospital patients in Canada. *CMAJ — Canadian Medical Association Journal*, 170, p. 1678-1686.

Davis, P., et al., 2001. Adverse events regional feasibility study: indicative findings. *The New Zealand Medical Journal,* 114, p. 203-205.

Kohn, L. T., 2000. *To err is human*. Washington, D.C.: National Academic Press.

Mendes, W., Travassos, C., Martins, M. & Noronha, J. C., 2005. Revisão dos estudos de avaliação da ocorrência de eventos adversos em hospitais. *Revista Brasileira de Epidemiologia,* 8 (4) p. 393-406.

Michel, P., Quenon, J. L., Sarasqueta, A. M. & Scemama, O., 2004. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. *British Medical Journal,* 328, p. 199-202.

Schioler, T., et al., 2002. Danish Adverse Event Study: incidence of adverse events in hospitals: a retrospective study of medical records. *Ugeskr Laeger,* 164, p. 4377-4379.

Thomas, E. J., et al., 2000. Incidence and types of adverse events and negligent care in Utah and Colorado. *Medical Care,* 38, p. 261-271.

Vincent, C., Neale, G. & Woloshynowych, M., 2001. Adverse events in British hospitals: preliminary retrospective record review. *British Medical Journal,* 322, p. 517-519.

Wilson, R. M., et al., 1995. The quality in Australian Health Care Study. *The Medical Journal of Australia,* 163, p. 458-471.

World Health Organization, 2006. *World Alliance for Patient Safety: forward programme*. [online] Geneve; 2006-2007. Available at: http://www.who.int/patientsafety/en/ [accessed 20 October 2007].

Patients' Perception of Hospital Safety

José Joaquín Mira*
Jesús Aranaz*
Julián Vitaller*
Susana Lorenzo**

ABSTRACT

Background: Almost 10% of patients hospitalized in developed countries suffer from discomfort, injuries or damage, irreparable in some cases, directly related to clinical trial or treatments (adverse events — AE). Referring medical mistakes or errors in clinical context proves to be complex due to emotional, economic, social, professional and legal consequences. However, failing to talk about medical errors is an act of irresponsibility, given it yields worse consequences. In this chapter we will review some studies that shed light on the participation of patients in clinical safety, such as: citizens' and patients' perception of safety, professional practice styles and physician-patient interactions favoring greater safety or study results on the identification of these adverse events by patients.

Main research question: We present the results of our own research performed with qualitative and quantitative methods as well as those of the R&D projects we are currently conducting.

General design and sample: Firstly, a observational study based in a survey administered at discharge to 12,389 registered patients at public hospitals in Spain was conducted. Secondly, a series of focus groups with patients and professionals was planned to gathered data in order to propose the content of the scale.

Data collection and analysis methods: In the first study we explore the relationship between patient satisfaction and the frequency of AE. In the second study we are developing a scale to know patient's perception of hospitals' safety (reliability and validity analysis will be conducted).

Main findings: Although we do not conclude a direct responsibility of information given in having an AE, data confirm the importance of a good

* Universidad Miguel Hernández — Agencia Valenciana de Salud (Jose.mira@umh.es).
** Fundación Hospital Alcorcón.

communication approach with patients, and suggest that clinical safety could be increased when patients are properly informed and when the patient has the possibility to take active part in clinical decisions related to his or her health.
Methodological challenges: An innovative focus under which it is advised to work is based on a "patients for patients' clinical safety" strategy. A first step is to analyze citizens' and patients' opinions, and a second step is to assess which characteristics of the physician-patient relationship play a role in safety.
Practical challenges: Governments and the WHO have set up efforts for the clinical safety of patients. Raising awareness among professionals and patients about the significance of this problem is a shared goal, and one of the foundations for a new safety strategy.
Main lessons for other researchers: The value of considerer the patients' point-of-view to increase safety.

The practical problem and the gap in knowledge this research addresses

Adverse events (AE) have now been defined as "the price to pay for modern diagnosis and therapy methods" (Barr, 1955). It is true that medical development, with its ever increasing specialization and complexity, carries a pricetag that involves injuries or involuntary harm to patients. Even though we are willing to assume this cost collectively, issues are very different on an individual level and accepting it proves to be tough.

Citizens expect hospitals to establish a framework that guarantees patient safety, since they logically consider that whoever comes to this kind of health facility expects to be in better condition when "leaving" than when "arriving." But we know this is not always the case by very different reasons, ranging from the most frequent system failures to the most infrequent errors made by one person (Leape et al., 1991).

From its very beginnings, health systems assume the goal of achieving patients' well-being. To reach this goal, they provide, among other options, the best possible therapy or, at least, the most appropriate therapy given the specific circunstances (available technology, scheduling pressure, availability of personal resources and materials, etc.). The set-up of scheduling protocols or the recent interest in establishing practice guidelines are nothing more than examples of this fundamental principle: "first, do not harm".

When patients suffer an AE with severe consequences as a result of specific medical attention, they are likely to interpret it as an act of nurses' or physicians' negligence. This fact is definitely, a mistake. Nevertheless, sometimes AEs they are hardly avoidable, other times, due to system failures not ascribable to any professional

in particular, other times, professional errors, and some times, they are attributable to the patient him/herself.

The most minor AE's usually have clinical impact on the patient that involves fear, discomfort and pain. However, in other cases they involve exposure to unnecessary and highly dangerous risks that end up causing harm or injury of not-so-minor consequences (Aranaz, Aibar, Gea & León, 2004) (and in a substantial percentage, the consequences are quite serious or fatal — Brennan, & et al., 1991). Furthermore, it is not unusual for these events to generate important press releases, legal and emotional effects, and influencing citizens' perception of the quality of medical assistance they receive and their trust in the work performed by health professionals.

Paradoxically, an AE may be an indirect cause of another AE, making it difficult to break this chain. Firstly, because the professionals, who feel threatened, do not report them (preventing them in potential patients). Secondly, because patients, who have suffered AE, are more distrustful and, in their wariness, may carry out risky behaviors.

If AE's are at one extreme, safety is at the other. For instance: Institute of Medicine (IOM, 2007) goal of reducing AE's to 50% in less than 5 years, or the program for patient identification established in many Spanish hospitals, or hand-washing with alcohol solutions common in almost every hospital; are all signs of the change of focus necessary to improve health care quality.

Although the greatest responsibility lies on health professionals, it would be unrealistic to say that clinical safety is an issue that only involves professionals. Patients might and should take an active role in their actions as well as their attitudes (Mira & Rodríguez-Marín, 2001). An active role in clinical decision making, or in health care itself, requires, as a first step, being positively aware of the importance of clinical safety to improve the help we provide.

Judging patients' "submissive" attitude as positive (considering better those who accept all explanations without question and who raise no objection to the incidents of the health care process) may soon be something of the past. On the contrary, a good patient is someone who actively participates in the clinical decision-making process in a responsible way and who actively contributes to his treatment and recovery (Guadagnoli & Ward, 1998; Jadad, Rizzo & Enkin, 2003; Mira, 2006).

Citizens' perceptions on clinical safety levels offered by hospitals thus play a decisive role in this change of attitude.

The data available on patients' perceptions, at least in Spain (Ministerio de Sanidad, 2004), indicate that patients (and consequently the citizens of our developed countries) are increasingly demanding from the Heath Care System more and better services. At the same time news on technological breakthroughs in the field of biomedical research promising encouraging results intensify these expectations for the

results of medical interventions. Logically, this cycle has consequences in the way in which patients and professionals interact.

Among the environmental barriers against clinical safety are the patients' unrealistic expectations on what medicine is capable of achieving, high doses of distrust towards professionals or towards the health care system, doubts regarding what could happen to them, or regarding death as the physicians' failure.

The study of patients' perceptions of clinical safety in hospitals can help identify guidelines to involve in clinical safety and help reduce the frequency of certain AE's, with the aim of better health care quality.

What helped us to design the study and to resolve some methodological difficulties. Main lessons for other researches

Patients' hospital safety: are there any reasons for concern?

In screening studies followed by a review of clinical reports, the AE rate ranges from 2.9% to 16%. It is believed that around 50% of them could have been avoided in a simple manner (Wilson, R. M. et al., 1995).

As far as we know, drug reactions are the most frequent AE's (19%) followed by nosocomial (in-hospital) infections of surgical wounds (14%) (Brennan et al., 1991).

Most AE's (almost three-quarters) have mild and perfectly reversible consequences. However 1 out of 10 has permanent or fatal consequences (Mostaza, Muinelo, Teijo & Pérez, 2005).

The recent ENEAS trial (Aranaz, Aibar, Vitaller & Ruiz, 2006) carried out in Spain, has shown that patients admitted to a small hospital have a 1.5 times higher risk of AE's than patients admitted to a large hospital. It has also been shown that patients older than 65 years with extrinsic risk factors have 2.5 times higher risk than younger patients, and hospitalization longer than 1-week involves a 3.5 times higher risk of AE.

Do patients who are admitted to hospital feel safe?

The European Union Eurobarometer (Commission Européenne, 2006) on citizens' perceptions of medical errors shows that 9% of Europeans believe that becoming a victim of errors is highly likely during a hospital stay in their country.

According to this study, 3 out of 4 Europeans perceive medical errors as an important problem in their own country, and 38% understand that it is a very relevant

issue. It is in South-Eastern and Eastern citizens show more concern about the likelihood of suffering medical errors.

Those who have experienced a medical error personally or in their families are logically the ones who consider that this constitutes a more important problem (53% vs. 33%). Italy (61%), Poland (54%) and Latvia (50%) top the list of countries whose citizens believe that medical errors are a very important issue, whereas Finland and Denmark are at the bottom. Women, senior citizens and people with lower educational level tend, in general, to be more concerned about medical errors.

According to data provided by the Kaiser Foundation (The Kaiser Family Foundation, 2000), the number of citizens worried about clinical safety in the United States is 15 percentage points higher than the number of Americans who state they are concerned about air safety. In two studies published in 2002 in the same country, we see both sides of the coin. In the first study, directed by Prof. Blendon (Blendon et al., 2002), medical errors were not viewed as an important problem of the American Health Care System. Conversely, in the study by Robinson et al. (2002) the reduction of medical errors was a priority deserving of special attention from health care authorities for 9 out of 10 respondents.

In Australian studies based on polls like the previous ones, 5.2% of the population expressed concern about clinical safety in case of hospital admission. Women, people between 40 and 59 years of age, metropolitan residents and those with incomes above 80,000 Australian dollars per year were the ones with the highest perception of insecurity.

Citizens' perceptions of clinical safety vary considerably. While in Europe 48% claim to feel safe when hospitalized only 23.3% of Australians consider themselves fully safe. The reasons for this discrepancy may be due to very different reasons, including the approach of poll-based studies. Nevertheless, they point out two fundamental issues. The first is that there is a large number of people worried about clinical safety. The second is that proximity to a "possible" medical error awakens this concern.

Patients' greater ability to participate in the decisions affecting their health, including the ability to chose the center and the professionals, probably affects that greater concern about quality and clinical safety. The Picker Institute (Coulter, A. & Magee, 2003) in its study of the European patient profile, concluded a couple of years ago that patients increasingly want to know more details about their illness, therapeutic alternatives, and possible risks and complications, as well as the qualification level of the professionals helping them.

The other side of the *safety coin* is the increase in lawsuits and the feeling among health professionals that health care is being unnecessarily submitted to the legal process (Vitaller, 2004). Neither aspect contributes to building the legal safety climate

for health professionals that would open discussion in terms of safety rather than in terms of errors.

To what degree do patients realize they are victims of an AE?

According to the Eurobarometer conducted between September and October of 2005 in the 25 countries of the European Union: 18% of Europeans (they or one of their family members) have been victims of medical errors during their hospital stay, and 11% have had problems due to errors from the medication prescribed. Latvia, Denmark and Poland are at the top of the list of countries with the most hospital mistakes. Latvia and Denmark also lead the blacklist of European countries in the specific case of medication errors.

Ten percent of Europeans (and the figure reaches 40%, if all those who show a certain degree of concern are taken into account) are very worried about becoming victims of a medical error, this concern being greater among those who have suffered or who have been close to an error. Swedish, Austrian and Dutch citizens are the least concerned.

On the other side of the Atlantic, the Kaiser Foundation and the Agency for Healthcare Research and Quality (2000) made public, five years earlier, their results on the safety perceptions of American citizens. In this case, 6% of respondents declared having been a victim of an AE within the 12 months prior to the study.

In a poll conducted in 2004 by Adams and Boscarino (Adams & Boscarino, 2004) of over 1001 former patients in the state of New York, 11.4% of respondents claimed to have been the victim of an AE (or to have known another patient who had suffered one) in the five years prior to the poll. When asked if someone in their family has suffered a medical error, the percentage of those who reply affirmatively rises to 21.1%.

In line with the data we already know, 5.9% of these supposed errors were related to medication, 7.2% to surgical procedures and 6.3% to errors in diagnosis.

Curiously, in this study by Adams and Boscarino the percentage of patients who were victims of a medical error was higher among those who sought information on the quality of the medical practice results before deciding what physician or medical center to select. Thus, seeking information multiplied the likelihood of being victim of a medical error by 1.96.

If we change continents and go south to Australia (Evans, Berry, Smith & Ester, 2006), the percentage of patients who had suffered an AE during their stay in hospital goes down to 7%. In something more than half the cases (59.7%) the AE had serious consequences for the patient. Evans et al., also reached the conclusion that the perception of safety was different in those who had suffered an AE (among patients

who had suffered an AE with serious consequences, the likelihood of their stating that there are hospital safety problems was multiplied by 2.38).

Are there practice styles that are risky?

In daily practice we see styles for guiding the visit and for communicating with the patient that seem to strongly influence the way a patient may interact with a physician. Let us look at three examples.

First, the visit during which the physician spends more time talking than the patient. In this case, the chances of a patient asking him/her questions or about possible signs denoting alarm will scarcely have time to appear.

Second, after receiving a written complaint, a physician disregards that information because s/he more busy justifying him/herself than learning from that experience.

Third, the case of a physician who only informs the patient of the possible effects of the drugs she has prescribed, ignoring systematically other possible drugs that the patient may be taking. Drug interactions would therefore escape the physician's control.

A relatively recent study conducted in the Primary Care setting of Spain by Barca et al. (2004) highlighted some risk factors derived from well-rooted professional practice styles. In the study patients were asked what the doctor had told them and the doctor was asked what he had told the patients after the office visit. The results reveal that the percentage of Primary Care physicians who tell their patients about the diagnosis amounts to 88%, and the percentage of those who talk about the therapy dosage is 73.91%. However, only 9.35% give information about possible complications or the precautions to take with that treatment.

On the other side of the table, we see that patients do not seem excessively worried about the potential consequences of inadequate therapy. While 30% ask their doctor about the etiology of their illness and 29% about the therapy to follow, only 9% ask questions about possible complications from the treatment. This study reveals something even more serious: although 19% of patients did not understand what the physician was explaining to them, 39% did not ask the doctor any questions or ask the doctor to explain the therapy in a different way so they could understand it better.

On the other hand, while the difference in percentage between physicians' and patients' replies as to whether the diagnosis had been explained reached 50 percentage points (74 to 24) and whether the treatment had been explained reached 60 percentage points (88 to 25), differences about whether possible complications or treatment risks had been explained were only 2 percentage points (9 to 7).

There is other information that reinforces the data under discussion

At least in Spain (Badía, Magaz, Guitierrez & Guilera, 2005), the patients' main source of information about the nature of the drugs they are taking is the information leaflet found in the drug container (specifically, for 76% of patients). The regular doctor is mentioned as the source of information by only 55% or respondents, and the pharmacist by 17%. Curiously, and in line with Barca's study above, 53% of patients consider themselves to be poorly informed about the treatment they are following.

This study also provides other data regarding practice styles that influence patients' clinical safety. 17% of respondents refer that their physician does not offer them information on the therapy prescribed, and 60% state that their physician customarily fails to inform them of possible therapeutic alternatives.

From another perspective, the importance of patients' complaints towards increasing clinical safety has been assessed. Some data suggest that the physicians with a higher number of lawsuits also have a higher number of written complaints against them from patients (Hickson et al., 2002). Based on this observation it has been hypothesized that patients' complaints could be useful in identifying potential clinical risks.

Murff *et al*'s study, conducted at Vanderbilt University Medical Center (Murff et al., 2006), follows these steps and showed a relationship between the number of complaints lodged by patients and the occurrence of post-operative complications, although these complaints do not focus on the intervention's results or on the surgical procedure *per se*, but on other questions related to post-operative care, treatment, comfort or information.

When reviewing the medical records of 16,713 patients, these authors identified an AE rate of 12.6%. The number of complaints lodged by these patients was 151 (0.9%). These complaints covered 509 different aspects of the attention received. Complaints about the care and therapy received amounted to 28% of the total (CI 95% 21-37) and it was the most frequent grounds for complaint.

Nineteen per cent of surgical patients who submitted written complains suffered a major complication after surgery, while, among those who did not complain about anything, the frequency of complications was 12.5% ($p = 0.01$). The number of hospitalization days was also higher in patients who complained (5 days vs. 4, $p = 0.02$).

In summary, the surgical patients who fill in a formal complaint were 1.74 times more likely to suffer an AE than the rest. These data contribute to considering patients' written complaints as a source to identify potential AE's.

Despite these data, we are far from being able to take advantage of the information provided by patients' complaints, and even less so from considering repeated complaints as a sentinel indicator that helps patients' clinical safety.

In a study conducted by Jarvis & Frizelle (2006) in New Zealand, they retrospectively analyzed the number of surgeons who had once received complaints and the number of complaints each had received. In this case, of the surgeons who had received any complaints, 60% had received only one. Three per cent of the surgeons participating in the study had received up to 4 complaints, 9% had received a total of 3 complaints and 11% had received 2 complaints. But when asked about it, none of the physicians thought that the information provided by the patient could be used to improve their surgical practice. On the contrary, 86% considered that their practice was now more defensive, 38% thought that as a consequence they had worsened their relationships with patients, aside from personal and family emotional costs.

In Spain Borrell et al. (2006) analyze the perception of medical errors by 238 primary care doctors with a mean professional life of 14 years. In sum the group totals 2,540 situations that had given rise to AE's or, in other words, 10.6 events per physician per year. Ten physicians (representing 4.3% of the sample) had had a lawsuit for this reason.

In terms of the frequency with which they declared being involved in an AE, physicians were classified into: deniers (did not declare any error), perceptive (admitted error in the last year) and hyperperceptive (admitted more than 28 errors). Additionally, they were also classified in terms of their locus of control: internal (attributed errors to themselves) and external (attributed errors to causes not their own).

In the first case, internal locus of control, the causes of error declared were becoming easily overwhelmed and not being sufficiently systematic in clinical exploration or in the application of protocols. In the second case, external locus of control, the most frequently mentioned cases were scheduling pressure and poor scheduling organization. The first group declared up to 2.1 times more errors than the second group. In the second group there were a significantly higher number of deniers.

Patients for patients' clinical safety: when is it right?

The WHO has set up an Alliance for the clinical safety of patients proposing different lines of action. One of them seeks precisely to involve patients in their own safety.

The proposals they want to bring to the practice are, among others: providing information, participating in decision making, and overcoming barriers so that patients will ask the professionals their questions.

To the best of our knowledge, patients who play an active role in their relationship with the physician have greater control over their condition. Data suggest that an active role by the patient relates to a good outcome of the medical intervention (Mira & Rodríguez-Marín, 2001).

However, not all patients want to play an active role. In fact, the more serious the disease, the more likely the patient will prefer the doctor to be the one who makes decisions (Ende, Kazis, Ash & Moskowitz, 1989) or, at least, plays a much more directive role.

The study of Arora, Allanan & Guadagnoli (2005) may shed light on whether all patients want or are able to become actively involved in clinical decisions. They asked a total of 621 patients attending a health care facility for acute processes whether they were ready to act in view of their disease (perception of self-efficiency being defined as "a person's conviction that they are capable of successfully performing a specific action"— Bandura, 1977) with respect to their attitudes and beliefs about the disease; and their confidence in the physician.

17.2% of patients did not want to participate in decision making and preferred that the physician to be fully accountable. On the contrary, 39.8% wanted to have information to participate actively.

Those who wanted to participate with the physician in the search for the best therapy were the ones who considered themselves as most capable of facing a disease ($F = 4.51$, $p < 0.01$) and also those who showed the greatest confidence in their physician ($F = 4.11$, $p < 0.01$).

According to these results it may be stated that when we find in the patient both a great confidence in the physician and the belief that, whatever they do, they will not be efficient in fighting a disease. The most likely issue is that that patient might prefer the physician to take the initiative and make all the decisions. In this case the patient will assume a very passive role with a low level of involvement in his/her own clinical safety.

On the contrary, self-confidence is a factor that predicts a patient's more active role, and therefore, they search for information, they demand higher attention to the health care process and, ultimately, greater involvement in the therapy.

Studies exploring patients' involvement in clinical decisions show certain unanimity that older people with only basic education and patients suffering from serious illnesses prefer their physicians to make decisions for them. Patients with this profile and those who perceive themselves as little capable of coping with disease would therefore be the ones more exposed to an AE.

Yet, not only the patient's personal traits, such as self-efficacy, play a relevant role. Thanks to social and technological changes we have information at our disposal that was very hard to access a few years ago.

New information technologies have a growing influence on the access to health care information. Some recent data (Eysenbach, 2003) indicate that in countries with a higher number of home Internet connections the number of patients who ask their physicians about information found on the web exceeds 50%.

One of the reasons to turn to the Internet is to contrast different information. Second-opinion Internet traffic has grown dramatically in recent years (Mira, Pérez-Jover & Lorenzo, 2004). Although different factors affect the demand for second opinions (a greater accessibility offered by new technologies, among others), one piece of data corroborated by several studies is that lack of information between physicians and patients is the basis for many of these requests (Van Dalen et al., 2001).

These easier ways to gain access to information, not always of appropriate quality, offers a new framework for the patient-professional relationship and open new possibilities to include the patient in the care of their health.

The initiative of "Questions are the answer" from the Agency for Healthcare Research and Quality (AHRQ) (2007) or the medical visit kit (Universidad de los Pacientes, 2007) in Spain will help patients get ready for their visits to a doctor and make the most of the time available during the visit. This kit includes basic recommendations for the patient about what to ask the physician, how to remember the physician's recommendations or the importance of clarifying questions before leaving the doctor's office. Essentially, these are recommendations that advocate for patients' involvement and responsibility for their own health.

The general research design. Methods and preliminary results

The practical problem and gap in knowledge this research addresses. Lessons learned from consumers, patients and health professionals

In a recent forum of consumers' and patients' associations that took place at the headquarters of the Valencia School of Health Studies *(Escuela Valenciana de Estudios para la Salud)* in Valencia, Spain, we conducted a session through qualitative techniques of context analysis where 8 of the most important (representative) patients associations put forward their viewpoints on the health care quality — safety binomial.

During the discussion it could be noticed that the general idea of medical errors is linked to monetary demands (disputes) after negative results of interventions. The concept of adverse event, as such, seems to go unnoticed.

When we refer to "medical errors" people understand that we mean specific issues concerning medical malpractice or unfortunate accidents with very serious consequences (death, amputation of a wrong limb or severe damage). From this standpoint, patients' chances to increase safety themselves are diminished.

At another level, representatives from consumers' associations declared that the number of complaints and lawsuits targeted at other sectors (cars, construction, hotel industry, etc.) are much more frequent and, therefore, fail to stir the public conscience. Still, they recognized the different value that a medical error had as opposed to these other kind of lawsuits.

The general idea deduced from the discussion was that the provision of health care services had an adequate quality level and that clinical safety depended on the degree of confidence in the physician. The so-called expert patient (or competent patient) profile (Donalson, 2003) was mentioned as an alternative to involve patients in their health. Curiously, it was also mentioned that a way to build confidence was to improve health care centers' accessibility and comfort conditions.

Probing the patients' concept of AE's a little further, we conducted a study in Saint Joan University Hospital (Alicante, Spain) in which 89 patients described, after discharge, the complications they had experienced during their hospitalization. In this case 11.2% of patients answered that they thought they had experienced an AE. However, when describing this AE as a consequence of an intervention, pharmacological treatment, assistance received, etc. we found a very wide variance in our cases.

A qualitative description of these supposed AEs reflects that the patients' idea is far from the professionals' perception of AE's. Thus, for example, we find patients who consider they had suffered an AE but in fact, what they describe are communication problems with their doctors, lack of access or difficulty in choosing a physician.

In a second group, we find patients who describe what we would consider an AE, but they have not experienced it as such. For example, we have the case of a woman who describes gluteal burns from an inappropriate position during delivery.

In a third group we included patients who referred an AE and described a compatible situation (undergoing a second intervention some hours after a prostatectomy with subsequent loss of erection capacity; or the case of a baby who as part of anti-anemia therapy, received a drug to which it was allergic) according to the records.

The results yielded by this study seem to indicate that patients often blame themselves for inadequate treatment reactions, or hold themselves accountable for not achieving certain results after a surgery. On the other hand, they clearly hold professionals responsible for more noticeable failures, like the ones mentioned above.

In a second forum organized with the same intention of exploring the health care context from their actors' perspectives, we had professional experts from different specialties.

When asked about what measures could be put to practice to obtain patients' involvement with clinical safety, the participating experts provided a total of 14 specific proposals. After the discussions, the experts' favorite ones were:

- Set up a campaign to help patients overcome current barriers and encourage them to ask questions of physicians and nurses. They even offered a motto for the campaign: Ask Us.
- Raise professionals' awareness of patients' rights, including the right to choose facilities and the right to information.
- In relation to the above, promote patients' participation in the clinical decision making process.
- Prepare protocols with the information to given the patient and specifically the information providing clinical content.
- Offer additional information about drug therapies focusing on possible allergic reactions.
- Describe to the patient the most common signs of alarm, according to the diagnosis and therapy prescribed.
- Inform patients about "what may be expected" in the most frequent processes, so that patients may be sentinels of their safety process.
- Publish informational support material specifying the health care center commitment to safety to raise patients' and professionals' awareness on this issue.

Our results about AE's as perceived by patients

In the last two years and in cooperation with the General Bureau of Quality and Assistance to Patients of the *Consellería de Sanitat* of our autonomous community (with responsibilities transferred on matters of planning, organization and rendering of health care services) we have been conducting a series of polls to know the opinion of patients about the health care they received during their stay in public hospitals. In the opinion analysis performed in 2005 and 2006, a question has been included to try to indirectly identify possible AE occurrences.

This information allows us to carry out an analysis of patients' perceptions of therapy complications and to analyze the relations between information, physician's accessibility, patient satisfaction and the perception of being a victim of an AE.

In the first of these studies, conducted in 2005, 12,389 patients were asked, through surveys carried out upon discharge, whether during their hospitalization, they had found it hard to talk with their physician, whether they had been able to ask their

physicians questions or tell them their concerns, whether they had been informed of possible complications and of precautions to be adopted to properly follow the treatment, and whether they had experienced unwanted or unexpected effects as a consequence of treatment. Additionally, they were asked a direct question to assess their satisfaction with the outcome of the medical intervention and another question to assess their overall satisfaction with the health care received during their hospitalization.

Patients admitted to medical units reported a frequency of possible AE's almost one percentage point higher than for surgical patients (4.8% vs. 3.7%). In our case did not find differences in the incidence of AE reports between men and women. Data tend to suggest (although with no statistical significance) that at an older age there is a higher frequency of AE's (an incidence of 4.7% in patients younger than 40 and 6.2% in patients older than 60).

A patient's dissatisfaction with the care received was directly related to the likelihood of suffering an AE ($p = 0.0001$). According to our data, if a patient negatively assessed the outcome of the medical intervention, their likelihood of suffering an AE was almost 4 times greater.

As regards accessibility and information variables provided by the physician, we found that among patients dissatisfied with the physician's accessibility, the number of potential AE's was 6 times higher; and among those dissatisfied with the information provided by their physician, the likelihood of AEs was 37% higher. When patients declared they had been able to clarify their questions with the doctor, the number of potential AE's reported by patients was reduced 8 times.

The following year, 2006, we surveyed 2,197 surgical outpatients (Mira, 2006). 71.3% were older than 40 and 29.4% were older than 60. In this case we analyzed together with perception of surgery-associated complications, the existence of possible barriers to talking with the surgeon and the ease in contacting health care staff once at home. Both variables were identified as essential for patients of this kind of surgery (Marchal et al., 2005). As an additional precaution (control of possible bias before considering these data valid) we also analyzed whether, as could be expected, there was a close relationship between patients' reports of having suffered an AE and that the surgery outcome would prevent them from going home on the day previously agreed upon.

In this study, 2.5% of surgical patients reported having been victims of some kind of AE, which almost coincides with the 3% obtained in a previous study of medical records review at one of the hospitals participating in the study. Clearly, the percentage of patients who suffered one of these AE's was significantly higher among those who were not able to go home on the intervention day, as had been predicted (52.5% vs. 1.5% ($\chi^2 = 425.55$, $p < 0.001$).

With regard to the relationship between the barriers that prevented talking to the physicians and the frequency of AE's, the data are crystal clear. The percentage of patients who were unable to talk with their physician to clarify their questions and reported suffering an AE reached 31.1%, while the percentage of patients who suffered an AE but were able to talk with their physician when they needed it was 1.5% ($\chi^2 = 261.93$, $p < 0.001$).

In line with these results, the percentage of patients who experienced an AE was much higher when they could not contact health care staff to solve problems once at home (32.3% vs. 1.6% ($\chi^2 = 236.32$, $p < 0.001$).

Main findings to date. What are we currently working at?

We are currently conducting two research projects financed by the *Fondo de Investigaciones Sanitarias* (Sanitary Research Fund) from *Instituto Carlos III* (Charles III Institute, the Ministry of Health Research Agency). We are working on two directly related topics: patients' participation in clinical decisions and safety perception in hospitals.

In the first study our starting point is the acknowledgement of patients' rights, first in the Charter of Rome and, secondly, in Law 41/2002 (Patients' Autonomy Law) that governs patients' participation in clinical decision-making in Spain.

It is apparent that health care occurs in a context with an obvious "asymmetry of information" between the professional and the patient, but it is also apparent that clinical safety appears as a paramount goal *(primum, non nocere)* of this relation.

In this project we started off with the idea that many patients want a bigger role in clinical decision-making. Now, we pose two issues. First: all patients want that level of involvement. Second: this greater autonomy should go hand in hand with greater responsibility and patient involvement in decisions and, therefore, in their obligations and in their identification of risk factors, complications, etc., by working together with the professional to achieve greater clinical safety.

Our starting hypothesis links professional practice styles with exercise of the patient's right to information and the right to participate in decision-making. Thus, practice styles that are very directive make all the responsibility for the patient's clinical safety fall on the professional. In these cases the physician often shows a defensive kind of practice in response to this greater responsibility. On the contrary, a more open style that makes it easier for the patient to ask questions and participate in the decision-making involves the patient, and the physician is not the only one accountable.

Among the objectives of our study are an analysis of the degree of knowledge of the Patient Autonomy Act, barriers patients face in participating in clinical decision-making, getting to know the professionals' opinion and in what cases they consider patient involvement as most advisable and not advisable, and assessing the current level of patient participation in clinical decisions.

This is a joint project of the *Universidad Miguel Hernández* (Miguel Hernandez University) and the Fundación del Hospital Alcorcón (Alcorcon Hospital Foundation) and we expect the first results late in 2008. The goal of the project is to put forth proposals to improve clinical safety for patients and legal security for professionals. The project development combines qualitative and quantitative research techniques.

In our second project our aim is to analyze in detail whether patients know the risks associated with medical and surgical treatments and their perception of the safety level in the public hospitals of our country, and to determine the strength of the association between the AE frequency perceived and patients' satisfaction with the treatment outcome.

There are currently 5 hospitals of different sizes participating in this trial, located at different geographical points in Spain.

Bearing in mind the previous results we have mentioned, we expect to confirm that Spanish patients trust their public hospitals and that this trust is closely linked to the confidence in their professionals.

In general terms, we suspect that patients' identification of AE's will be lower than the AE rate identified in screening trials of medical records, and we pose the question of whether there are crossed effects in the perception of patients from medical units and surgical units.

The preliminary results of this study will be available at the end of this year, although we expect to have more conclusive results by mid-next year.

What can we conclude?

The analysis of errors made in the health care context contains a negative load that prevents the implementation of proactive behavior to improve patients' safety.

One out of ten patients admitted to hospital suffer some harm or have a longer hospital stay as a result of an AE. Half of these events could have been avoided if there had been a better use of the knowledge derived from the reporting of error occurrence.

European citizens show themselves as moderately concerned about medical errors, and 9% believe they could suffer an AE when admitted to an EU hospital.

Patients' assumption of a more active role could be a positive change for the sake of clinical safety. However, there are barriers attributable to professionals and to the patients themselves that make a safety increase harder to obtain. In the first case, we could name styles of very directive practice or defensive medicine that prevent learning once it has been detected that something went wrong. In the case of patients, their perception is that whatever they may do, it will be to no avail.

AE's occur as a consequence of health care practice and are, in many cases, inevitable. Learning from mistakes to prevent them is a much smarter strategy than the more usual strategy of denial. To accomplish it we need a change of attitude by professionals and patients, a framework that guaranties legal safety for professionals and means so that AE victims are duly informed and compensated. The Apology Act, recently passed in 19 U.S. States, may be a good starting point, and from its implementation we may learn appropriate formulae for each legal and professional framework.

Acknowledgements

The Spanish Carlos III Institute (FIS) is supported our study on Patients' Perception of Hospital Safety (reference PI06/90043).

References

Adams, R. & Boscarino, J. A., 2004. A community survey of medical errors in New York. *International Journal of Quality in Health Care,* 16, p. 353-362.

Agency for Healthcare Research and Quality, 2000. Questions are the answer [internet] Rockville, MD: AHRQ. US Department of Health & Human Services. Available at: http://www.ahrq.gov/questionsaretheanswer/index.html [accessed 12 March 2007].

Aranaz, J. M., Aibar, C., Gea, M. T. & León, M. T., 2004. Los efectos adversos en la asistencia hospitalaria: una revisión crítica. *Medicina Clínica* (Barcelona), 123, p. 21-25.

Aranaz, J. M., Aibar, C., Vitaller, J. & Ruiz, P., 2006. Estudio nacional sobre los efectos adversos ligados a la hospitalización: ENEAS 2005. Madrid: Ministerio de Sanidad y Consumo.

Arora, N. K., Allanan, J. Z. & Guadagnoli, E., 2005. Examining the relationship of patients' attitudes and beliefs with self-reported level of participation in medical decision-making. *Medical Care,* 43, p. 865-872.

Badía, X., Magaz, S., Guitierrez, L. & Guilera, M., 2005. Información de medicamentos de preinscripción: encuestas a la población general española. *Atención Primaria,* 36, p. 93-99.

Bandura, A., 1977. Self-efficacy: toward a unifying theory of behavioural change. *Psychological Review,* 84, p. 191-215.

Barca, I., et al., 2004. La información al paciente y su participación en la toma de decisiones clínicas. *Atención Primaria,* 33, p. 361-367.

Barr, D., 1955. Hazards of modern diagnosis and therapy: the price we pay. *JAMA,* 159, p. 1452.

Blendon, R. J., et al., 2002. Views of practicing physicians and the public on medical errors. *New England Journal of Medicine,* 12 (347), p. 1933-1940.

Borrell, F., et al., 2006. Errores clínicos y eventos adversos: percepción de los médicos de atención primaria. *Atención Primaria,* 2006; 38: 25-32.

Brennan, T. A., et al., 1991. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. *New England Journal of Medicine,* 324, p. 370-376.

Commission Européenne, 2006. Les erreurs médicales. [online] *Eurobaromètre Special.* 241. Janvier 2006. Available at: http://europa.eu.int/comm/health/ph_information/documents/eb_64_fr.pdf [accessed 12 March 2007].

Coulter, A. & Magee, H., 2003. The European patient of the future. Buckingham: McGraw Hill Education. Open University Press.

Donalson, L., 2003. Expert patients usher in a new era of opportunity for the NHS. *BMJ.* 326, p. 1279-1280.

Ende, J., Kazis, L., Ash, A. & Moskowitz, M. A., 1989. Measuring patients' desire for autonomy: decision making and information-seeking preferences among medical patients. *Journal of General Internal Medicine,* 4, p. 23-30.

Evans, S. M., Berry, J. G., Smith, B. J. & Ester, A. J., 2006. [online] Consumer perceptions of safety in hospitals. *BMC Public Health,* 6, p. 41. Available at: http://www.biomedcentral.com/1471-2458/6/41/prepub [accessed 12 March 2007].

Eysenbach, G., 2003. The impact of the Internet on cancer outcomes. *CA: A Cancer Journal for Clinicians,* 53, p. 356-371.

Guadagnoli, E. & Ward, P., 1998. Patient participation in decision-making. *Social Science and Medicine,* 47, p. 329-339.

Hickson, G. B., et al., 2002. Patient complaints and malpractice risk. *JAMA,* 287, p. 2951-2957.

IOM, 2007. Crossing the quality chasm: a new health system for the 21st century [internet] Available at: http://www.iom.edu/CMS/8089.aspx [accessed 12 March 2007].

Jadad, A., Rizzo C. & Enkin M., 2003. I am a good patient believe it or not. *BMJ*, 2003; 326: 1293-1294.

Jarvis, J. & Frizelle, F., 2006. Patients' complaints about doctors in surgical training. *The New Zealand Medical Journal,* 119 (1236), p. U2026.

The Kaiser Family Foundation, 2000. National Survey on Americans as Health Care Consumers: An Update on the Role of Quality Information. Menlo Park, CA: The Kaiser Family Foundation. Agency for Healthcare Research and Quality, 2000 [internet]. Available at: http://www.kff.org/kaiserpolls/3093-index.cfm [accessed 12 March 2007].

The Kaiser Family Foundation, 2000. *National survey on Americans as health care consumers: an update on the role of quality information.* Menlo Park, CA: The Kaiser Family Foundation. Agency for Healthcare Research and Quality, 2000. Available at: http://www.kff.org/kaiserpolls/loader.cfm?url=/commonspot/security/getfile.cfm&PageID=13571 [accessed 12 March 2007].

Leape, L. L., et al., 1991. The nature of adverse events in hospitalized patients: results of the Harvard Medical Practice Study II. *New England Journal of Medicine,* 324, p. 377-384.

Marchal, F., et al., 2005. Post-operative care and patient satisfaction after ambulatory surgery for breast cancer patients. *European Journal of Surgical Oncology (EJSO),* 31, p. 95-499.

Ministerio de Sanidad, 2004. Barómetro Sanitario 2004 [internet]. Available at: http://www.msc.es/Diseno/sns/sns_sistemas_informacion.htm [accessed 12 March 2007].

Mira, J. J., Pérez-Jover, V. & Lorenzo, S., 2004. Navegando en Internet en busca de información sanitaria: no es oro todo lo que reluce... *Atencion Primaria,* 33, p. 391-399.

Mira, J. J. & Rodríguez-Marín J., 2001. Análisis de las condiciones en las que los pacientes toman decisiones responsables. *Medicina Clínica (Barcelona),* 116, p. 104-110.

Mira, J. J., 2006. El paciente competente: una alternativa al modelo paternalista. *Monografías Humanitas,* 8, p. 111-123.

Mira, J. J., 2006. Implicaciones para la cirugía mayor ambulatoria de las medidas de calidad percibida por el paciente. Madrid: Hospital Santa Cristina de Madrid. II Curso de organización y gestión clínica. Febrero 2006.

Mostaza, J. L., Muinelo, I., Teijo, C. & Pérez, S., 2005. Prevalencia y gravedad de efectos adversos durante la hospitalización. *Medicina Clínica* (Barcelona). 124, p. 75-79.

Murff, H. J., et al., 2006. Relationship between patient complaints and surgical complications. *Quality and Safety in Health Care,* 15, p. 13-16.

Robinson, A. R., et al., 2002. Physician and public opinions on quality of health care and the problem of medical errors. *Archives of Internal Medicine,* 28 (162), p. 2186-2190.

Universidad de los Pacientes, 2007. Kit de visita médica [internet]. Barcelona: Universitat Autònoma de Barcelona. Available at: http://www.universidadpacientes.org/aula-es1/ [accessed 12 March 2007].

Van Dalen, I., et al., 2001. Motives for seeking a second opinion in orthopaedic surgery. *Journal of Health Services Research and Policy,* 6 (4), p. 195-201.

Vitaller, J., 2004. El nuevo escenario de la gestión de la asistencia sanitaria y su aplicación en la seguridad de los pacientes: relaciones profesional y paciente. En: Aranaz, J., Vitaller, J., ed., 2004. De las complicaciones y efectos adversos a la gestión de los riesgos de la asistencia sanitaria. Valencia: EVES, 2004.

Wilson, R. M., et al., 1995. The quality in Australian Health-Care Study. *The Medical Journal of Australia,* 163 (9), p. 458-471.

PART 5
WRITING THESIS AND PUBLICATION

How to Write-up your Research, Turn your Data into a Publication and Get Published

JOHN ØVRETVEIT*

Introduction

Considering how to get your research published is mostly about meeting basic scientific requirements and presenting the research in the best way to the reader. This chapter is therefore also relevant to researchers who do not aim to publish. It also highlights how important it is to look ahead and plan your research to ensure it is limited to a specific subject and will contribute to knowledge and action in that subject.

Research is judged in the academic field in terms of whether it is published, and in which journals it is published in. The number and quality of a researcher's or a research unit's publications decides appointments, promotion, whether research grants are awarded and the power of a unit within an academic organization. For better or worse, increasing productivity and quality of publishing is the key to success. This note concentrates on how to publish in scientific journals.

Most data and ideas are not published, often for good reasons. However, most researchers and research units have data and ideas which could be published but which lie unused because of lack of time and know-how about how and where to publish. Working on these data and ideas develops skills to do better research in the future.

Structuring the material for presentation in a paper and getting peer reviews is a powerful way to learn how to do better research next time — "if only we had done the research in a different way it would be so much more publishable!"

The purpose of this chapter is to:

— encourage researchers to work-up the data and ideas you already have into a publishable form;
— give guidance about to how to structure the paper and get it published.

* Director of Research at Medical Management Centre, Karolinska Institutet, Stockholm, Sweden (jovret@aol.com).

It does not address the question of making time to do so. "No time" is often because researchers do not know how best to write the paper, or do not have the motivation, or fear confronting their limitations, do not know a suitable journal, or feel it may be a waste of time because no one will publish it, or do not realize how important publishing is, or all of these. The answer to all of these is that it is better to address "the publication issue" sooner rather than later — better to know whether you have what it takes to be a researcher and want your career in this field than to discover these years later. Writing is thinking. Publishing is presenting the final organized version of your thinking.

Writing the paper

Step 1: overview

One way to start is to make an electronic file (or a handwritten note) with the main headings of the paper. Then write a few words under each heading and move on to the next heading, and back to earlier headings: do not spend longer than ten minutes on one heading because more important at this stage is how the items under each heading relate to each other. This helps focus the main theme of the paper.

TWO GENERAL PURPOSE STRUCTURES FOR A PAPER

For a study involving empirical data
Title
Abstract
Introduction
Previous research
Methods and design for data collection and analysis (details in appendix)
Presentation of findings (details in appendix)
Discussion (last section of this is limitations of the research)
Conclusion
References
Appendices

For a conceptual paper
Title
Abstract
Introduction: first statement of the problem/question. Why it is important — what difference it would make if we had a better way to... etc.

Previous research (and strengths and limitations)
Proposed new idea/approach/key point. Explanation. What this contributes and what difference it would make
Discussion: limitations of this idea. Proposals for future research or debate on the subject
Conclusion
References

Seven types of paper are described by the Journal of Health Organisation and Management, each of which would need a different type of structure: Research paper, Viewpoint, Technical paper, Conceptual paper, Case study, Literature review, General review.

Many papers I review are first drafts, where the person has written down everything they can think of on the topic, but there is no logical linking between the parts and many different questions and subjects being mentioned. Usually you have to decide which is the main theme you will concentrate on and put the first draft to one side to go back to later if you need to. Then cut all the material which does not relate to the main theme and central question and see what the paper looks like. How do you decide the main theme or central question? By choosing your "strongest" data (valid, reliable, sample, and connected with an important question) — everything before and after you present your data must relate to the data you will present.

Making this overview using the headings helps you to ensure that the paper does these two most important things:

- *Vertical linking:* what is written under each heading of the paper should follow-on from the one before and link into the next heading. It is this which shows how the research question follows from the review of previous research and leads into the methods for collecting data to answer the question, the analysis and findings presentation (to answer the question), and the discussion and conclusions.
- *Horizontal linking:* how the research relates to previous and future research. The previous research section should review studies in the field addressed, and the discussion should show how the findings are similar or different to others or test theories or fills gaps.

Step 2: decide structure and style for chosen journal

Making this overview draft focuses your paper and shows you if you have something publishable. The second step is to decide whether to keep this general

structure or use one which is used or recommended by the journal you are aiming for. This step also involves getting clear the style you need to use for the journal. "Style" is a combination of these things:

- *Detail of explanation:* journals differ in the level of knowledge they assume of their readers about different subjects. Some assume little knowledge, some do not want detail but for you to give references for items for the reader to follow up if they want to know more. You need to know the readers and the items you need to give details about and which items you can assume they know.
- *Conciseness:* conciseness is giving only the detail required but also using sentences and phrasing which use the fewest words and use active words. The most concise style in journals I know is the British Medical Journal, and The Economist. Other journals are more forgiving of repetition and allow longer papers. Generally a first draft can be up to 10,000 words, but a publishable paper is usually 2000-6000 words in most journals. Journals vary in the number of references they allow and want.
- *Objectivity:* all scientific journals require a scientific style, where you as an author are dispassionately reporting what you have observed and are open minded about whether the observations and conclusions are correct — this applies to a conceptual as well as to an empirical article. Never use "I", be careful about using "very", "much" or emphasis unless it is really justified and follows after you have presented the evidence. Reviewers are allergic to any paper which gives even the slightest suggestion that the writer has a crusade or even an opinion, and is just using the data to get their cause into print. Objectivity also includes an honest discussion of the limitations and alternative explanations for the findings. However, journals do vary in how rigorously objective a style they require.

Step 3: fill in the headings

Having decided the best headings and style for the paper the next step is to start filling in the details under each heading. Aim for one page per heading, but two for findings and discussion. Put details in an appendix and add them later to the body of the text if you need to. At this stage do not take time filling in references but use brackets to note reference (Davis et al.) or that you will need a reference to give more details or to substantiate your point — mark these ( ) so you can find and fill them in later.

If you have not already done a review of previous research you need to do this earlier rather than later. The section after the introduction presents previous research, and your discussion section shows how your study is similar and different to what has been reported. Your review will help you decide what your study contributes to what is already known: typically new empirical data where there was none, or data which supports, disproves or questions what was previously thought. Looking at previous studies helps you decide what the main focus and question of your paper should be, and also which journals to aim for. It also helps you see the style and structure used for this subject before. Journals differ in how much of a review they want — set a time and words limit for this and add to it later.

Presenting the findings concisely, but giving the reader enough detail to see if the conclusions do follow from the findings is always a challenge. It is particularly difficult to present qualitative data — often qualitative papers present the analysis of the data at a high level of abstraction. At least some quotes are needed to show both the range of extremes and the typical responses.

Use examples and illustrations to explain and make things clear, but also to keep the paper interesting, so that people see the relevance, and to test whether you are saying anything useful.

When to write the abstract? Some do it last of all. Some do it right at the beginning, to outline the paper and see how it could all fit together.

Step 4: revising

A paper should go through many draft versions. At least three or four times, before you ask a colleague to give comments. Always leave a week after working on a paper before doing a redraft. The more time you leave after drafting before doing a revision, the more you can see the paper as if someone else wrote it and the less determined you will be to keep ideas which seem precious to you but not to anyone else (or which you have not made clear).

Getting another persons comments is essential to writing a paper. As them to tell you straight and clearly what is good and what needs to be changed and do not be defensive. Do not insult them by arguing with them — just listen (write it down) and reflect on what they say.

Revising is of two types: the first is to get the theme clear and the vertical linking so that each part follows from the last and leads into the next. This involves ruthlessly cutting out side-stories or items which do not clearly contribute to the main theme/ question. Less is more and anything which does not clearly contribute actually damages the paper. Unlike TV drama, a series of parallel stories do not add to the

interest of the paper. The second type is polishing by cutting out repetitions and shortening lengthy sentences or those with more than one idea in them.

Use the seven scientific requirements questions above to decide what you need to change and add each time you revise.

SUGESTIONS

Get to know the type of articles they publish and how difficult to get accepted.
Have an A and B status list of journals and a plan A, plan B and plan C for submitting a paper.
Use examples and illustrations.
Define terms.
Expect rejection, but the reviewers' comments will be very helpful.
Offer to do reviews yourself.
Write headings, decide length of each, and then fill in. Put aside for at least a week then draft again.

Step 5: submitting the paper

You are ready to submit when the paper is as short as you can get it, with one clear single theme and question running from beginning to the end, and as many of the above scientific requirements have been met. At an early stage you should have made your list of target journals: decide which to submit to and have a plan B and C ready: expect the journal to either reject outright or require revisions and your plan B is which journal then to submit to if rejected or revisions are too demanding or impossible to meet. The review can take anything from 2-6 months before you hear. Reviewers usually demand adding a number of items, as well as shortening the paper!

Finding a journal and meeting journal requirements

You may already have decided which journal to aim for right at an early stage. Even if you have, the work of reviewing previous research will take you to other journals which might publish your paper.

You need to know the preferences of different journals for papers in terms of: subject area, methods bias, and degree of scientific rigour demanded. Some journals are happy to publish qualitative small sample studies, and these journals vary in the level of scientific rigour and evidence presentation demanded of the qualitative study.

Journals vary along a spectrum from "practitioner only" to "researcher/academic only". Journals also vary along a spectrum of "low scientific demand" to "rigorous scientific demands", and the latter high status scientifically. Some journals which have mostly practitioner readership also have rigorous scientific demands, such as some clinical journals (e.g. British Medical Journal, Quality and Safety in Health Care) Some journals which are entirely or largely for and read by practitioners (clinicians or managers) often publish papers which do not meet such high methods rigour (e.g. small sample, simple before after study) or which do not give explanation or prediction (e.g. Journal of Health Organisation and Management). Journals' demands do change over time.

Get a feel for the style and type of paper the journal tends to publish by looking at a few editions. But also go through the guidance to authors section in the Journal and follow these requirements closely.

Conclusions

The above gave some guidance about how to write and publish research. We can conclude by summarizing the above in a reverse form: these will gurantee your research is not published and are common mistakes

— Over 6000 words (if it is still an early draft in search of a focus and probably has two or more objectives it is trying to fulfill).
— Not focusing the paper on one limited subject or question — trying to cover too wide an area or different subjects so that the reader gets confused about the main point or subject.
— No clear linking between problem, research reviewed, research question, methods and design, findings, discussion and conclusion: irrelevant material or new ideas introduced not related to the single theme.
— Repetition.
— Too-little or too-much detail, typically about data collection methods, findings, and background.
— Poor linking to previous research, important previous research missing.
— Limitations or alternative explanations not discussed sufficiently.

— Title, abstract or conclusion does not describe what was presented.
— Referencing method to cite other research not followed consistently.

FURTHER NOTES ON PUBLISHING YOUR WORK

- Ways of knowing and evidence:
- What is the primary knowledge base you are drawing on and contributing to? For scientific publication: your work has to contribute to a body of knowledge, and one which is normally covered by the journal you are targeting.
- Do you know this literature? Can others help you overview and guidance to the many different sources where knowledge is published on this subject and to which you can relate your work?
- More systematic documentation and evaluation as part of your projects could produce good enough data for scientific articles. There are different types of evidence: experience, individual knowledge, documented learning, collective learning, research.
- But a medical journal will accept a publication which points to a non-medical body of knowledge which has important implications for medical science and practice.

Checklist:

Basics: types of article, typical article headings, style of writing and boxes, references, spacing etc.

Important problem/issue scientifically/practically

Of interest to readers.

Links to other research.

Defined question/hypothesis.

For the question/hypothesis it uses accepted data gathering and analysis methods and design.

Methods described and appear to be used in proper way.

Findings presented in a way which allows independent assessment.

Discussion relates findings to others research or lack of it.

Self critical listing of limitations.

Conclusions follow from the findings.

Logical linking structure from first section through to end.

Some want practical implications or even recommendations spelt out (so what? what difference would this make?).

References

Hall, G., ed., 1994. *How to write a paper.* London: BMJ Publishing Group.

Gustavii, B., 2004. *How to write and illustrate a scientific paper.* Cambridge: Cambridge University Press.

Moxley, J., 1992. *Publish don't perish.* London: Greenwoord Press.

University of Washington, 1997. *Writing a literature review* [online]. Washington, D.C.: University of Washington (published 1997). Available at: http://depts.washington.edu/psywc/handouts.html [accessed 20 October 2007].

Reliability and Validity of the Hospital Survey on Patient Safety Culture at a Norwegian Hospital

ESPEN OLSEN*

ABSTRACT

Main research question: The core aim of the present study was to translate and test the validity and reliability of the Hospital Survey on Patient Safety Culture (HSOPSC) in Norwegian health care.

General design and sample: The study was conducted among employees at a large hospital in Norway that offers a wide range of hospital services organized into 10 clinics.

Data collection and analysis methods: HSOPSC (Sorra & Nieva, 2004) was translated into Norwegian and used to measure the safety culture among the main target groups — namely, health and mercantile workers employed in the health care environment; 1919 questionnaires were returned, resulting in a response rate of 55 percent. Confirmatory factor analysis (CFA) was utilised to investigate the fit of the proposed factor structure, and Cronbach's alpha was determined to examine the internal consistency of dimensions. Furthermore, the intercorrelation among concepts and MANOVA were conducted to investigate discriminate validity. Finally, concurrent validity was examined to verify the degree to which the safety culture dimensions influenced the outcome variables included in HSOPSC.

Main findings: Confirmatory factor analyses indicated that the factorial model fitted the data well. One dimension, "Organizational learning — continuous improvement", indicated unsatisfactory internal consistency, although the internal consistency improved when the item "mistakes have led to positive changes here" was removed from the dimension. Contrary to established expectations, the safety culture dimension exerted several negative influences on "Number of events reported (last 12 months)", indicating that this outcome

* Risk Management and Societal Safety, Department of Media, Culture and Social Sciences, University of Stavanger, Norway (espens.olssen@uis.no).

variable is invalid. "Patient safety grade" and "Overall perceptions of safety" seemed to be the most valid outcome variables.
Methodological challenges: The greatest methodological challenges facing the study were using the correct statistical procedures and methods so that Type I and Type II errors were avoided.
Practical challenges: The most practical challenge of this study was achieving a high survey response rate among hospital staff.
Main lessons for other researchers: The study indicated that the psychometric properties of HSOPSC are satisfactory and that the instrument can be used in Norwegian hospital settings. However, users should be aware that the internal consistency is lower for the dimension "Organizational learning — continuous improvement" and that "Number of events reported" is probably dysfunctional as an outcome measure.

Keywords: Safety culture; Safety climate; Validity; Reliability; Safety performance; Patient safety

Introduction

Safety culture has been regarded as one of the most important premises for the further improvement of patient safety in health care (Corrigan et al., 1999). Generally speaking, the survey method appears to be the predominant strategy for studying organizational cultures, and their effects on safety. The term culture is often replaced with climate when questionnaire surveys are used to assess an organizations' culture. The survey method is well suited for studying individual attitudes and values as well as practices — "the way people do things around here" (Hopkins, 2006: p. 878). Interest in measuring the safety culture has generated several instruments for use in health care settings. Such instruments normally incorporate several dimensions; most adopt a "generalist" focus designed to address several safety issues in a variety of hospital areas, while psychometric techniques are commonly used to ensure potential users that instruments will be a good predictor of safety events and provide actionable information (Singla et al., 2006).

Grasping the concept of safety culture is challenging as it is concerned with work practices concerning safety as well as how individuals think, act, and cooperate concerning safety (Cooper, 2000). Confusion within scientific areas often relates to a lack of evidence concerning reliability and validity. As health care safety culture/ climate instruments are increasingly being used on a large scale throughout health care organizations, it is becoming increasingly important to obtain information

about the psychometric properties of such instruments (Flin et al., 2006). Validity concerns come even more into question when questionnaires are translated into other languages, especially since environmental differences might exist at the national level (Hutchinson et al., 2006). These factors make it important to investigate whether or not the dimensional structure of safety culture instruments can be replicated in various organizational and international contexts. To ensure that survey instruments are valid and reliable, instruments developed in one context should ideally always be validated before extensive use in a new context (Pronovost & Sexton, 2005).

The main research question of this paper is to investigate the psychometric properties of Hospital Survey on Patient Safety Culture (HSOPSC) in a Norwegian health care setting. Results from a review of different instruments have shown that HSOPSC, when compared with other instruments, meets more psychometric criteria (Flin et al., 2006). Still, it is unclear if the reliability and validity of the instrument will be replicated in a Norwegian health care setting.

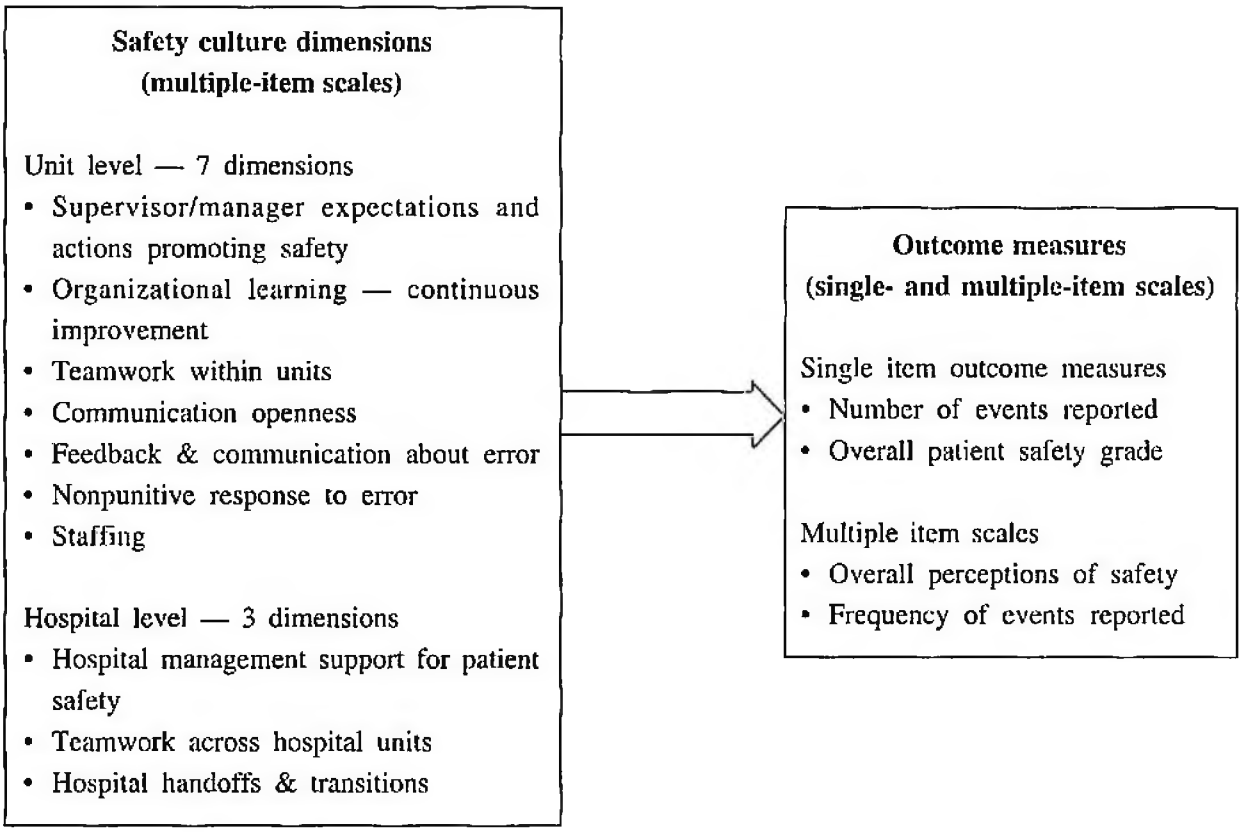
Data collection and analysis meyhods

Selection of instrument

A review of available safety culture instruments was conducted; HSOPSC was selected as the instrument for testing for three reasons. First, the dimensionality of HSOPSC covers general topics revealed as part of a broader patient safety project (Thomassen et al., 2005). Second, as noted earlier, studies show that HSOPSC has met more psychometric criteria compared to other instruments (Flin et al., 2006). Third, benchmark statistics of HSOPSC can be retrieved from the internet (AHRQ, 2007).

HSOPSC was developed based on a literature review, an examination of existing published and unpublished safety culture instruments, and psychometric analyses from the Veteran's Administration Patient Safety Questionnaire and the transfusion safety culture survey. The final version of HSOPSC consists of two single-item outcome measures and two overall patient safety outcome scales that were assessed to validate ten safety culture dimensions (Sorra & Nieva, 2004). All items in HSOPSC are rated on Likert-type scales with verbal anchors. "Number of Events Reported (last 12 months)" is measured on a scale from 1 to 6; all other concepts are measured on scales from 1 to 5. Figure 1 summarizes the safety culture dimensions and outcome variables measured with HSOPSC. Additional details about HSOPSC are available at www.ahrq.gov/qual/hospculture.

Figure 1
Safety culture dimensions and outcome variables measured with HSOPSC



Translation and pilot testing

HSOPSC was translated into Norwegian and then translated back into English by two independent researchers to ensure validity of the translation. In the translation process, it was stressed that the same meaning and "strength" should be reproduced in the translation into the Norwegian language. In order to test if respondents understood the meaning of all items, HSOPSC was pilot tested in a group of eight health care workers.

Distribution of items

The degree of missing values on items and skewness were also used as indicators of usability. The missing criteria were estimated so that no more than 10 percent of respondents would skip items. Skewness was defined so that 85 percent of the sample would not answer on one end (1 or 5/6) of the scales.

Sample

The study was carried out in a Norwegian hospital. The target group included health workers at the hospital and other personnel employed in the same working environment as the health care personnel. A total of 1919 workers answered the survey, resulting in a response rate of 55%. Of these respondents, 89% had direct patient contact, whereas 62% worked between 20 and 37 hours per week. Nurses with or without specialist education represented 45% of the total sample. The pilot testing of HSOPSC in a group of health care workers (N = 8) did not reveal any problematic items. In addition, informal dialogs with health care workers supported the usability and relevance of items to the broader patient safety and safety culture issues.

Analysis of data

Conventional validation strategies were undertaken (DeVillis, 2003; Hinkin, 1995, 1998; Netemeyer et al., 2003) in order to assess the validity of HSOPSC. Utilising CFA with Maximum Likelihood method, the dimensionality was assessed to investigate if all dimensions loaded as expected on their respective items. Items were treated as continuous variables, missing responses were deleted list-wise, and covariation was allowed between dimensions.

As explained by Sorra & Nieva (2004), composite scale scores for the 12 safety culture dimensions were created by obtaining the mean of the responses to items in the dimension after reverse coding of the reverse items. One is the lowest possible score on composite scores, and five is the highest. Several analysis were conducted after negative items were reversed. The Cronbach's alpha was estimated to determine if factor scales yielded acceptable alpha coefficients and internal consistency. Pearson's *r* was estimated to examine the discriminate and convergent validity among measures. MANOVA (Wilks' Lambda) was conducted to examine if different work characteristics had overall effects on HSOPSC concepts (discriminate validity). A regressions analysis was conducted to investigate if the safety culture dimensions influenced the outcome variables as expected (concurrent validity). CFAs were conducted using linear structural relation (LISREL) analysis, whose core aim was to judge the goodness of fit of the factorial model. Structural Equational Modelling made simple (STREAMS) was used for the LISREL analysis (Gustafsson & Stahl, 2000). The remaining results were generated using SPSS 13.0.

Results

Descriptive statistics

Table 1 presents the mean statistics and 95% confidence interval (CI) for each of the measurement concepts. The mean score for all concepts ranged from 1,84 to 3,84; most of the mean scores on the safety culture dimensions were above 3, which is the midpoint of the measurement scales. All items were satisfactory when it came to the missing and skewness criteria, indicating no need to remove any items based on these criteria. Skewness was highest for the variable "Frequency of event reporting"; 45% of respondents did not report any events during the last 12 months.

Internal consistency reliabilities

With one exception, Cronbach's alpha scores ranged from .64 to .82, which are considered satisfactory (Table 1). Only "Organizational learning — continuous

Table 1

Number of items, descriptive statistics, 95% CI and Cronbach's alpha

MEASUREMENT CONCEPTS	NUMBER OF ITEMS IN SCALE	MEAN	95% CI	ALPHA
4 outcome measures				
Patient safety grade	1	3,44	3,41 to 3,47	—
Number of events reported (last 12 months)	1	1,84	1,80 to 1,89	—
Overall perceptions of safety	3	3,50	3,46 to 3,53	.76
Frequency of event reporting	3	2,89	2,85 to 2,93	.82
Safety culture dimensions — unit level				
Supervisor/manager expectations & actions promoting safety	4	3,82	3,79 to 3,85	.77
Organizational learning — continuous improvement	3	3,37	3,34 to 3,40	.51
Teamwork within hospital units	4	3,84	3,82 to 3,87	.77
Communication openness	3	3,71	3,68 to 3,74	.68
Feedback and communication about error	3	3,24	3,21 to 3,27	.70
Nonpunitive response to error	3	3,81	3,79 to 3,84	.64
Staffing	4	3,35	3,32 to 3,38	.65
Safety culture dimensions — hospital level				
Hospital management support for patient safety	3	2,90	2,87 to 2,94	.79
Teamwork across hospital units	4	3,11	3,09 to 3,14	.65
Hospital handoffs and transitions	4	3,21	3,18 to 3,23	.65

improvement" had a lower alpha score (.51); however, the alpha score for this dimension did increase to .60 when the item "mistakes have led to positive changes here" was removed from the dimension[1].

Confirmatory factor analyses

CFA were conducted to determine if latent variables loaded as expected on the observed variables. Widely used goodness of fit indices indicated that the measurement model acceptably fitted the data (RMSEA = 0.044, CFI = 0.97, GFI = 0.91, AGFI = 0.90). Generally speaking, factor loadings were satisfactory. The lowest loading — from "Organizational learning — continuous improvement" with the item "mistakes have led to positive changes here" — was 0.29.

Discriminate and convergent validity

Correlations among the 10 safety culture dimensions varied between .17 and .59 ($p < .01$). It was expected that "Overall perceptions of safety" would be highly correlated with "Patient safety grade". This was supported in the data (.68, $p < .01$). "Feedback and communication about error" and "Communication openness" were the highest correlated dimensions (.59) among the safety culture dimension.

Results using MANOVA (Wilks' Lambda) indicated that all work characteristics (work area, length of services at the hospital, length of services in work area, hours per week, position, and patient contact) significantly ($p <.001$) explained the variance of the HSOPSC concepts.

Concurrent validity

Concurrent validity concerns to what degree phenomenon covariates with other related phenomena at the time of distribution (Netemeyer et al., 2003). Four regression analyses were conducted to determine if the ten safety culture dimensions influenced the four outcome variables as expected (Table 2). "Number of events reported (last 12 months)" was most weakly influenced by the safety culture dimensions. The ten safety culture dimensions had generally positive effects on the other three outcome variables;

[1] In order to test the original factorial model of HSOPS, this item was not removed before conducting the remaining analyses.

Table 2
Regression analyses testing the concurrent validity of HSOPSC

PATIENT SAFETY DIMENSIONS	OUTCOME VARIABLES			
	NUMBER OF EVENTS REPORTED (LAST 12 MONTHS)	PATIENT SAFETY GRADE	FREQUENCY OF EVENT REPORTING	OVERALL PERCEPTIONS OF PATIENT SAFETY
	Beta	Beta	Beta	Beta
1. Supervisor/manager expectations & actions promoting safety	−.03	**.14*** **	.05	**.13*** **
2. Organizational learning — continuous improvement	**.08****	**.13*** **	**.11*** **	**.16*** **
3. Teamwork within hospital Units	**−.06***	**.06***	.00	.02
4. Communication openness	.01	**.08****	−.05	.03
5. Feedback and communication about error	**.11*** **	.03	**.33*** **	.04
6. Nonpunitive response to Error	**.06***	−.04	−.01	**.08*** **
7. Staffing	**−.11*** **	**.27*** **	**−.07*** **	**.36*** **
8. Hospital management support for patient safety	−.04	**.17*** **	**.12*** **	**.18*** **
9. Teamwork across hospital units	−.03	**.09*** **	−.05	**.06****
10. Hospital handoffs and Transitions	−.05	−.00	.01	.02
Explained variance (R2)	.04	.42	.18	.54

* $p < 0.05$;
** $p < 0.01$;
*** $p < 0.001$. Significant coefficients are bolded.

however, contrary to expectations, the results revealed some significant negative influences on the outcome variables. "Hospital handoffs and transitions" was the only variable with no significant influences. The explanatory power was greatest for "Patient safety grade" (.42) and "Overall perceptions of patient safety" (.54).

Discussion

Pilot testing

In order to make HSOPSC useful in a Norwegian health care setting, it is crucial that items be clear and unambiguous to workers. Generally, results did not reveal any problematic items; therefore, it is fair to justify the usability of the Norwegian

version of HSOPSC. In addition, informal dialogs with health care workers concerning the content of HSOPSC provided further important support for the usability of HSOPSC.

Distribution on items

Statistical variation on items is important as it is a fundamental condition required for many statistical tests (Stone, 1978). Moreover, using items with low variance is problematic because doing so takes away the intention of conducting benchmark studies if no variance exists between groups. Therefore, it is important to note that all items were satisfactory in regards to the missing and skewness criteria, indicating no need to remove or adjust any items based on poor distribution.

Internal consistency reliabilities

Previous studies have revealed that using five-point scales on items is advantageous for achieving satisfactory coefficient alpha scores (Lissitz & Green, 1975). In the current study, with the exception of "Organizational learning — continuous improvement", alpha scores ranged from .64 to .82, which is considered satisfactory. Removing the item "mistakes have led to positive changes here" increased the alpha score on "Organizational learning — continuous improvement" from 0.51 to 0.60, which suggests that this item should be removed. However, removing it will reduce the possibility of comparing results on this dimension with benchmark data. Moreover, less than three items for a dimension is usually not recommended (Flin et al., 2006).

Confirmatory factor analysis

Cultural and contextual differences between the United States and Norway made it far from certain that the factorial structure of HSOPSC would be reproduced at the Norwegian hospital. Widely used goodness of fit indices indicated that the measurement model acceptably fitted the data. Results from CFA, therefore, support the argument that the factorial structure of HSOPSC is replicable in a Norwegian health care setting. It is reassuring that the factorial structure of HSOPSC is robust across different cultures and after translation into Norwegian, as this makes it possible to compare studies in Norway with U.S.

benchmark data (AHRQ, 2007). In addition, the robustness of the factorial structure makes it more reasonable to conduct longitudinal studies for measuring change over time on stable factors.

The lowest loading (0.29), from "Organizational learning — continuous improvement" on "mistakes have led to positive changes here", corresponds with the low alpha score (0.51) on this dimension. Therefore, it is not surprising that this dimension loaded relatively low on this item. Loadings below 0.30 are not considered optimal because it means that less than 9 percent of that item's variance is shared with the factor (Comrey and Lee, 1992).

Discriminate and convergent validity

Composite scores on the 10 safety culture dimensions ranged from 1.0 (lowest) to 5.0 (highest). The HSOPSC dimensions all measure various aspects related to the phenomena of safety culture. It was expected that composite scores on the safety culture dimension would correlate to some degree. However, correlations should not be too high as this will indicate that dimensions measure almost the same concept (Sorra and Nieva, 2004) and show low evidence of discriminate validity (Hinkin, 1998).

Correlations among the 10 safety culture dimensions varied from between .17 to .59 ($p < .01$). These correlations are considered satisfactory and do not indicate problematic associations among dimensions. The strongest correlation was between "Feedback and communication about error" and "Communication openness" (0.59). Considering that both dimensions share some attention towards communication, this outcome was not surprising; because it was conceptually meaningful, these concepts were not integrated into one concept.

Similarly correlating constructs support the evidence for convergent validity (Hinkin, 1998). It was therefore expected that "Overall perceptions of safety" would be highly correlated with "Patient safety grade", as in the pilot study conducted by Sorra and Nieva (2004). This finding was reproduced in the current study (.68, $p < 0.01$). "Overall perceptions of safety" and "Patient safety grade" are highly associated concepts; the high correlation between these concepts indicates convergent validity for both concepts.

Results using MANOVA provided further support for the discriminate validity of HSOPSC as the different work characteristics had generally significant effects on HSOPSC concepts. This is important as it was expected that perceptions of safety culture varied based on different worker and organisational characteristics (Huang et al., 2007).

Concurrent validity

Regression analysis revealed that "Number of events reported" does not function well as an outcome variable. Forty-five percent of the sample did not report any events. Contrary to Sorra and Nieva (2004), we see no reason to believe that the lack of association with this outcome variable is due to a lack of variability or extreme skewness in the number of events reported. A more probable reason is that "Number of events reported" does not capture the actual risk level due to the poor culture of reporting in health care. Nevertheless, the present data concur with Sorra and Nieva's assertion that the best use of this one-item measure is to use it as a change measure in order to monitor if staff members report more events over time.

With a few exceptions, the ten safety culture dimensions had positive influences on the other three outcome variables. This should be interpreted as better scores on safety culture dimensions positively influence safety outcomes: higher levels on the "Patient safety grade", higher "Frequency of (no harm) event reporting", and higher levels on "Perceptions of patient safety".

Some safety culture dimensions negatively influenced the outcome variables. Several reasons can explain this. Perhaps improved staffing will decrease near misses, thereby reducing the need to report and consequently the frequency of reported events. The negative influences on "Number of events reported (last 12 months)" probably relates to the general problem with the use of this measure as a criterion variable.

The general impression is that "Number of events reported (last 12 months)" does not function well as a criterion measure. The consistent influences on "Patient safety grade", "Frequency of event reporting" and "Overall perceptions of patient safety" supports both the validity of the safety culture dimensions and these outcome measures. Most significant influences are associated with the "Patient safety grade" and "Overall perceptions of patient safety"; therefore, these variables seem to be the most valid outcome measures. Meanwhile, "Hospital handoffs and transitions" seem to be the dimension with the lowest explanatory power. The evidence for the validity of this dimension therefore seems weaker.

Limitations of the study

The validity of this study is limited to self-reported outcome variables (concurrent validity), which is not optimal due to the possibility of common method bias (Podsakoff et al., 2003). Some precautions should therefore be taken as HSOPSC dimensions have not been validated against other patient safety indicators, such as actual reporting of adverse events on subsequent occasions (predictive validity). Until

HSOPSC has been validated against other criterion measures, the full impact of resulting data on different organisational risk areas cannot be known.

In the present study, traditional analysis were used in order to test the reliability and validity of HSOPSC. The combination of analysis used is not comprehensive when it comes to assessing the content validity of all concepts. Based on the researchers' experiences, the content validity of HSOPSC seems good with one exception: the outcome variable "Frequency of event reporting". Items included in this dimension concern mistakes made and reported; however, items are limited to any mistake that "is caught and corrected before affecting the patient", "has no potential to harm the patient", and "could harm the patient". In other words, this dimension measures something more limited than a broader frequency of event reporting — namely, the reporting of near misses. Based on these arguments "Frequency of no harm reporting" is probably a more suitable name for this dimension.

Implications and conclusions

The current study's results demonstrated that the factorial structure of HSOPSC was replicated at a Norwegian hospital, and results generally complied with conventional reliability and validity criteria. The examination of new measures in independent samples is important (Stone, 1978) as it gives further evidence for the psychometric properties of measures. Based on this study, the general impression is that the factorial structure of HSOPSC can be generalized and HSOPSC is usable in a Norwegian hospital context.

One dimension showed to have weaker internal consistency: "Organizational learning — continuous improvement" (.51). Based on results from the CFA and the low alpha score, practitioners and researchers should consider removing the item "mistakes have led to positive changes here"; however, practitioners and researchers should also be aware that removing an item will reduce the possibility for exact comparison with benchmark data. The ten safety culture dimensions were less strongly related to the outcome dimensions "Number of events reported (last 12 months)" and "Frequency of event reporting". The "Patient safety grade" and "Overall perceptions of patient safety" seem to be the most valid outcome measures of HSOPSC.

This work has been part of a PhD thesis and the aim of the present study was to translate and test the validity and reliability of the Hospital Survey on Patient Safety Culture in Norwegian health care. In order to achieve this aim, applying all relevant methods are challenging; however, it is important to do so in order to make correct scientific conclusions and avoid Type I and Type II errors. Other researchers should also be aware of the difficulty of achieving high response rates among hospital staff,

which we consider to be the most practical challenge after conducting this study. Future research should explore associations between HSOPSC dimensions and other organizational outcomes in health care settings, combining studies with other methods. Another important research task will be to conduct longitudinal studies (Carroll, 1998). Longitudinal studies will give answers to the stability and changeability of safety culture. Also, by using longitudinal designs researchers can investigate which interventions most effectively improve the levels of safety culture in health care settings.

References

Agency for Healthcare Research and Quality, 2006. *Comparing your results: preliminary benchmarks: Hospital Survey on Patient Safety Culture,* 2006 [internet] Rockville, MD: Agency for Healthcare Research and Quality. Available at: http://www.ahrq.gov/qual/hospculture/prebenchmk.htm [accessed 05 March 2007].

Carroll, J. S., 1998. Safety culture as an ongoing process: culture surveys as opportunities for enquiry and change. *Work and Stress,* 12, p. 272-284.

Comrey, A. L. & Lee, H. B., 1992. *A first course in factor analysis.* Hillsdale, NJ: Lawrence Erlbaum.

Cooper, M. D., 2000. Towards a model of safety culture. *Safety Science,* 36, p. 111-136.

Corrigan, J., et al. (1999). *To err is human: building a safer health system.* Washington: National Academy Press.

DeVillis, R. F., 2003. *Scale development: theory and application.* Thousands Oaks, CA: Sage Publications.

Flin, R., et al., 2006. Measuring safety climate in health care. *Quality and Safety in Health Care,* 15, p. 109-115.

Gustafsson, J. E. & Stahl, P. A., 2000. *Streams user guide: Version 2.5 for Windows.* Mölndal, Sweden: Multivariate Ware.

Hinkin, T. R., 1995. A Review of scale development practices in the study of organizations. *Journal of Management,* 21, p. 967-988.

Hinkin, T. R., 1998. A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods,* 1, p. 104-121.

Hopkins, A., 2006. What are we to make of safe behaviour programs? *Safety Science,* 44, p. 583-597.

Huang, D. T., et al., 2007. Perceptions of safety culture vary across the intensive care units of a single institution. *Critical Care Medicine,* 35, p. 165-176.

Hutchinson, A., et al., 2006. Use of a safety climate questionnaire in UK health care: factor structure, reliability and usability. *Quality and Safety in Health Care,* 15, p. 347-353.

Lissitz, R. & Green, S., 1975. Effects of the number of scale points on reliability: a Monte Carlo approach. *Journal of Applied Psychology,* 60, p. 10-13.

Netemeyer, R. G., Bearden, W. O. & Sharma, S., 2003. *Scaling procedures: issues and application.* London: SAGE Publications.

Podsakoff, P., et al., 2003. Commom method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of Applied Psychology,* 88, p. 879-903.

Pronovost, P. & Sexton, B., 2005. Assessing safety culture: guidelines and recommendations. *Quality and Safety in Health Care,* 14, p. 231-233.

Singla, A., et al., 2006. Assessing patient safety culture: a review and synthesis of the measurement tools. *Journal of Patient Safety,* 2, p. 105-115.

Sorra, J. & Nieva, V. F., 2004. *Hospital survey on patient safety culture.* [online] Rockville, MD: Agency for Healthcare Research and Quality. U.S. Department of Health and Human Services. (AHRQ Publication; 04-0041). Available at: http:// www.ahrq.gov/qual/hospculture/hospcult.pdf [accessed 05 March 2007].

Stone, E., 1978. *Research methods in organizational behavior.* Glenview, IL: Scott, Foresman.

Thomassen, R., et al., 2005. *Pasientsikkerhet: en analyse av feilmeldinger ved SUS [Patient safety: an analyses of adverse event at SUS].* Stavanger, Norway: University of Stavanger (report 7).

Improving Quality through Developing Interprofessional Learning in the Context of Shared Medical Appointments for Diabetes Care: a Research Proposal

SUSAN KIRSH*
RENEE L. LAWRENCE**
DAVID C. ARON**

ABSTRACT:

Research questions: Our underlying research question is: can high quality chronic illness clinical care be combined with high quality learning for trainees? We use the shared medical appointment/group visit (SMA) for diabetes care as our test environment. We have the following specific aims (1) evaluate effect of SMAs on trainees' *attitudes, self-efficacy and knowledge* of diabetes and effectively working in a multi-disciplinary team; (2) evaluate the effect of SMAs on patients' intermediate *clinical outcomes for diabetes (A1c, LDL-cholesterol; and systolic blood pressure)*; and (3) evaluate SMAs as a method to improve the quality of care for a chronic illness such as diabetes in the presence of physician trainees.

Study design: The study design for physician trainees will be a pre-test/post-test design. This will be embedded in a sequential partial cross-over randomized controlled trial of patients that also involves a pre and post-test.

Data collection and analysis: We settled upon: Knowledge Assessment Test, Attitude Survey, P-ACIC/self efficacy survey, measure of interprofessional domains, and a semi-structured interview. Patient outcome measures include A1c, systolic blood pressure, LDL-c. foot and eye exam rate, aspirin use and satisfaction. T tests will be utilized to compare control and intervention clinic groups. T tests will also be used for individual patients/trainees where the outcome measures allow the subject to be his/her own control (i.e., pre-test/ post-test design). Repeated measures will be assessed with ANOVA models and statistical process control.

* Medical Director of Primary Care (Wade Park) and Center for Quality Improvement Research (Susan.kirsh@va.gov).

** Center for Quality Improvement Research.

Main findings: To date we have found that patients A1c and systolic blood pressure are improved compared to those matched controls who have not participated in diabetes SMAs. Initial feedback from trainees, both formal (focus groups) and informal, has been positive.
Methodological challenges: Major challenges were identifying a survey instrument that could most closely match our research hypotheses and determining and obtaining enough exposure to the SMAs to see an effect.
Practical challenges: The major practical challenge was getting enough other health professional trainees to participate. We have been unable to overcome this challenge so we have focused on physician trainees.
Main lessons for other researchers: Focusing on the research and addressing its challenges was difficult, given the amount of time I had to work on research. I had little research experience at project inception and was responsible for administering the large primary care clinic into which these changes were made. Developing and using conceptual models has helped me to work through this in an organized way. This model of high quality patient care with high quality clinical training is exciting to work on because not much exists in this area yet and this sustains my efforts.

Introduction

This chapter was written as the research on shared medical appointments (SMAs), also referred to as a cluster or group visits, reached different stages for different parts of the process — literature review, planning, pilot testing, and initial implementation. Moreover, it should be recognized that this process was not linear. Rather, this process with its multiple PDSA cycles and feedback loops proceeded along several lines simultaneously. However, though all of these influenced and intersected each other, we will use a more conventional approach to relate our progress. My research advisors and mentors worked closely with me in writing this chapter.

Background

Social and demographic changes as well as advances in medical care have resulted in chronic illness care becoming an increasing burden on health care systems worldwide. The American health care system, designed primarily around the need for addressing acute health care issues, is no longer sufficient for treating the growing number of individuals requiring ongoing care for their chronic conditions where

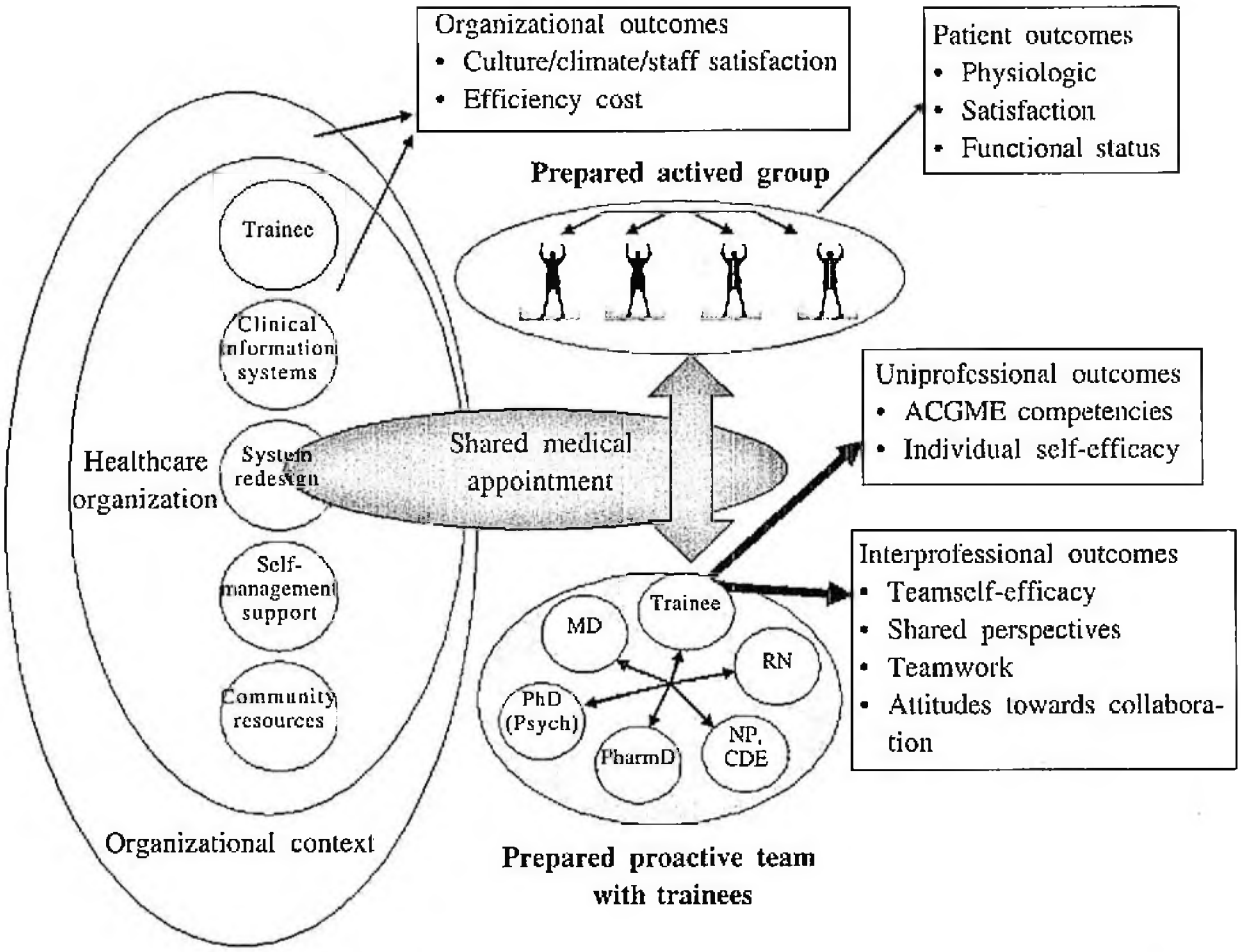
despite much effort, large quality gaps persist. These issues have challenged the Veterans Healthcare Administration in general and its primary care clinics in particular. The chronic care model developed by Wagner et al., uses 6 components and an activated patient to engage providers into an approach that has shown improved outcomes for patients with chronic diseases, e.g., diabetes. The Shared Medical Appointment (SMA) is a type of planned visit, which emphasizes a different delivery system design and utilizes key concepts of the chronic care model to achieve high quality patient-centered care. Studies of SMAs for patients with diabetes for example, have shown improved patient intermediate outcomes of Alc, LDL-c and systolic blood pressure as well as improved patient self-management skills and patient and provider satisfaction (Kirsh, S. et al., in press).

Since 2004 the Veterans Health Administration has promoted (and in fact mandated) the establishment of SMAs. This was done in order to increase clinic capacity and reduce waiting times, while providing high quality multidisciplinary care. SMAs allow group learning and teaching while still providing individual patient goal setting and medication management. In April 2005, the Louis Stokes Cleveland Veterans Administration Medical Center primary care clinic implemented a weekly Diabetes SMA to improve our diabetes care. We targeted those individuals seen by Primary Care Providers (Attending staff physicians, Internal Medicine Resident Physicians, Nurse Practitioners) and who were at highest risk for complications: (poor glycemic and/or blood pressure and/or cholesterol control) and invited them to participate in a SMA. All patients who agreed to participate were informed that this was a group appointment. Our SMA includes between 8 and 20 patients and is staffed by a combination of: 1 attending physician, 1 to 4 residents and/or medical students, 1 nurse practitioner, 1 psychologist, 1 pharmD and 1 RN. Our Shared Medical Appointment is based on the chronic care model and its conceptual model is shown in Figure 1.

Chronic care management needs approaches that educate, sensitize, support, engage, and help nurture activated patients and prepared proactive health care teams. Yet there exists an imbalance between acute and chronic disease in the undergraduate and graduate medical education (Nair & Finucane, 2003). Moreover, medical students begin medical school with positive attitudes toward caring for the chronically ill, but this perception depreciates with clinical experience. (Davis et al., 2001). Graduate trainees also need exposure to chronic care of patients and patient centered care. Their attitudes toward chronic care reveal low comfort levels with reaching patient goals and low satisfaction in caring for these patients. In addition, trainee performance in meeting recommended targets for glycemic control lag behind that of faculty physicians. Both care delivery and clinical training needs redesign. The chronic care model with its six key elements (delivery system design, clinical information system,

Figure 1

Conceptual model the shared medical appointment based upon the chronic care model and how it might serve as a site for training of physicians



Note: We focused on the uniprofessional outcomes for physician trainees as defined by the six core competencies outlined by the Accreditation Committee on Graduate Medical Education and began to conceptualize interprofessional outcomes related to teamwork.

decision support, self-management, community, health care organization) provides a conceptual framework around which the care delivery and even the clinical education system can be redesigned to meet the needs of patients and healthcare professionals at every level of training.

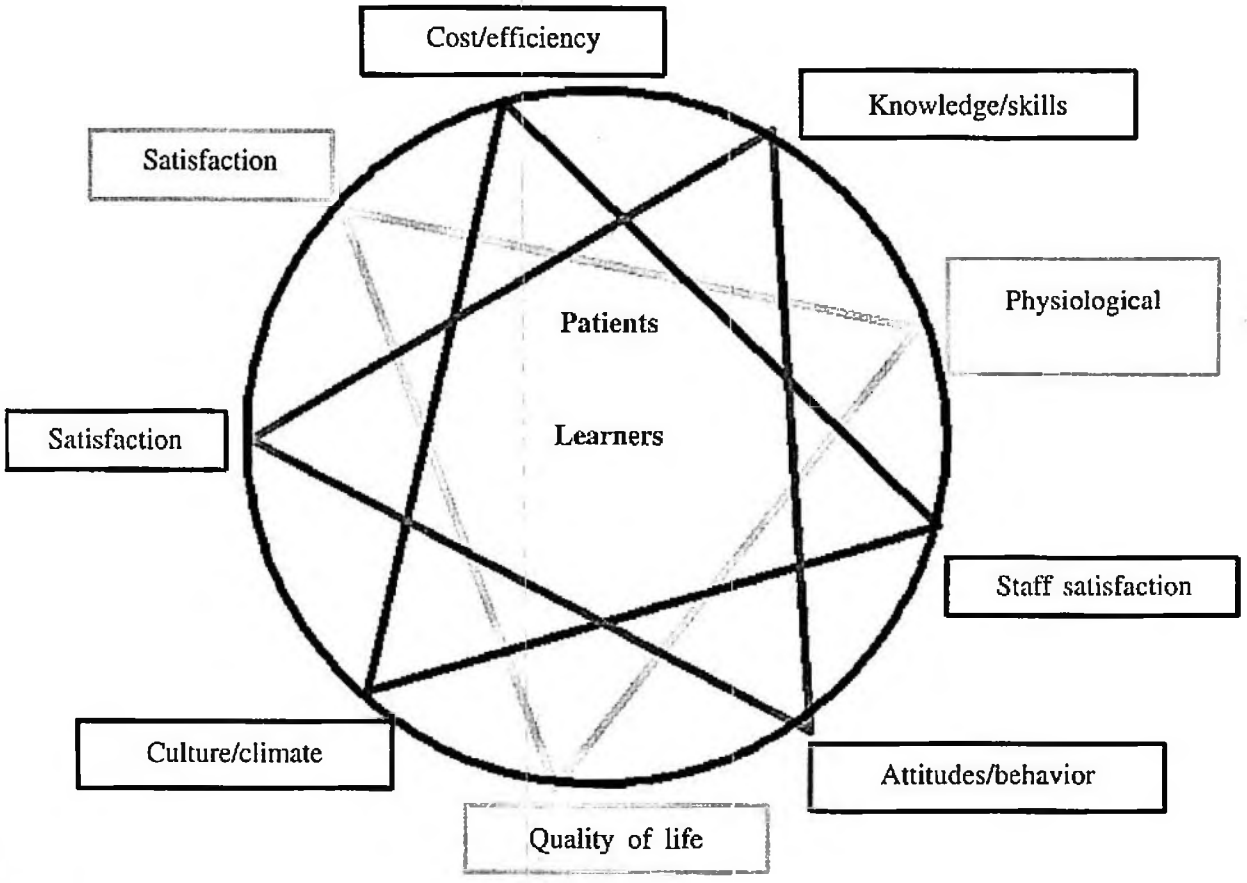
As we began to see successful patient outcomes, we considered using this patient centered multidisciplinary team model as a venue to train physicians. Managing patients with a chronic disease such as diabetes in the context of a multidisciplinary team as well as facilitating a shared medical appointment requires skill sets not adequately provided in current medical curricula, either undergraduate or graduate. Participation in the shared medical appointment may also offer trainees valuable opportunities for managing diabetic patients, feeling more comfortable caring for

patients with a chronic illness and help in developing skills needed to provide patient centered care.

Providing high quality clinical care where exemplary clinical training is occurring is a strong interest of mine. For a site to meet a high standard, it must achieve excellent outcomes for the patients, the learners, and the organization as a whole with tools to assess those sites along multiple dimensions. Our conceptual model is shown in Figure 2. The needs for such sites are particularly great in chronic care because most healthcare professionals have been trained and continue to be trained in an acute care-oriented health care system. Our efforts to develop sites of exemplary care and learning have focused on shared medical appointments/group visits which are based on the chronic care model.

Our preliminary data suggest the usefulness of this approach, raising questions about the mechanisms by which such an approach might work. Our conceptual model for the

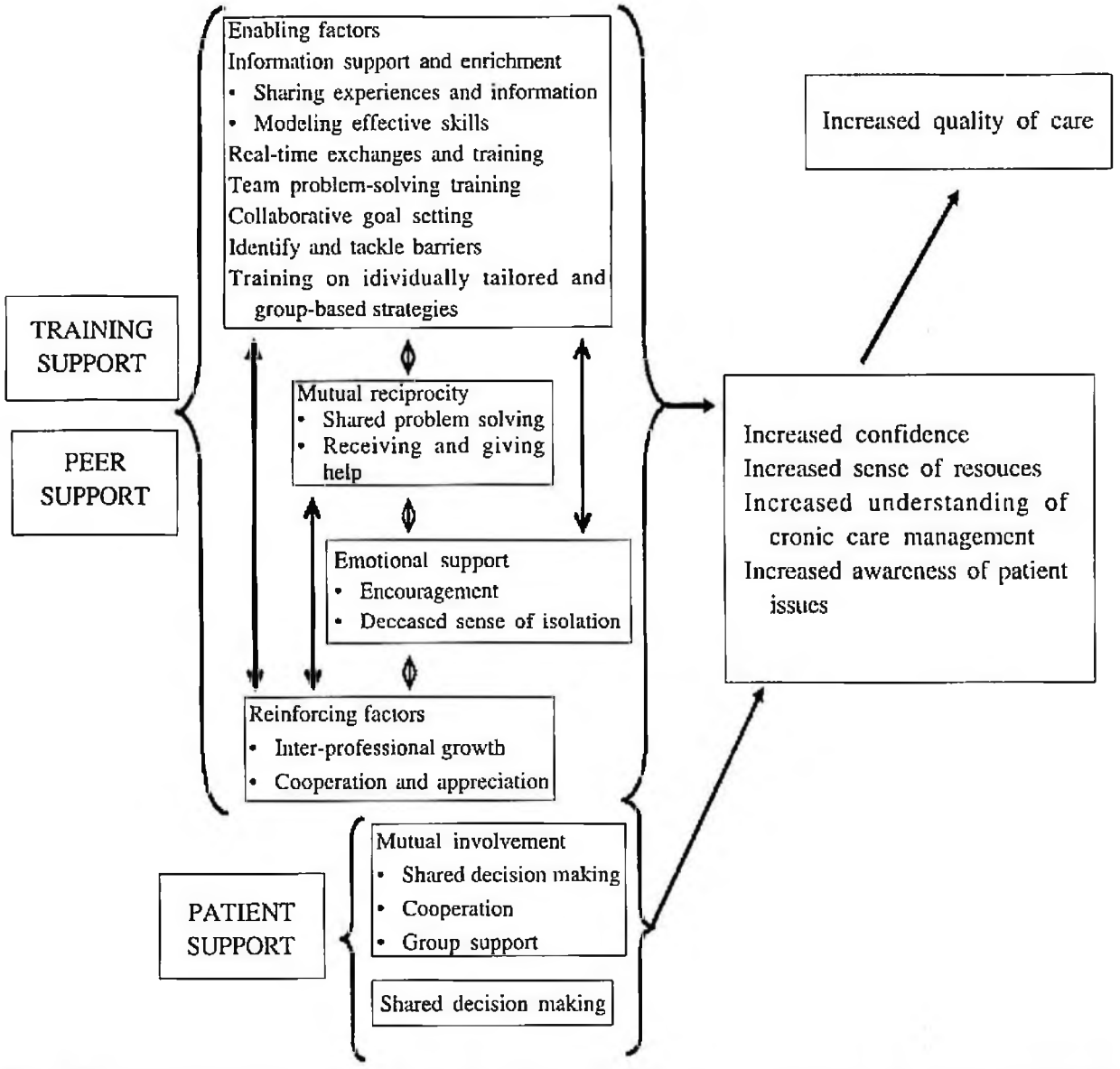
Figure 2
A conceptual model where quality care,
quality education and value to the institution exist simultaneously



Note: We suggest that outcomes need to be assessed at three different levels — patients, learners, and staff. Some suggested domains of measurement are shown for each level. This framework could apply to any setting where care and training are done together — inpatient, outpatient, or community.

"physiology" of training in shared medical appointments takes into account the principles of adult learning as well as issues related to peer support and the presence of multiple perspectives. Our model is shown in Figure 3. We hypothesized that Internal Medicine resident physician participation in a multidisciplinary diabetes SMA could improve attitudes and knowledge toward chronic illness care. We began testing this in a study and through feedback and interactions of these resident physicians with other health professionals, we conceptualized a model where trainees would not only learn about providing high quality chronic illness care, but the value added and perspective of other health professionals. This led us to the issues of interprofessional trainees learning together.

Figure 3
Conceptual model of factors associated with learning in an SMA

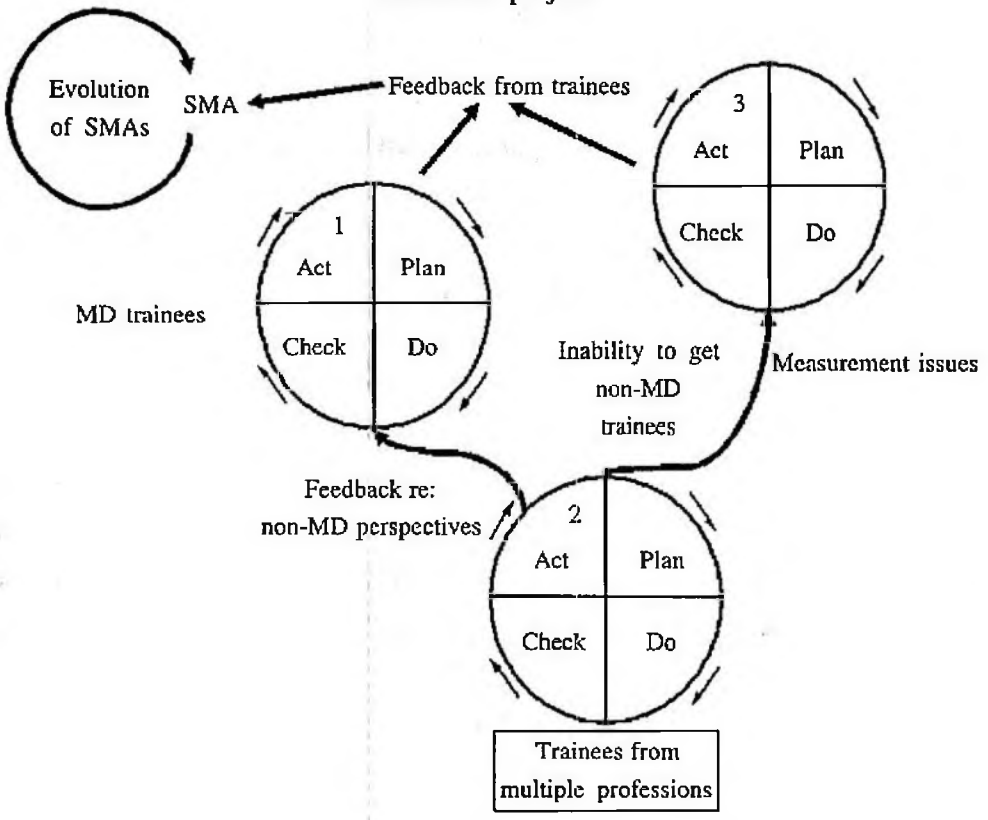


Note: We theorize that peer support is a critical factor for both trainees and patients and have designed the SMAs accordingly. Measures of these different aspects will be the subject of future studies.

The knowledge gap

Our long term goal is to address the gaps in knowledge about how to improve training in management of patients with chronic disease, recognizing that this must be done without compromising quality of care and ideally while improving the quality of that care. Quality care for chronic disease is best delivered in the context of a multidisciplinary team of providers. This has several implications. First, there are knowledge gaps related to training methods in general. Second, there are knowledge gaps about the effectiveness of interprofessional training. Third, given that effective chronic care management involves a team based approach, it was important to consider training outcomes in a more broad sense, for example, the importance of other team members and valuing what other team members bring to the table. SMAs as we have operationalized them are interprofessional and involve representatives from 5 health professions (medicine, nursing, psychology, pharmacy, nutrition). This project's path involved three major PDSA cycles, into which numerous small PDSA cycles were nested. This path, shown in Figure 4, first involved physician trainees

Figure 4
Evolution project



Note: Evolution of the project from one focusing on MD trainees (PDSA-1) then trainees from multiple professions (PDSA-2) and back to MD trainees (PDSA-3), all of which occurred in the context of the evolution of SMAs themselves.

(PDSA 1), then trainees from different disciplines (PDSA 2), and then back to physician trainees only (PDSA 3). We began at one place and after a number of efforts came back to where we started, albeit at a higher, more sophisticated level.

In order to pursue this long term goal, we had several PDSA cycles which were as follows: PDSA — was designed to assess shared medical appointment as a venue to improve knowledge, attitudes and self-efficacy in management of patients with chronic diseases like diabetes. Based on feedback from PDSA 1, PDSA-2 was designed to assess trainees in several disciplines. We wanted to identify attitudes toward chronic illnesses like diabetes as well as concepts such as teamwork and shared perspectives. We had some challenges which have led us to PDSA-3 which was designed to expand upon PDSA by assessing training and quality of care and determine if both could be improved simultaneously. In addition, we aimed to assess new measures related to interprofessional care.

PDSA 1

Research Design and Methods: The study design was a pre-test/post-test design to test the following hypotheses with questionnaires:

a) Trainees will demonstrate improved attitudes toward chronic care and improved self-efficacy in providing high quality diabetes management.
b) Trainees will demonstrate increased knowledge of the chronic illness of diabetes.
c) Resident physicians' participation in SMAs will be positively associated with the quality of care for patients with diabetes in the residents' primary care practice.

We settled upon the following measures for Internal Medicine residents participating in SMA's: Knowledge Assessment Test, Attitude Survey, P-ACIC/self efficacy survey, and a semi-structured interview. This study was embedded in a sequential partial cross-over randomized controlled trial of patients that also involves a pre and post-test. This study was designed to take place in the Cleveland VA Primary Care Clinic which consists of two academic group practices to which patients are randomly assigned. One group practice chosen at random will be assigned to the intervention group initially. During this time, the other group will serve as a control. When the eligible patients in the intervention group have all been approached to participate, then patients in the non-intervention Firm will be approached to participate. Thus all patients suitable for shared medical appointments will be

included. The only difference between the groups is the timing of the intervention; one group will precede the other.

Results: Because the 3rd PDSA cycle is a modification of the 1st PDSA cycle, we will discuss the results together below. However, this modification arose following another PDSA cycle.

PDSA 2

Research Design and Methods: Our initial plan was to study trainees of several types: physician trainees, nursing students, psychology interns, and pharmacy interns. However, we were unable to get enough participants from disciplines other than internal medicine. The reasons are discussed in the Methodological Challenges section. We have been forced to focus on physician trainees — medical students and residents in their 1st, 2nd, and 3rd postgraduate years of training. We used both quantitative and qualitative assessment methods.

Results: Consecutive internal medicine postgraduate trainees (PGY1 = 7, PGY2 = 4, PGY3 = 5) were included, representing 17% of internal medicine residents. These trainees participated in ≥ 1 SMAs. Attitudes towards interdisciplinary care were assessed quantitatively (Likert scale-based questions); this self-reported

Table 1
Baseline results of survey of attitudes towards interdisciplinary care

Attitudes toward interdisciplinary care: **1 = Strongly disagree to 4 = Strongly agree**	Mean ± 1SD
Interdisciplinary teamwork often fosters fragmentation of responsibility for patient-care decisions	2.19 ± 0.83
Receiving care from an interdisciplinary team creates confusion for patients and families	2.56 ± 0.73
Training with students from other disciplines dilutes the quality of training in one's own field.	1.81 ± 0.98
Comfort working in an interdisciplinary team: **1 = Very uncomfortable to 4 = Very comfortable**	
Negotiating patient care responsibilities with providers from other disciplines.	3.13 ± 0.96
Relying upon providers from other disciplines to perform tasks that you are capable of undertaking.	2.81 ± 0.91
Comfort seeking advice form various provider types: **1 = Very uncomfortable to 4 = Very comfortable**	
Nurse practitioner students	3.56 ± 1.36
Pharmacy residents	3.5 ± 0.89
Health psychology interns	3.8 ± 1.28

quantitative measure (11 items) had Chronbach's Alpha = .70. Baseline results are shown in Table 1.

Confidence when working with other provider types was assessed on a scale: 1 = Very confident to 4 = Very unconfident; lower numbers reflect greater confidence. Confidence in conveying the logic underlying your clinical recommendations to providers from other disciplines rose from 1.75 ± 0.77 prior to participation to 1.64 ± 0.74 Post and in understanding the distinctive perspectives of providers from other health care disciplines rose from 1.88 ± 0.62 Pre to 1.79 ± 0.70 Post, but the differences were not statistically significant.

Two focus groups were conducted by an individual not directly associated with the SMAs. Themes revealed in qualitative analyses of participants' perceptions: 1) Patient benefits ("shared expertise", "gain insight from one another", "discover universal issues", "well suited to poorly controlled patients"); 2) learning from team members/ patients ("ability to rely on shared experiences", "multidisciplinary approach", "responsibility is divided"); 3) SMA as a mode of healthcare delivery is well-suited to chronic disease ("tackles all aspects of patient care", "chronic care is a lot of work", "must address all issues surrounding disease"); and 4) patient-centered approach ("patients learn differently", "create solutions to achieve compliance", "allow patients to identify areas of change"). We concluded that trainees can be integrated into SMAs. Residents generally have positive attitudes toward interdisciplinary care and are comfortable working with individuals outside of their discipline. SMAs are well-suited to developing, in residents, an appreciation for the role that interdisciplinary care plays in managing the chronically ill and perhaps for reinforcing positive attitudes toward interdisciplinary care for chronic disease.

PDSA 3 — interprofessional training in a high quality patient care site

The specific aims of PDSA-3 include not only the aims related to trainees outlined in PDSA-1, but also aims related to improvement in patients' intermediate outcomes (A1c, LDL-cholesterol, and systolic blood pressure) and process measures (aspirin use, documentation of foot examination and consultation for eye examination).

Research Design and Methods: Ninety internal medicine residents are eligible to participate in the SMA and will do so 1-2 a week during their ambulatory block rotation which is 4 weeks long. The number of medical students is estimated to be approximately 75. Each student and resident will participate in between 2 and 8 SMAs. The questionnaires will address not only uniprofessional outcomes, but also include items related to attitudes towards interprofessional care (teamwork, shared

perspectives, and respect for other professionals). Questionnaires will be obtained at baseline and at the end of the block rotation (4 weeks.) Semi-structured interviews of IM residents will be conducted after participation in each SMA. These interviews will probe, in greater detail, issues raised in the surveys and identify other themes. In addition, the formative evaluation includes a weekly de-briefing involving SMA staff (as routinely conducted now, i.e., conducted independent of the issue of whether or not a research project is being performed). In order to assess resident utilization of skills obtained during SMA participation, patient outcome data (A1c, LDL-c, SBP, documentation of SMA goals, and use of a template diabetes clinic note) will be collected for residents who have continuity clinic patients and who have participated in ≥ 2 SMA's. Control groups will be drawn from patients receiving usual care and will include those who have had education classes but not SMAs and they, along with other non-participating veterans, will serve as a control group.

The following outcome measures will be obtained for patients participating in SMAs: A1c, systolic blood pressure, and LDL-c. In addition, the following process measures will be obtained: aspirin use, documentation of a foot and eye examination within the last year, a documented self-management goal as well as patient satisfaction. These measures will be collected at baseline, after third SMA, 6 months and 12 months after the first SMA. Patients will be referred back to their primary care provider after they reach targets of A1c < 9, SBP < 130 and LDL-c < 100 or they have participated in up to 8 SMAs. Information regarding the following will also be collected. Ethnicity/race will be self-identified by respondents. Respondents' age will be recorded in years and will be obtained from patient records along with gender. Marital status will be coded as married, separated, widowed, divorced, or never married. Living arrangements will also be coded. Educational attainment will be measured in number of years completed. For those respondents indicating 12 or more years of school completed, a follow-up question will clarify if they have a high school diploma or GED and if they have a college degree.

Data analysis

Quantitative: Statistical process control will be used for the patients in the groups. T tests will be utilized to compare control and intervention Firms. T tests will also be used for individual patients/trainees where the outcome measures allow the subject to be his/her own control (i.e., pre-test/post-test design). All questionnaires will be given to resident trainees by a member of the research team. In addition, repeated measures will be assessed with ANOVA models and statistical process control.

Qualitative analyses will be performed based primarily on key informant interviews all of which will be audiotaped and transcribed. Most of the initial coding and analysis of data will be conducted in collaboration with an expert qualitative analyst and by the project co-directors, who have significant expertise from previous research. Preliminary coding of these data will be performed using *Atlas.ti* software. A data code book will be constantly revised to reflect new ideas and evolving themes as they become apparent. Several forms of triangulation are built into the qualitative data analysis, and will be used to establish the trustworthiness of the data. A data display table will be used to ensure completeness of sampling and observational events. Additional corroboration, refutation, and interpretation of meaning of both the qualitative and the quantitative data will occur when different members of the analysis team perform individual coding and interpretation, and then share and confirm or refute their interpretation during team analysis meetings. A final form of triangulation will be between the quantitative questionnaire data and the qualitative observational data.

Methodological challenges in our current project as it has evolved into PDSA 3

After PDSA 1 and feedback from physician trainees about their interest in the roles of other healthcare professionals in the SMAs, we broadened our conceptual model to include trainees from multiple health professions. Our first challenge was to define what constituted interprofessional/interdisciplinary outcomes. A literature review indicated that multidisciplinary, interdisciplinary and intradisciplinary had different definitions. These differences while notable and explicit in a recent conference of the Institute of Medicine were difficult to differentiate in practice, especially in terms of measurement. The definitions tend to blur so we chose to adopt the term interprofessional with measures of domains that we felt were important. Another question that arose surrounding measurement of interprofessional learning was which instrument to choose. This instrument had to be applicable to all potential trainees. We found many instruments with varying degrees of validity and they address a variety of domains. We thought that trainee self-efficacy was a critical domain. Therefore, we chose a measure that included a measure of self-efficacy in interacting with other health professionals. In addition, we felt that "teamness" was a critical aspect of interprofessionalism and chronic illness care and selected a relevant measure of this. Cognizant of respondent burden, we limited our measurement to these two domains and developed our attitudes survey by using selected questions from different instruments.

When addressing study design, we identified a potential problem with dose of intervention and the subsequent ability to see an effect size. We could easily get

medical trainees to participate in 4 to 8 SMAs, but was this enough to change attitudes, knowledge and skills in a chronic disease like diabetes and have an appreciation for other health professionals with or without trainees that were present? Additionally, we wanted to see if participation in this training venue had any impact on resident physician's performance in his/her own continuity clinic performance however, there were so many confounders that it would be difficult to attribute changes to the SMA experience. Securing collaboration from other disciplines was also difficult in that it took some time to meet with leadership and get them to agree to trainee participation and to commit to doing so in a sustained fashion given all of the individual training program demands. Thus, we had difficulty recruiting an adequate number of trainees from disciplines other than medicine, i.e., medical students and Internal Medicine residents.

This research project has played out as the SMAs themselves evolved over the last two years within our local context. The team makeup was modified by adding a nutritionist after we recognized dietary education as a key patient need. The patient population also changed over time. We started with a large group of patients' not meeting performance measures in particular there were a high percentage of patients seen with an A1c > 9, although patients could also be referred by their primary care physicians regardless of their laboratory results. As we improved glycemic control and primary care providers recognized progress made in their high risk patients, they began to refer patients with A1c's between 7 and 8 such that these patients became a majority in the group. Additionally, many small improvements resulting from small tests of change resulted in increased efficiency. Among these were flow redesigns to perform glucometer downloading and blood pressure checks at the beginning of the SMA instead of at nurse intake at a remote location. The groups now accommodate up to 18 patients in a 2 hours session instead of 6 or 8. Each of these has the potential for affecting the learning experience for trainees as well as the experience for patients.

Resolving methodological challenges — the help of individuals and the literature

A number of modalities helped me resolve some of our methodological challenges. They include: people, collaborative experiences, my attendance in a research course, and specific references from the literature. One individual became an important project collaborator. She was a nurse researcher who had specifically focused on interprofessional training in an outpatient setting. She was instrumental in assisting us in measurement selection. She guided us to appropriate measures given our challenges in intervention dose and effect size. Although she attempted to involve nurse trainees, she was unsuccessful. One of our long- standing colleagues is an education researcher

with particular expertise in program evaluation. She served as a sounding board for our ideas and was able to point us in the direction of the evaluation tools from which we chose specific measurement instruments. Several meetings provided an environment of interaction with a number of colleagues whose interests overlapped mine. These meetings included the Health Professions Education Collaborative (HPEC) and the Academic Chronic Care Collaborative (ACCC) both IHI collaboratives. The HPEC consisted of representatives from multiple disciplines within multiple institutions. This collaborative both encouraged and assisted the development of interprofessional models of both inpatient and outpatient care that could serve as good training sites. The ACCC was a collaborative effort to improve chronic illness care in institutions that have a large role in residency training. Several scientific meetings provided a venue for similar kinds of discussions. These meetings included IHI scientific symposium and the annual meetings of VA HSR and D center of excellence. Established researchers willingly shared their expertise with me. For example, I spent two days at the Ann Arbor VA with members of their HSR and D center of excellence reviewing my research. I also participated in an intensive summer research course during which time I had to maintain focus on the project while experiencing enormous frustrations learning a new statistical program with which I was completely unfamiliar. Not withstanding, shared frustrations led to interactions with other interested colleagues. Prior to and throughout this project, I have continued to review the literature on interprofessional training and its measurement. This literature included both the peer reviewed literature as well as the gray literature. One particular reference from each of those literatures served as the underpinning of much of my approach. These included the paper by D'Amour et al. (D'Amour, Ferrada-Videla, San Martin Rodriguez & Beaulicu, 2005) with a very useful conceptual model for interprofessional training.

Practical challenges in implementation

As director of the primary care clinic, there were competing demands on my time. There was enormous pressure from top management to improve the clinic process and outcome measures; this was a much higher priority for management than was my desire to incorporate this as part of the residency curriculum. Our residency program centers more on training in the inpatient setting than the outpatient setting, therefore, additional efforts needed to be made to convince education directors of the value added with this new experience, i.e. getting their buy-in. Moreover, we had to engage a variety of colleagues and obtain buy-in from individuals in other disciplines/ professions. We needed to demonstrate that sending trainees (resident physician,

medical student, psychology graduate student, pharmacy residents and NP students) to participate and observe SMA for patients with diabetes was a valuable experience. Then we needed to try and coordinate a schedule with all the education leaders so that a few learners from each discipline were present at each diabetes SMA. Multiple trainees in a clinical venue with a multidisciplinary team approach to outpatient care was the initial research plan. Some of the staff participants could not break out of their comfort zone and work in a model that lacked the traditional hierarchy of authority. This resulted in the necessity to reassign one of the staff members. An additional challenge was the reporting silos (separate lines of authority) inherent in our medical center's organizational structure.

Perhaps our major challenge from a research standpoint was estimating the effect size of our intervention and identifying instruments with sufficient sensitivity for those effect sizes. We had very little information on which to base effect size estimates. For practical reasons, the one month length of participation in the SMAs was necessitated by the length of the resident rotation. We did a literature search to identify potential instruments and we will have to do some pilot testing. Sensitivities of these instruments are not well described in the literature. We will have to do some pilot testing and may have to develop new questions.

Resolving practical challenges — the help of individuals and the literature and taking advantage of the political context

This research project was constructed upon a clinical quality improvement project as clinic director. Improving performance measures for patients with diabetes met a specific need of top management. The chief of staff at my facility was very helpful in securing collaboration and resources of professionals from other disciplines as well as staff from support services. This was only a start. Shared medical appointments were featured at a breakthrough series type quality improvement collaborative. At this meeting, I met Ed Noffsinger who has played a large role in the development of shared medical appointments. I have obtained a great deal of practical information from him and other participants at this meeting. This meeting, which was part of an initiative to reduce waiting times in outpatient clinics involved several staff of the Cleveland VA representing multiple disciplines. This time away from the daily pressures of work, allowed us to devote energy to this project and facilitated the camaraderie that has helped to sustain this project. A researcher not involved in the operations helped me to identify methodologies to which I was unfamiliar. Although finding survey instruments was a challenge, I was familiar with methods of quantitative analysis. I have learned, with the help of this individual, about qualitative

methods and the richness of information that could be gleaned by these methods, information that would be critical in refining our design and analytic plan. Moreover, this would help us deal with the issue of the insensitivity of our quantitative surveys.

Progress to date (SMAs for interprofessional training and improved care; development of a research career)

Although we have successfully conducted SMAs continuously for over 2 years, the research aspect has progressed in fits and starts, periods of relative inactivity broken by periods of great progress. For example, certain aspects of the project required approval of the IRB, which has contributed to significant delays. However, during this time we were able to work on other issues such as literature review, identification of survey instruments and conceptualization. We have now received IRB approval. During periods of progress we have presented this material to national/international audiences. This has given me the opportunity to network with others health professionals, both in and out of my own discipline, on a national and international level. This has helped to broaden my perspective. Routine participation of residents and medical students in the SMAs demonstrates their feasibility in participating in this type of care. Medical students at CWRU have a chronic disease month where SMAs for diabetes, heart failure and hypertension are central to the rotation. At our institution we have conducted SMAs that have involved 6 disciplines and for several different chronic diseases. SMAs have even been used to teach first year medical students about this model of care. The students had a mock SMA demonstrated with several patients who have participated in the SMAs and felt that they significantly helped to improve their care. We have developed a team of interested individuals at the Cleveland VA as well as collaborators at a VA Health Services Research & Development Center of Excellence. This project continues to not only hold my interest, but has increased that interest over time. I have also been able to secure some research funding for this project from VA HSR&D. In short, this project has helped me to develop as a researcher and clarified the path which I have chosen for my academic career.

Lessons learned

Some of the lessons I have learned are more personal than others. This project is of great interest and passion for me and I have learned how important this fact is, for my interest and passions have sustained me when confronting challenges. Though it seems obvious to me that training should occur in settings where high quality is

delivered in outpatient settings by healthcare professional teams and that it makes sense for trainees from different disciplines to learn together, I have learned how difficult this is to achieve in practice. Disciplines/professions often work in silos; we are trained to think in a very uni-professional/uni-disciplinary fashion. I underestimated the difficulty in getting all disciplines to commit over time; there are so many competing demands on training programs. Perhaps the most important lesson for me, in developing as a researcher, is the importance of being open to new ideas and of challenging received wisdom. Addressing issues raised in the course of this research has forced me to devote much energy to developing my skills in systems thinking. I find myself continually developing and refining my conceptual models and linking different aspects of systems. In looking back over my activities, I have recognized that the time needed to develop conceptual models was lengthy, but worth all the time spent. I am not a full time graduate student and I have a number of time consuming clinical and administrative responsibilities. Even though this is a research project, it is closely linked to the clinical operations for which I am responsible. This contributes to making the most efficient use of my time. In addition, this project is closely linked to clinical operations that are highly valued by my employer. This has enabled me to obtain resources (e.g., time, staff assistance) that have made the project feasible within the constraints of my job. An example of this is the development of a model that is not only more integrative in terms of systems in an interesting way, but also reflects core values and can be shown as well to reflect institutional values (Figure 2). While the last is clearly important, it is the combination of the first two that I believe will sustain me in this long term process, of which this research project is only a small step.

Suggested references

1. Bodenheimer, T., Wagner, E. & Grumbach, K., 2002. Improving primary care for patients with chronic illness. *JAMA,* 288 (14), p. 1775-1779.
2. Bodenheimer, T., Wagner, E. & Grumbach, K., 2002. Improving primary care for patients with chronic illness: the chronic care model, Part 2. *JAMA,* 288 (15), p. 1909-1914.
3. Turner, J., et al., 2005. It's always continuing: first year medical students' perspective on chronic illness and the care of chronically ill patients. *Academic Medicine,* 80 (2), p. 183-188.
4. D'Amour, D., Ferrada-Videla, M., San Martin Rodriguez, L & Beaulieu, M., 2005. The conceptual basis for interprofessional collaboration: core concepts and theoretical frameworks. *Journal of Interprofessional Care,* (Suppl. 1), p. 116-131.

5. Oandasan, I., et al., 2004. Interdisciplinary education for collaborative, patient-centred practice: Research and Findings Report. Toronto: Health Canada.

Resources which would help other Q&SI researchers

Bammer, G., 2003. Integration and implementation sciences: will developing a new specialization improve our effectiveness in tackling complex issues? Draft of 19 Sept., http://www.anu.edu.au/iisn/. Posted to the Science, Environment and Development Group web site, 19 September.

Davis, K., 2001. Preparing for the future: a 2020 vision for American health care. *Academic Medicine,* 76 (4), p. 304-306.

Gharajedaghi, J., 2005. Systems thinking: managing chaos and complexity: a platform for designing business architecture. 2nd ed. Burlington, MA: Butterworth Heinemann.

Greenhalgh, T., et al., 2004. Diffusion of innovations in service organizations: systematic review and recommendations. *The Milbank Quarterly,* 82 (4), p. 581-629.

Grol, R. P., et al., 2007. Planning and studying improvement in patient care: the use of theoretical perspectives. *The Milbank Quarterly,* 85 (1), p. 93-138.

Kirsh, S., et al., 2007. Shared medical appointments based on the chronic care model: a quality improvement project to address the challenges of diabetic patients with high cardiovascular risk. *Quality and Safety in Health Care,* (in press).

Litaker D, et al., 2006. using complexity theory to build interventions that improve health care delivery in primary care. *Journal of General Internal Medicine,* in press.

Litaker, D., et al., 2006. Using complexity theory to build interventions that improve health care delivery in primary care. *Journal of General Internal Medicine,* 21 (Suppl. 2), p. S30-S34.

Nair, B. R. & Finucane, P. M., 2003. Reforming medical education to enhance the management of chronic disease. *Medical Journal of Australia,* 179 (5), p. 257-259.

Rubenstein, L. V. & Pugh, J., 2006. Strategies for promoting organizational and practice change by advancing implementation research. *Journal of General Internal Medicine,* 21 (Suppl. 2), p. S58-64.

Rycroft-Malone, J., et al., 2002. Ingredients for change: revisiting a conceptual framework. *Quality & Safety in Health Care,* 11 (2), p. 174-80.

Sales, A., Smith, J., Curran, G. & Kochevar, L., 2006. Models, strategies, and tools: theory in implementing evidence-based findings into health care practice. *Journal of General Internal Medicine,* 21 (Suppl. 2), p. S43-9.

Stetler, C. B. et al., 2006. The role of formative evaluation in implementation research and the QUERI experience. *Journal of General Internal Medicine,* 21 (Suppl. 2), p. S1-8.

Stetler, C. B., et al., 2006. Role of "external facilitation" in implementation of research findings: a qualitative evaluation of facilitation experiences in the Veterans Health Administration. *Implementation Science,* 18, p. 1: 23.

Stroebel, C. K., et al., 2005. How complexity science can inform a reflective process for improvement in primary care practices. *Joint Commission Journal on Quality and Patient Safety,* 31 (8), p. 438-446.

Thomas, P., 2006. Integrating primary health care: leading, managing, facilitating. Oxford, UK: Radcliffe Publishing.

PART 6
COMPLETED RESEARCH AND LESSONS FOR OTHERS

Successful Hospital and System Quality Programs: How Did They do it, and Is There Evidence of Improvement?

ANTHONY STAINES*

SUMMARY

Research questions: Can world class quality programs show evidence of improved clinical results? If yes, which methods and tools do these hospitals use to succeed and how are they implemented?
General design and sample: Case study design of three hospital's strategies chosen through a panel of 10 international experts.
Data collection and analysis methods: Semi-structured interviews, observation, gathering quantitative data from the improvement programs. Electronic database designed for the analysis.
Methodological challenges: Defining the right research question, locating data on processes and outcomes, data analysis, attribution of causality.
Practical challenges: Word processing for about 400 pages document; time management — sticking to a 24 months timeframe for the PhD program.

Introduction — from the initial idea to the research question

This chapter was written a few months after the defence of the PhD dissertation, the research having been completed. As a Chief Executive Officer of hospitals for ten years, before moving into research, and having been involved in leading quality improvement programs in these institutions, my initial idea was to investigate which improvement framework or methodology was best suited for what context.

I started with an exploratory phase, including interviews of researchers and consultants involved with quality improvement in hospitals, as well as of hospital Chief Executives. I asked the hospital leaders which process they had gone through

* Associate professor at IFROSS, University of Lyon 3, France, consultant, and Vice Chair of sanaCERT, accreditation body of the Swiss hospitals (anthony.staines@bluewin.ch).

when choosing a methodology for their institution's quality improvement initiatives and what had been their criteria to select their strategy.

The findings came as a surprise. None had actually followed a process to define the improvement program. In one case, someone in the organization had suggested an initiative, it had been tried and then had been carried on and had evolved into a program. In another case a framework hand been selected because CEO colleagues had mentioned it and no other framework was known at the time. In all instances, quality improvement had been an evolution, sometimes even with a fairly unclear origin, rather than a strategy building process. One Chief Executive, whose improvement program was amongst the most ambitious in Switzerland, even warned me that providing hospital leaders with a decision tool to select their methodology could be harmful: it could lead the decision-makers into relying on a simplistic decision tree, rather than carrying out their own feasibility study. The risk of a model, he said, would be to trivialize, to give an illusion of a safe selection process. The conclusion of the exploratory interviews was that my initial question was outside of what the professionals wanted to know.

As regards to what they wanted to know, there was consensus. They all wondered what the evidence was for improved results for patients through quality improvement programs. The exploratory literature review also showed the weak relevance of the initial question. According to Walshe, all methods work to a degree, but their effectiveness is highly variable, depending on the context in which they are used (Walshe & Freeman, 2002) *"(...) the approach to quality improvement used in an organisation probably matters less than how and by whom it is used"*.

The literature showed no evidence of improved clinical results from quality improvement programs, on an institutional level. There was evidence of some projects having led to some results in a given unit, but no evidence of sustained improvements on an institutional scale. The systematic review of quality improvement strategies, done by the Health Evidence Network for the World Health Organization, showed no scientific evidence of any strategy being superior to any other, in terms of efficiency or cost (Øvretveit, 2003). Shortell et al., in a research done in year 2000, including 3045 CABG patients from 16 different hospitals, evaluated the correlation between the deployment of Total Quality Management and clinical results (such as length of stay, adverse events, mortality). The results varied by a factor of 2 to 4, but with no correlation to TQM implementation (Shortell et al., 2000).

At that stage I found myself in a fairly difficult position. What I initially intended to research did not fit with what the people the investigation intended to serve identified as a priority for research. On the other hand, what professionals perceived to be a gap in knowledge (the impact of improvement programs on clinical results) was in fact answered already (no evidence of impact to date). However, having led quality management initiatives myself, I could feel how frustrating the findings from research were for

healthcare executives and how a large part of the gap in fact remained open. First, no evidence of impact did not mean evidence of no impact. Second, what should decision makers do with this lack of evidence? Should they stop all improvement initiatives? Could they still assert that they managed, if they stopped trying to improve their organization?

Drawing from the above, I considered taking the following stand:

- to consider that there was indeed no evidence of impact of quality improvement programs on clinical results, but also to recognize that previous research had not covered all types of programs.
- to hypothesize the existence of exceptions, of positive outliers, of world class improvement programs (at least a few on the planet) that could show evidence of improved clinical results.

At this point in the process, I had received encouraging comments from my research director, but still felt insecure about this strategy. New to quality improvement research, I felt I needed to discuss this perspective with another senior researcher. I decided to look for somebody with specific experience in quality improvement research in hospitals, with an international perspective, and with insights about the choice of research methodologies to use. The literature led me to John Øvretveit, at the Medical Management Center from the Karolinska Institute in Stockholm. With over 280 scientific papers and more than 20 books, a reviewer for and editorial board member of eight scientific health journals, including a book on evaluating health interventions (Øvretveit, 1998) and one on action evaluation of health programmes and change (Øvretveit, 2002a), this leading researcher appeared to be exactly who I was looking for. Furthermore, I detected from his comments in his papers a rigor in the approach that I really felt I could benefit from. A bit intimidated by this author's impressive biography, I nevertheless e-mailed a request for an interview. I have kept all the e-mails. His first reply said he would be happy to meet sometime but at present was rather tied up with travel. I felt that if this leader in improvement research could realise how much I could benefit from a short interview, he would certainly accept to meet. So I decided to keep trying and my next message to him was "I fully understand the time constraints linked with travel. I am ready (...) to meet you in an airport transit area (...) any arrangement that is not time consuming for you (let's call it recycling waiting time) (...) please feel free to suggest meeting you anywhere at any time, even if it means unusual places or schedules. (Yes I really feel an hour of your time could help me a lot.)"

After this message, I received an invitation to meet in London. The hour and a half I spent with John Øvretveit turned out to be the most worthwhile and the most inspiring in my whole research process. He encouraged the study of an elite through

a case study and gave me some specific ideas on how to design the case study. Only then was I able to clearly define what would become the research question:

Can world class quality programs show evidence of improved clinical results? If yes, what methods and tools do these hospitals use to succeed and how are they implemented? What is the perception of these programs within the institution and what are their evidence-based results? What is the reason for these results? And if world class programs do not show improved results, what is the phenomenon that prevents the goal from being reached or from being assessed?

LESSONS AT THIS STAGE

- Healthcare executives call for evidence about the impact of quality and safety improvement programs in their field, but literature reviews show no evidence of improved clinical results, on an institutional basis, from improvement programs.
- The exploratory phase of the research is fundamental. It allows to identify irrelevant research questions or research methods and to check on the existing body of knowledge on the subject.
- Researchers are part of a community. When stuck, it is worthwhile to identify through the literature, researchers that have experience on the specific difficulty that is at stake, and to search for insights.
- People can only help if they are advised on why specifically their help is sought for, what they can contribute, and within what timeframe.

A case study design focussing on world-class improvement programs

A case study design was chosen for this research. The literature review had revealed enough quantitative data showing that no global impact of quality improvement programs on clinical results could be proven to date. One way to take this further was to look for exceptions, in a very specific niche: the elite. This, by definition, implied a small number of programs to study. But it implied a very deep investigation into these programs, their content, their implementation, and their context.

According to Yin, "A case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident". This was precisely the situation I was in. Quality improvement programs had existed for a number of years and were still running. They were and are a contemporary phenomenon. They could not be isolated from their context and had therefore to be investigated in-vivo. Yin adds that

(Yin, 2003: p. 13) "The case study inquiry 1) copes with the technically distinctive situation in which there will be many more variables of interest than data points, as one result 2) relies on multiple sources of evidence, with data needing to converge in a triangulating fashion, as another result 3) benefits from prior development of theoretical propositions to guide data collection and analysis". This again completely described the context of our research. Improvement programs overflow with quantitative and qualitative variables, their volume being a challenge for the analysis. Their investigation within the triangulation process is the only way to obtain insights about the contents of the programs, their implementation and the reason for their results.

A case study is justified when research questions focus on "how" and "why" (Wacheux, 1996: p. 89; Yin, 2003: p. 1). It allows to describe a management situation across time (Hlady-Rispal, 2002: p. 65) reconstructing sequences (Wacheux, 1996: p. 89) assessing local causalities, exploring the emergence of a strategy and of its recursive causalities as well as of coming up with hypotheses for an explanation. It is particularly relevant for complex and dynamic contexts (Fitzgerald, 1999) or when there is little prior research allowing to draw hypotheses. In such a case, no single method can on its own capture all aspects of an intervention (Keen & Packwood, 1995). Case studies, however, draw on a range of methods.

Keen and Packwood emphasize the high potential of case studies for evaluation purposes, in situations involving judgements about the relevance of an intervention by examining its outputs in relation to its inputs and processes (Keen & Packwood, 2000: p. 51). «Asking participants about their experiences, or observing them in meetings and other work settings, can provide rich data for descriptive and explanatory accounts of organisational processes, work practices, and the impact of change". Keen & Packwood consider case studies to be particularly relevant when the change that is introduced occurs in the chaos that characterizes the real world and when it is important to understand why an intervention succeeds or fails. In such conditions, experimental methods cannot be considered. Another reason for choosing case studies is when the success of the intervention is highly dependant on the commitment of all stakeholders (Keen & Packwood, 2000: p. 52). Each stakeholder will have its own interpretation of events. Case studies can capture these interpretations through interviews or other qualitative methods. Case studies incorporate the views of the stakeholders (Tellis, 1997).

Having chosen a case study design, the next step was to select the cases. Hlady Rispal considers three criteria as crucial for sample selection (Hlady-Rispal, 2002: p. 82):

1. Theoretical representativeness, meaning alignment to the research question.
2. Variety, meaning a selection of cases very different from one another in different stages of development, different relational contexts, in order to develop an understanding of all dimensions of the problem and of its complexity.

3. Potential for discovery, meaning a selection of cases with a high potential for data collection on the problem that is studied.

Yin, on the contrary, rejects the concept of sampling. For this author (Yin, 2003: p. 45-48), sampling implies listing the population and going through a statistical process of selecting a representative sub-population to investigate. For Yin, case selection follows a logic of replication rather than sampling.

My research question being focused on world-class quality improvement programs, it was natural to search for this population. Theoretical representativeness would come from being able to locate programs qualifying for that status. This therefore became our first selection criteria.

To select hospitals with such improvement programs, I contacted a panel of 10 international experts by e-mail. I chose experts having published about quality improvement programs for hospitals on an international basis within the industrialized world, with a majority being non native English speaking, from four different continents.

MESSAGE TO THE EXPERTS

Could you help me by naming 3 to 5 hospitals that, to your knowledge, are among those that have made the most out of Quality Management? (I expect these to be hospitals that have probably pursued Quality Management for at least 5-8 years, whatever the label used for it (CQI, TQM, Six Sigma, MBNQA or EFQM, ISO, Six Sigma, accreditation (...) or a mix of many of these). I am also thinking of institutional programs involving most departments and most types of professionals.

Your list of 3-5 hospitals will be aggregated with the lists of the 9 other experts and contacts will be taken on the basis of an anonymous aggregated ranking. This document will be sent to all the experts that have provided a list.

Thank you very much for the 2 minutes you took to read this message and even more for the next 3 minutes you might want to contribute with, answering this mail by October 31st, with 3-5 hospital names and locations.

Nine out of the ten experts answered and provided a list of hospitals. The aggregated results lead to a list of 22 hospitals, one being on 4 different lists, one being on 3 and 5 hospitals being on 2 lists.

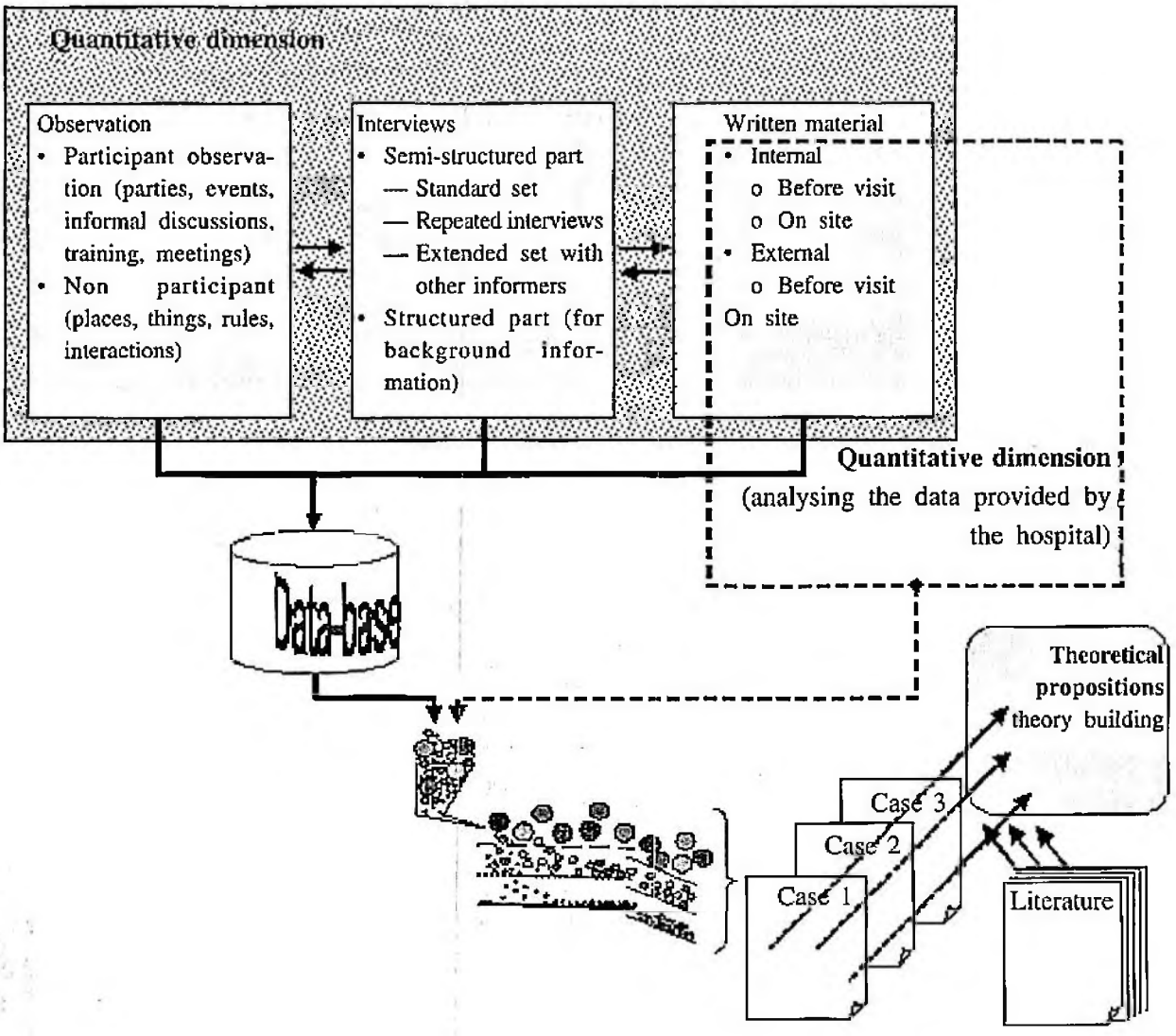
I pre-selected 3 hospitals. This was done starting with the one recommended by 4 experts and the one recommended by 3 experts. The first being in the US, I used the second criteria (diversity) to decide that the third hospital that I would select

should be from outside the US. This decision brought the list of hospitals with 2 recommendations down to a single hospital. I then checked the pre-selection against other criteria of diversity and noticed substantial differences in size, integration, political culture, language spoken, public versus private status. This allowed settling for three hospitals that met Rispal's requirement and I contacted their CEOs, who fortunately all agreed to be part of the case study.

Methods for data collection and analysis

Data collection was carried out through semi-structured interviews, observation and collection of written material (see Figure 1).

Figure 1
Strategy for data collection and analysis



To select interviewees, I constructed a standard set of interviewee profiles in order to cover a variety of departments and hierarchical levels, including observers internal and a few external to the organization, leaving however room for an extended set of interviews, as a development from the standard set. If I felt I was getting too many people with given characteristics (people whose project succeed, for example) I would ask for disappointing projects and contact these people. The key to the interviewing strategy was based on Øvretveit's recommendation (Øvretveit, 2002b). *"The researcher gathers data about the effects of the program by interviewing health personnel to find out what difference the program made for them and their views about the effects for patients (what would have happened without the program?)"* To be able to triangulate data Øvretveit suggests, *"One technique is to ask informants if they know of any evidence which would prove or disprove their perceptions. In addition, the researcher asks informants for their ideas about which factors helped and hindered the intervention — their theories of causality — and for evidence which might support or disprove their ideas".*

For observation purposes and for the collection of written material, I had prepared check-lists with a minimum data set that I needed to collect, but also had left open for additional data that the concurrent analysis would lead us to collect.

For data analysis, I constructed a data-base. Yin urges researchers to do so (Yin, 2003: p. 95) "every case study project should strive to develop a formal, presentable data-base, so that, in principle, other investigators can review the evidence directly and not be limited to the written reports. In this manner, a case study database markedly increases the reliability of the entire case study".

In this case, the data-base would not only insure reliability, it would become the instrument for cross-comparisons of a high volume of qualitative data. All interviews were transcribed and the same was done with observations. Meaningful extracts of qualitative written material were also entered in the data-base. This data was then encoded after creating a dictionary of themes (keywords). Five sets of keywords were used for the encoding:

- Methods and tools, including all methods, frameworks, systems and tools used by the organizations that were studied.
- Management, including management concepts and practice.
- Actors, listing key actors for which data could be gathered.
- Fields of application, including subjects chosen for improvement projects.
- Attitudes, including attitudes of the people we interviewed or observed.

Entering all the qualitative data in the data-base meant a lot of work. It allowed, however, for a very diverse set of queries at a later stage, crossing over all the dimensions. For example, I could extract all the summaries and the quotes collected

Figure 2
Screen — architecture of the data-base

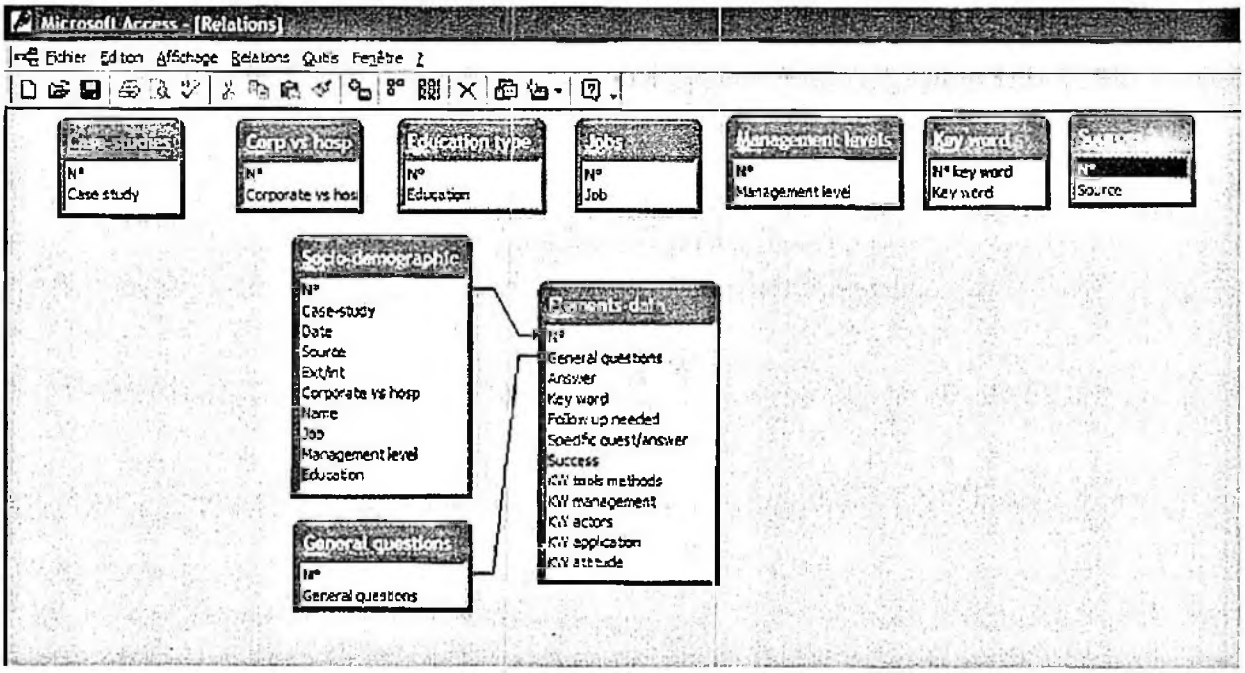


Figure 3
Socio-demographic data

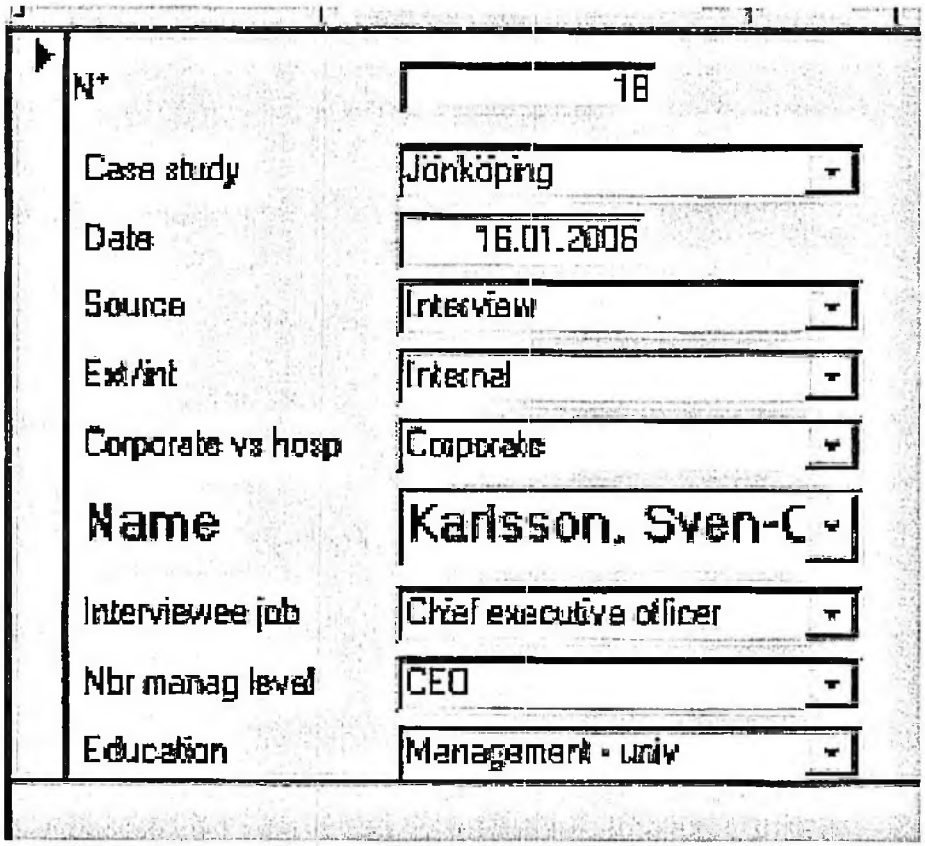
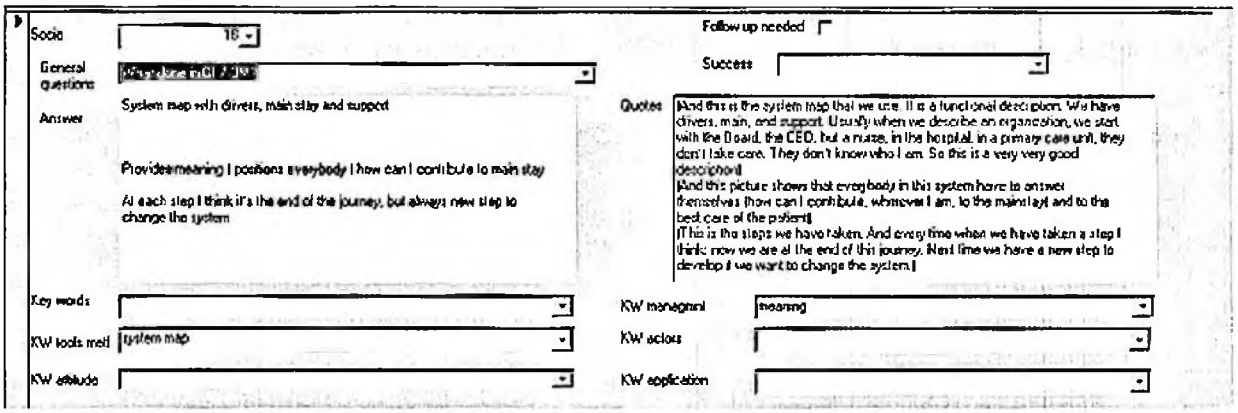


Figure 4
Data entering and encoding



on the strategy and sort them by hierarchical level to see how the content and the meaning had spread in the organization.

Methodological challenges

The first methodological challenge has already been described. It was to come up with the right research question. The exploratory phased proved to be essential in this respect.

On site, a dilemma has been how to deal with the hospital's public relation effort. In two of the three healthcare systems, public relations were well developed. In one, there was even a sort of unconscious way of all staff to unite in praising the quality improvement program. I was lead to meet extremely interesting people and to visit very successful initiatives. What proved essential to be in a position to get around public relations was to stay long enough on the premises. After two weeks, I would know enough people to find out by myself which departments to visit and which people to meet. I would know at what time to arrive in the staff cafeteria to informally join a table and start a conversation, introduce myself as a researcher on the QI program and ask people if they had been involved in QI (and run to a quite place as soon as possible afterwards to record as many details as possible on my tape recorder).

Another challenge has been locating data on processes and outcomes. In two of the three organizations, people sometimes referred to figures, to data, but then often said they didn't have the data themselves, mentioning about somebody else who would have it. To solve this, as a researcher, I had to build my own network within the organization. This also emphasizes the importance of staying long enough on site.

Data analysis came out as a major challenge. I started transcribing the interviews on the premises, with keywords linked to all items for analysis and with follow up

questions for further interviews. However, I ended up with about 100 interviews (on average about 75 minutes each), observation notes, printed material and quantitative data on processes and outcomes. I considered manual processing with paper boards and key ideas on notes. The work seemed overwhelming. I then decided to test building an electronic data-base. It took a good week and expert help to build the data-base and then a good month to enter all the information, but this then allowed to sort key words, abstracts of ideas and quotes by many different criteria and compare views by organizations, hierarchical level, profession, internal vs. external interviewees, education level of interviewees, improvement methods used, etc.

Being in a theory-building process, one of the challenges was also the attribution of causality. In a live organizational setting, everything changes all the time. Can an improved outcome be linked to a improved process measure? Can the improved process measure be attributed to the improvement intervention?

I had to remember that causality can never be proven with certainty. In many instances, many other factors than the intervention could be possible explanations for the improvement. it was not even necessary to consider the issue. However, in some instances, I felt it would not be fair to hide behind the fact that causality could not be demonstrated. To be fair to the program, I had to consider the probability of causality. In some instances, for example, although no causality could be proven, I saw that there was nevertheless a higher probability of causality than a probability of no causality. Despite the potential controversy I decided, in the research report, to discuss probabilities of causality.

LESSONS LEARNED

- Staying long enough on the premises helps to get around the public relations effort and to build one's network to locate data.
- Designing or using an electronic data-base creates a great potential for data processing and analysis.
- Causality can never be proven with absolute certainty. Probability of causality is worth discussing.

Practical challenges

One important practical challenge was word processing. The more the document grew (ending up with 468 pages) with charts, pictures, tables, the more unstable it

became. Layouts were lost a number of times. Fortunately, a colleague from the IFROSS institute was extremely generous with his time, almost setting up a private 24/24 hotline for me in the last ten days.

Another challenge was time management. As I had started my PhD at the age of 44, without any grant, leaving a position of CEO of a group of hospitals, I had decided that I wanted to carry out the full process full time in 24 months. I found myself in an environment that was built around completely different constraints and concerns. Most of my fellow doctoral students were involved in consulting projects, in lecturing and considered it extremely fast to complete the PhD requirements in 4 years.

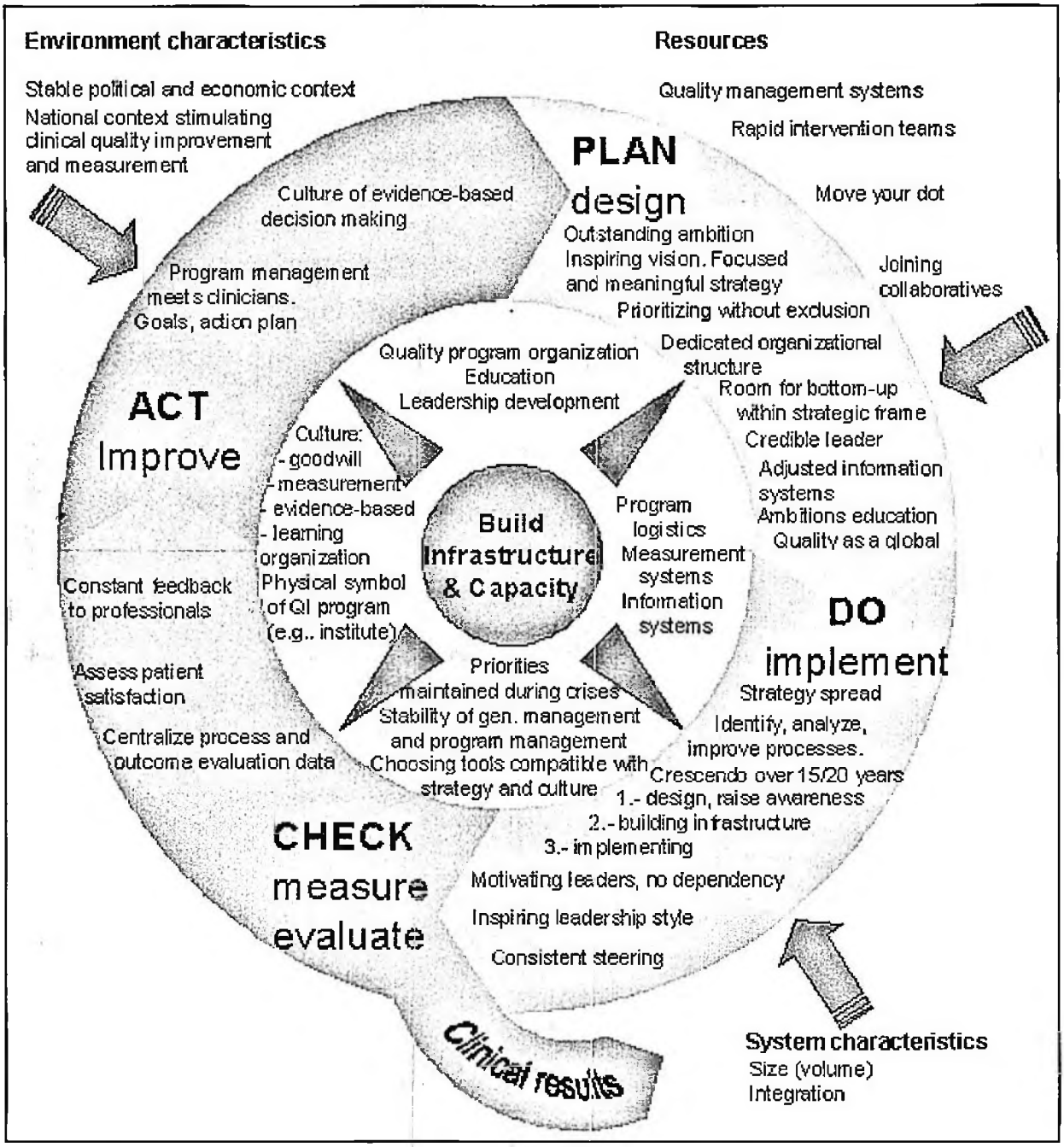
Findings from the research

Figure 5 shows the factors that have immerged from the theory building process as influencing the impact of quality improvement programs on clinical results for patients. Some factors have been grouped around the four steps of the Deming wheel. They belong to the running of the improvement program. Outside the wheel, there are system characteristics that favour or hinder the improvement program, resources and environment characteristics. The central part plays a crucial role, as revealed by the case study. I have called it "building capacity and infrastructure" for the program. This part takes a lot of time to complete. It is however crucial to invest in these aspects, because it is only once they are in place that the program will start being in a position to lead to evidence of improved results for patients.

One of the programs that were studied showed improved outcomes in the majority of its departments. One showed improved outcomes in one department and improved process measures in a number of departments. The third revealed improved process indicators in only a few departments. For all of the programs, it had taken more than ten years to come to a stage where evidence of sustained clinical improvements would start to show.

The study of the quality improvement interventions and of the program results lead to the concept of a threshold in the quality improvement process. I have called it the "investment threshold". Below that investment threshold, the hospital is not in a position to show evidence of sustained clinical results improvement on a global scale. It is as if the organization was in a zone of noise, where the signal of improved results is covered by noise. The investment we refer to is precisely what we have identified above as "building capacity and infrastructure". This phase implies raising awareness, developing leadership will and commitment, thriving through the political process of freeing up resources for QI, training staff, building culture, setting up indicators and

Figure 5
Factors influencing the achievement of improved clinical results



data collection systems, adjusting computer systems to capture and process the data and testing QI tools.

The zone of noise can be due for example to too short time series to show evidence of improvement, to too few cases for correlations to be significant or to measurement systems that have not yet reached maturity (not yet dealt with disagreement). It can also be due to inverse correlations showing up, for example because the system for the reporting of adverse events has improved, people start to report more adverse

events, whereas the evidence that was hoped for was a decrease in adverse events.

Understanding the concept of the investment threshold has important implications. Without this understanding, healthcare leaders will too soon turn to new strategies, new methods or new tools. And the new framework will just imply a new threshold, never allowing the necessary time for capacity and infrastructure to be in place and for results to start appearing. Being in the noise level prevents hospitals from realizing the benefit of their improvement program. Recognizing the threshold phenomenon allows them the required determination and encourages them to focus their effort in emerging out of the noise level.

Publications that could inspire quality & safety improvement researchers

West, E., 2001. Management matters: the link between hospital organisation and quality of patient care. *Quality in Health Care,* 10, p. 40-48.

Walshe, K., Rundall, T. G., 2001. Evidence-based management: from theory to practice in health care. *The Milbank Quarterly,* 79 (3), p. 429-457.

Mant, J., 2001. Process versus outcome indicators in the assessment of quality of health care. *International Journal for Quality Health Care,* 13 (6), p. 475-480.

James, B. C., 2002. Physicians and quality: improvement in hospitals: how do you involve physicians in TQM. *The Journal for Quality and Participation,* 25 (2), p. 56-63.

Grol, R., Baker, R. & Moss, F., 2004. Quality improvement research: the science of change in health care. In Grol, R., Baker, R., Moss, F., ed. *Quality improvement research: understanding the science of change in health care.* London: BMJ Books. 1-5.

References

Fitzgerald, L., 1999. Case studies as a research tool. *Quality and Safety in Health Care,* 8 (2), p. 75.

Hlady-Rispal, Martine, 2002, *La méthode des cas: application à la recherche en gestion.* Bruxelles: De Boeck.

Keen, Justin & Packwood, Tim, 1995. Qualitative research: case study evaluation. *BMJ,* 311 (7002), p. 444-446.

Keen, Justin & Packwood, Tim, 2000. Using case studies in health services and policy research. In Pope, Catherine & Mays, Nicholas, ed. *Qualitative research in health care.* 2nd ed. London: BMJ Books. 50-58.

Øvretveit, John, 1998. *Evaluating health interventions*. Buckingham: Open University Press.

Øvretveit, John, 2002a. *Action evaluation of health programmes and change: a handbook for a user focused approach.* Oxford: Radcliffe Medical Press.

Øvretveit, John, 2002b. Producing useful research about quality improvement. *International Journal of Health Care Quality Assurance,* 15 (7), p. 294-302.

Øvretveit, John, 2003. What are the best strategies for ensuring quality in hospitals. Geneva: WHO Regional Office for Europe's Heath Evidence Network (HEN).

Shortell, S. M., et al., 2000. Assessing the impact of total quality management and organizational culture on multiple outcomes of care for coronary artery bypass graft surgery patients. *Medical Care,* 38 (2), p. 207-217.

Tellis, Winston, 1997. Introduction to case study. *The Qualitative Report,* 3 (2).

Wacheux, Frédéric, 1996. *Méthodes qualitatives et recherche en gestion.* Paris: Economica (Collection Gestion; 290).

Walshe, K. & Freeman, T., 2002. Effectiveness of quality improvement: learning from evaluations. *Quality and Safety in Health Care,* 11, p. 85-87.

Yin, R. K., 2003, *Case study research: design and methods.* Beverly Hills, CA: SAGE Publications.

Being a Quality Improvement Practitioner-Researcher: Advantages and Challenges

JOHAN THOR*

One of the key challenges in conducting organizational and management research is to access relevant field data (Gummesson, 1991). Most management research can not feasibly be conducted in laboratory settings. Being a highly applied field, it instead needs to take place where management efforts happen in real life. Since I wanted to pursue my doctoral research on the issue of quality improvement (QI) in healthcare, I therefore seized an opportunity I saw to combine my practical work supporting QI at a large university hospital with research on healthcare QI. With a view to helping colleagues with similar interests, this chapter conveys some of the lessons learned — some benefits and challenges of being a practitioner-researcher — over the ten-year period that it took to complete my doctoral research and ties them to the literature on this topic. The chapter is an adaptation of parts of my doctoral thesis (Thor, 2007).[1]

Why pursue this research?

The research reported in my thesis was inspired by a deceptively simple question: Does quality improvement (QI) work in healthcare? In other words, does application of QI principles and methods enable clinicians and managers to manage and improve healthcare? The answer, as it turns out, was not quite that simple, as I elaborate on in the thesis.

I became interested in QI — or, really, I was driven into it — for two reasons: My desire to provide good care to my patients, and my experience as a young physician that the healthcare system I found myself working in was not well designed

[1] The thesis is available electronically at: http://diss.kib.ki.se/2007/978-91-7357-274-3/thesis.pdf.

* The Medical Management Center, Karolinska Institutet and The Hospital Director's Office, Karolinska University Hospital, Sweden (johan.thor@ki.se; johan.thor@karolinska.se).

to enable me, or others, to provide such good care. Frustrated, I saw a need to improve the state of affairs. Then, I ran in to another source of frustration: It turned out to be rather difficult to improve healthcare as a junior physician (and for many others too, I later found out). My suggestions to colleagues and superiors for ways to improve typically met with anything from mild disinterest to outright rejection. The feeling that I needed to strengthen my ability to improve healthcare was brewing inside me.

When I had the opportunity to learn about QI at Harvard School of Public Health, and later at the Institute for Healthcare Improvement, in Boston, I began to be hopeful again. The QI principles and methods I was introduced to there resonated with my experience and my values. I felt that this was by far the best approach to addressing the problems I had witnessed in healthcare. Eager to apply QI at home, I encountered the next challenge: Many colleagues were skeptical of QI, and asked for evidence that it really works. That brings us back to the question stated above: does QI work in healthcare? The thesis reports my research and my understanding of the literature on this topic thus far.

Framing the research problem

Although I was interested in understanding whether, and how, QI "works" in healthcare, I realized, for reasons I will discuss below, that I was in a better position to address a different, but related research question: How does an organization begin to apply QI in practice? There are a myriad of decisions and actions that go into establishing QI, and thus many opportunities to "go wrong". The sum of all these choices and actions over time constitute the degree to which an organization applies QI in practice. Understanding the degree of application is key for assessing the impact of QI (Øvretveit & Gustafson, 2002). Therefore, my doctoral research aimed not primarily at demonstrating whether QI causes better performance, but rather at illuminating how a healthcare organization begins to apply QI, thus laying a foundation for QI effectiveness, in line with suggestions in the literature (Greenhalgh et al., 2004; Øvretveit & Gustafson, 2002; Walshe & Freeman, 2002).

The research opportunity

The research reported in my thesis drew on data collected during the introduction of process management — a method for QI — at the Huddinge University Hospital (HUH) from 1997 through 2001 (reported in three separate articles (Thor et al.,

2004a; Thor et al., 2007a; Thor et al., 2004b). The thesis also included a systematic literature review of how Statistical Process Control has been applied to healthcare QI (Thor et al., 2007b).

The introduction of process management — defined at the hospital as "a systematic way to organize, lead and continuously improve the processes of an organization" — provided what is called a "natural experiment". In 1997, the hospital management initiated improvement efforts in 12 clinical processes at the hospital. By the end of 2001, over 25 processes were involved, 93 improvement projects had been initiated and over 65 of them had been completed or otherwise ended. The HUH (which merged to form the Karolinska University Hospital in 2004, after the study period), was a tertiary care academic medical center, part of the Stockholm County Council publicly funded healthcare system. During the study period the HUH had over 6000 employees, including some 2500 nurses and 1000 salaried physicians. Caring for approximately 45,000 inpatients annually, it had 800 beds, and its clinicians provided some 500,000 outpatient visits, including in a busy Emergency Department. Care was provided in 50 departments, which were organized into six clinical divisions. The hospital had close ties with Karolinska Institutet, the medical university in Stockholm: many students in the health professions received their clinical training at the hospital and much research was carried out there. In April of 2000, the hospital was incorporated but it remained wholly owned by the County Council.

I thus had the opportunity, together with colleagues, to both participate in the process management efforts and collect data documenting those efforts with the explicit intent to subsequently use the data for research purposes. This was an observational study of a natural experiment: the actions taken and the decisions made by various stakeholders in the process management initiative were guided by the needs and circumstances of the hospital and its employees and managers; not by the researchers. This is in contrast to an experimental study where researchers design and govern the application of the intervention of interest, or to action research, where the researchers and their research findings guide the intervention as it unfolds.

The study of process management at HUH was, in this sense, opportunistic. The hospital leadership had decided to make an effort to introduce process management; I was recruited to help with that effort. My interest in adding a research dimension to that effort was received favorably. Given the uncertainty about the effectiveness of QI in healthcare, and perhaps facilitated by the fact that this was a university hospital where research — even if not typically management research — was seen as part of the identity, the idea to evaluate the process management efforts through research was accepted. Consequently, the group of facilitators that I joined arranged to collect data documenting the efforts to introduce process management from the outset of the initiative. A guiding design principle was to make data collection part of the regular

work, and to make it useful both to the practical efforts at the hospital and to subsequent research efforts. I also discussed research design issues broadly with one of my advisors, although the specific studies did not take shape until later. Our data collection was rather broad, partly because it fit with our practical needs, partly perhaps because we wanted to cast our net wide to enable the pursuit of different future research questions.

Study data and design

As part of the efforts to introduce process management, the outcomes of all team efforts were documented electronically. Frequently, for instance, improvement teams used brainstorming to generate ideas — on problems, their causes, or potential solutions — at different stages in their improvement efforts. They wrote down these ideas individually on sticky notes (Post-it®) which they subsequently organized thematically on a white board or a wide paper taped onto a wall. After each meeting, a facilitator (or occasionally a team member) dictated all the text from the sticky notes, and their categories. A medical secretary transcribed the dictations. The facilitators and team representatives checked these transcriptions for accuracy and made corrections before they were circulated to all stakeholders. The facilitators stored all these electronic documents on a secure hospital server.

We facilitators documented our observations during team meetings, and other significant events or communications, such as key e-mail messages, in electronic progress notes. These were kept in a database (File-Maker®) designed by the lead facilitator. We usually dictated, or sometimes typed, these progress notes immediately after the conclusion of a team session, although sometimes time did not permit this and notes were written later or even, occasionally, not at all. A medical secretary transcribed the dictations, and the facilitator concerned checked these transcriptions for accuracy and made any changes necessary. We facilitators relied on these progress notes for our ongoing efforts — mirroring how clinicians use progress notes documenting patients' care — to keep track of what had been done at different meetings, commitments made etc. We also used the notes in periodic evaluations of our own work, to identify areas in need of improvement, and sometimes also specific ideas for how to improve.

In addition to these data, which were collected as part of the process management efforts at the hospital, two elements of data were collected more specifically for this research: First, in the course of the first year and a half of the process improvement initiative, I interviewed all the heads of the six clinical divisions at the hospital and the two hospital CEOs who were in charge during this period. The interviews were

semi-structured with questions regarding the respondents' views on the rationale for the process management initiative at the hospital, and the recent history leading up to it, their views on the potential of this initiative, as well as the main challenges and barriers to success. Second, members of the facilitator group at the hospital gathered over a weekend, after the end of the study period, to review and reflect on their efforts and capture lessons learned. These lessons were documented on sticky notes, transcribed into electronic format, shared with all participants, and stored electronically.

Study designs

The three studies based on data from the HUH were all observational studies, undertaken within an overall framework of a case study. We chose a case study design because it is "the preferred strategy when 'how' or 'why' questions are being posed, when the investigator has little control over events, and when the focus is on a contemporary phenomenon within some real-life context" (Yin, 1994: p. 1).

Ethical considerations

Plans for the thesis research were submitted to the regional board in Stockholm for vetting the ethics of research involving humans in accordance with Swedish law and regulations for research at Karolinska Institutet. The board concluded that this research did not meet the criteria under the law for research that it needs to vet (because it did not involve data deemed to be of a sensitive nature, about individuals), and added that, from its perspective, it had no objections to the research.

The main object under study in this project was the organization — particularly how process management was introduced there — rather than the particular individuals who worked there. While it is the actions and decisions of many individuals that constitute an organization, it was not their individual behavior that was of prime interest here, but instead the organizational phenomena that those actions and decisions represent.

The vast majority of research data from the HUH was initially collected as part of the day-to-day operation of the hospital. The process management efforts were undertaken irrespective of this research initiative. The potential ethical problems in this research, therefore, do not have to do with the process management efforts that employees of the hospital were "subjected" to, but rather with the act of using the data documenting those efforts for research purposes.

One important way to prevent ethical problems is to request informed consent from research participants and keep participation voluntary (Eriksson & CODEX, 2000-2006). This was practiced in the case of the interviews with senior managers at the hospital and for the follow-up session with the group of facilitators. Information was given verbally, and participation in the subsequent interview, or review session, was documented as an expression of informed consent. For the data documenting the process management efforts, approval to use them for this research was obtained from the hospital management and this, in keeping with the national guidelines for social science research (Vetenskapsrådet, 2002), was deemed sufficient, since the research does not concern issues of a private, or sensitive, nature.

The main ethical concern is any potential embarrassment for members of the organization when data stemming from their efforts are analyzed and used to generate research findings. We sought to prevent this risk by trying to avoid identifying individual actors when reporting the research. Furthermore, all data were stored securely in electronic format with password protection, in accordance with Swedish law on treatment of data concerning individuals and regulations at KI.

The potential ethical concerns are, in our view, outweighed by the benefits of this research, not only to the community of healthcare stakeholders and researchers of healthcare and change management, but also to the employees of HUH who were involved in the process management initiative. This research has been made available to them and can potentially help them gain a deepened understanding of their experience at the time and of change management more generally.

Discussion

The thesis illuminated how QI was established in a healthcare organization by way of continuous learning and adaptation — combining the insights and motivation of healthcare professionals, harnessed through a bottom-up approach, with managers' strategic views and accountability, conveyed through a top-down approach — and with the specialized assistance of QI facilitators. While the HUH case revealed weaknesses regarding QI measurement, and the negative implications of that for the conduct and evaluation of QI efforts, the systematic literature review (Thor et al., 2007b) demonstrated that SPC is a versatile tool that can help stakeholders manage and document change in healthcare processes. For further details on the findings and the discussion, including a model put forth in the thesis for how QI is established in a healthcare organization, the interested reader is referred to the thesis (Thor, 2007). What follows is a discussion of some benefits and challenges related to being a QI practitioner researcher.

Methodological considerations

As noted initially, the HUH research represents an opportunistic observational study of a natural experiment. Most of the data were collected as part of the hospital's process management initiative, which was run by the hospital, not by outside researchers. This brings both advantages and disadvantages. The advantages include that data were being collected because they were needed for practical purposes in the QI efforts. This should, in theory at least, increase the accuracy and completeness of the data collection. The facilitators furthermore sought to make the data useful for *both* practical *and* subsequent research purposes which should serve to increase data quality. There were no additional efforts required of anyone generating or documenting these data because of the research plans; data were collected the ways they were irrespective of those plans. On the other hand, the facilitator group members, in particular, were mindful of the plans for subsequent research which probably contributed to a higher degree of motivation in collecting and documenting data than would have been the case without any such plans. As Solodky et al., observed: "Psychological theory suggests that partners are better data collectors than subjects." (Solodky et al., 1998: p. AS19) The accuracy of the data collected was also strengthened by the way that transcripts of dictations were checked against the original sticky notes by the involved facilitator, and validated by review of the involved team members. Similarly, the facilitators reviewed transcriptions of their progress notes, to ensure their accuracy.

The disadvantages of data collection being a part of the hospital's process management initiative include that it may have failed to capture data of relevance to this research (but not to the QI efforts in practice), although few such instances have become apparent in this research, apart from the weakness of measurement of QI indicators noted above. Also, the fact that participants collected data on their own efforts, for subsequent research, raises the possibility of bias — conscious or not. It is not clear, however, in which direction such bias would operate, nor how. It may, in theory, have been tempting to fail to document instances that made the person documenting the data "look stupid", or to favor data perceived as "socially desirable" (a term used for surveys and interviews connoting the risk that respondents modify their responses to "please" the researcher or to "look good") (Robson, 2002). The fact that the research questions were not specified from the outset, and were finalized after the end of the data collection period, should have reduced the likelihood that data collection was somehow influenced to yield particular answers. Instead, the research questions were developed pragmatically to "fit" with the available data, in light of the research problem identified. For instance, we were not able to carry out quasi-experimental research based on these data, and instead focused on the "how" questions

specified above, which could be meaningfully addressed through a case study analysis of the available data.

The fact that as a researcher, I was also intimately involved in the organizational change efforts under study can prompt similar questions about objectivity and neutrality, but it was, arguably, an asset in the research reported in the thesis. The justification for such a role is that it simultaneously offers *access* and *understanding of the context* for the issue under study, without which the study would be much more difficult, if not impossible, to perform. In the words of Gummesson: "Access refers to the opportunities available to find empirical data (real-world data) and information"; what he calls the management researcher's "Number 1 Challenge" (Gummesson, 1991: p. 11). Organizational research requires understanding of the context because "organisational behaviour can only be understood in context" (Ferlie, 2001: p. 25).

When analyzing the data, I benefited accordingly. I believe that I was able to "read between the lines" in the data. Reading accounts of various sessions I had attended frequently brought memories back to me of the situation — the room, the people, and sometimes even the atmosphere — in ways that reach beyond what I can articulate. This raises, of course, the possibility too, that I have "misread between the lines" and added meanings not actually justified by the data. This is one reason to interpret this kind of research with a certain degree of caution. Ways to balance this kind of risk, which we employed, include *triangulation*, using multiple data sources and having multiple researchers review the material (Fulop et al., 2001). In line with the arguments above, I think it would have been much more difficult — perhaps prohibitively so — for someone who had not been part of the process management initiative to interpret and analyze the data.

In a similar vein, we noted in our systematic literature review (Thor et al., 2007b), that our reading of the 57 SPC articles was informed by our understanding of QI in general, and of SPC in particular, as well as of healthcare. Having background experience in both areas, as most of us did, was helpful for interpreting and analyzing the SPC data.

This leads us now into a consideration of the research methodology. As stated by Yin, the case study design is useful for studies of *how* a phenomenon of interest evolves over time, in a particular "real world" context, in situations over which the researcher has limited control (Yin, 1994; Yin, 2003). Case studies allow, and are strengthened by, the combination of multiple data collection methods, spanning both qualitative and quantitative data, as exemplified in this thesis. Thus, "the case study's unique strength is its ability to deal with a full variety of evidence — documents, artifacts, interviews, and observations" (Yin, 2003: p. 8).

Much of the data collection was carried out as *participant observation*. With roots in anthropological field research, this is — as the term implies — when one or several

researchers "engage in the daily life of the group or setting under study. They watch, listen and record what happens in the everyday interactions, involving themselves [in] ongoing activities" (Murphy, 2001: p. 45). This is what the facilitators did in the HUH QI efforts. Given that several facilitators, including myself, also engaged in the research concerning those efforts, does the approach not qualify as *action research?* No. A distinctive difference between these two approaches is that action research "involves opportunistic planned interventions in real time situations and a study of those interventions as they occur, which in turn informs further interventions" (Coghlan & Casey, 2001: p. 674). While parts of this definition apply also to the HUH case study research, that research was not used to inform the conduct of the QI program, and it was observational, rather than interventionist. As stated before, as researchers, we had limited control over the decisions and actions taken at the hospital. The HUH QI program was not a research intervention; it was a natural experiment in which we were participant observers. To complicate the distinction further, however, QI efforts — such as those studied here — share many features with action research, in that both involve a "participatory process [which is] educative and empowering, involving a dynamic approach in which problem identification, planning, action and evaluation are interlinked" (Waterman et al., 2001: p. 11, "A definition of action research").

Doing research at home

Nevertheless, the literature on action research, particularly on doing research in one's own organization, offers insights that apply also to the case study research discussed here. Issues particular to this situation include the pre-understanding of context and phenomena of interest prior to initiation of a research project; role duality when being both a member of the organization and a researcher studying the same organization; and the frequent need to manage organizational politics in these dual roles (Coghlan & Casey, 2001). Being an insider to the organization offers valuable pre-understanding and contacts, although it may also lead the researcher to assume "too much" about what is going on in certain situations, and it can impose certain barriers that would not apply to a researcher from outside, such as the ability to access certain people, groups, or settings. Role duality can involve having "to deal with the dilemma of writing a report on what you have found, and dealing with the aftermath with superiors and colleagues, if you do, on the one hand, and doctoring your report to keep your job, on the other" (Coghlan & Brannick, 2001: p. 51).

Pre-understanding — the understanding of an issue that a researcher brings to a study of that issue — is important, since it enables, but also restricts, the learning that

can result from such study (Gummesson, 1991; Alvesson & Sköldberg, 2000). There is, argue Alvesson and Sköldberg, "in interpretation always an irreducible moment of reshaping, of subjective creativity with its point of departure in the researcher's already pre-existing frames of reference. The researcher is never *tabula rasa*" (Alvesson & Sköldberg, 2000: p. 68). So, what pre-understanding did I bring upon entering this research?

My perspective was that of a junior physician — steeped in the biomedical research tradition that dominates medical school — who had become increasingly concerned about, and fascinated by, how (healthcare) organizations function, and how their members manage change. My perspective was also influenced by my management studies at Harvard prior to the research project. Although based at a medical university, my research has more of a social science character than a natural science one, in that it concerns how human beings collaborate in an organization to achieve intentional change, notwithstanding that such change often aims to increase the use of biomedical technologies which are founded on research closer to the natural science paradigm. As indicated above, I saw QI as a promising approach to addressing many of the problems I had encountered in healthcare, personally or through the literature. While I certainly entered this research with a positive attitude towards QI, I was, and am, genuinely interested in understanding whether and how it works, and can be enhanced. My hope is to ultimately help improve healthcare, including the working conditions for those who take care of patients, and thereby to improve human health.

While reporting here on research to illuminate how QI was established in a healthcare organization, it is worth noting that this type of question can be addressed in other ways. Non-fiction prose, for instance, as exemplified by the insightful essays on quality and safety in healthcare by Harvard surgeon Atul Gawande (Gawande, 2002; Gawande, 2007), which represent a form of journalistic, and partly autobiographical, inquiry. As a final note here, then: what makes the work reported in my thesis a form of research rather than, say, journalism? Some hallmarks are the explicit approach to framing the problem (as a "research problem") and to data collection; the explicit use of research methods; the explicit reflection on that use, and the attempt to link the work both internally from the initial question all the way to the conclusion, and externally to the wider research literature.

Notes on being a practitioner-researcher

The model, for me, of a practitioner-researcher is the multifaceted role many of my clinical colleagues inhabit, where they combine, to varying degrees, clinical practice with research and teaching, and sometimes also management responsibility.

I consider this one of the assets of the medical profession — the close ties between these different domains.

I have had the privilege to combine the role as researcher with the role of being a QI practitioner in a hospital, working in close collaboration with both clinicians and managers. So for me, the practice part has entailed supporting QI efforts, including process management, EBM and patient safety efforts. While I have not been seeing patients, I view my work as helping those who do take care of patients to improve their work, thereby helping patients as well.

It has been valuable to alternate between theory and practice — what I have learned through my research, I have been able to relate to the real world. I have also benefited in my role at the hospital from being exposed to research-based knowledge. For me, theory and practice have been mutually reinforcing. I agree with others that this can be a model for the future, even though it will not be easy: "Integrating [Evidence-Based Medicine and Evidence-Based Management] also requires practitioners who are aware of and able to draw on evidence from both. Few physicians read management studies; few managers read clinical studies; and few persons read all relevant studies within their own field" (Shortell et al., 2007: p. 674).

The obvious challenge of being a practitioner-researcher is that it takes time. This helps explain why it took me over 10 years to complete my thesis. A potential risk therefore is that the data becomes out-dated. A potential solution would be to speed up the research efforts by conducting parts of the project on a full-time basis.

Conclusion

Being a practitioner-researcher brings particular advantages and challenges to the research endeavor. It can bring unique access to, and familiarity with, a real-world setting conducive to management research of high relevance. The challenges include the potential conflicts inherent in the role-duality, the risk of bias when collecting and interpreting data, and the time required to engage in both research and practice. Practitioner-researchers arguably have the potential to generate new knowledge which outsiders might not be able to generate, and also to provide a direct link back from the theoretical sphere to the practice setting.

References

Alvesson, M. & Sköldberg, K., 2000. *Reflexive methodology: new vistas for qualitative research.* Thousand Oaks, CA: Sage.

Coghlan, D. & Brannick, T., 2001. *Doing action research in your own organization.* London: Sage.

Coghlan, D. & Casey, M., 2001. Action research from the inside: issues and challenges in doing action research in your own hospital. *Journal of Advanced Nursing,* 35, p. 674-682.

Eriksson, S. & CODEX (2000-2006) *Informed consent [*in Swedish: *Informerat samtycke].* Vetenskapsrådet (The Swedish Research Council). The Centre for Bioethics at Uppsala University. Karolinska Institutet. Available at: http:// www.codex.vr.se/oversikter/manniskor/samtycke.html [accessed 2007-08-05].

Ferlie, E., 2001. Organisational studies. In Fulop, N., Allen, P., Clarke, A. & Black, N., ed. *Studying the organization and the delivery of health services: research methods.* London: Routledge. p. 24-39.

Fulop, N., Allen, P., Clarke, A. & Black, N., ed., 2001. *Studying the organization and the delivery of health services: research methods.* London: Routledge.

Gawande, A., 2002. *Complications: a surgeon's notes on an imperfect science.* New York: Metropolitan Books.

Gawande, A., 2007. *Better: a surgeon's notes on performance.* New York Metropolitan.

Greenhalgh, T., et al., 2004. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Quarterly,* 82, p. 581-629.

Gummesson, E., 1991. *Qualitative methods in management research.* Newbury Park, CA: Sage

Murphy, E., 2001. Micro-level qualitative research. In Fulop, N., Allen, P., Clarke, A. & Black, N., ed. *Studying the organization and the delivery of health services: research methods.* London: Routledge. p. 40-55.

Øvretveit, J. & Gustafson, D., 2002. Evaluation of quality improvement programmes. *Quality and Safety in Health Care,* 11, p. 270-275.

Robson, C., 2002. *Real world research: a resource for social scientists and practitioner-researchers.* Oxford: Blackwell Publishers.

Shortell, S. M., Rundall, T. G. & Hsu, J., 2007. Improving patient care by linking evidence-based medicine and evidence-based management. *JAMA,* 298, p. 673-676.

Solodky, C., et. al., 1998. Patients as partners in clinical research: a proposal for applying quality improvement methods to patient care. *Medical Care,* 36, p. AS13-AS20.

Thor, J., 2007. Getting going on getting better: how is systematic quality improvement established in a healthcare organization?: implications for change management theory and practice. Stockholm: The Medical Management Center. Department of Learning, Informatics, Management and Ethics. Karolinska Institutet. Dissertation.

Thor, J., et. al., 2004a. Getting going together: can clinical teams and managers collaborate to identify problems and initiate improvement? *Quality Management in Health Care,* 13, p. 130-142.

Thor, J., et. al., 2007a. Evolution and outcomes of a quality improvement program. Submitted. Manuscript.

Thor, J., et al., 2007b. Application of statistical process control in healthcare improvement: systematic review. *Quality and Safety in Health Care,* 16, p. 387-399.

Thor, J., et al., 2004b. Learning helpers: how they facilitated improvement and improved facilitation. lessons from a hospital-wide quality improvement initiative. *Quality Management in Health Care,* 13, p. 60-74.

Walshe, K. & Freeman, T., 2002. Effectiveness of quality improvement: learning from evaluations. *Quality and Safety in Health Care,* 11, p. 85-87.

Waterman, H., Tillen, D., Dickson, R. & de Koning, K., 2001. Action research: a systematic review and guidance for assessment. *Health Technology Assessment,* 5, p. iii-157.

Vetenskapsrådet, 2002. *Principles for research ethics in the humanities and social sciences (in Swedish: Forskningsetiska principer inom humanistisk-samhällsvetenskaplig forskning).* Utgivare: Vetenskapsrådet. Available at: http:// www.vr.se/download/18.6b2f98a910b3e260ae28000360/HS_15.pdf [accessed 2007-08-04].

Yin, R. K., 1994. *Case study research: design and methods.* Thousand Oaks, CA: Sage.

Yin, R. K., 2003. *Case study research: design and methods.* Thousand Oaks, CA: Sage.

Preventing Ventilator-Associated Pneumonia in 18 Hospitals in Thailand Using Collaborative Quality Improvement Methods

AKEAU UNAHALEKHAKA*
SILOM JAMULITRAT**
VIRASAKDI CHONGSUVIVATWONG**
JOHN ØVRETVEIT***

Introduction

Hospital-acquired infections (HAI) are infections that occur because patients acquire pathogenic organisms when they receive treatment in hospital (Mayhall, 1999). HAI occurs in at least 5-10% of patients hospitalized to hospitals in developed countries and the prevalence is greater in developing nations. HAI clearly affect morbidity, mortality and the economic burden of patients and healthcare (Wenzel, 1991; Jarvis, 1996; Leroyer et al., 1997; Wenzel, 1995). The highest prevalence of HAI is in the intensive care unit (ICU) (Emmerson, 1995).

Ventilator-associated pneumonia (VAP) is the most important HAI associated with high case fatality rates, ranges from 24 to 50% and can reach 76% in some specific settings or when lung infection is caused by multi-drug resistant pathogens. As many as 15% of ICU mechanical ventilated patients develop VAP (Institute for Healthcare Improvement). Mechanical ventilated patients who developed VAP appear to have a 2- to 10-fold higher risk of death compared with those without pneumonia (Bergmans & Bonten, 2004; Craven & Steger, 1995).

A surveillance study of HAI conducted in 55 ICUs of 46 hospitals in 8 developing countries during 2002-2005 found VAP posed the greatest risk (41% of all device-associated infections) with a rate of 24.1 per 1,000 ventilator-days. The crude mortality rate of VAP was 44.9% (Rosenthal, Maki, Salommao, Alvarez-Moreno, 2006).

A study of the financial impact of nosocomial pneumonia was conducted in 6 adult intensive care units in Argentina from July 1998 to June 2002. The mean extra length of stay of VAP cases was 8.95 days, the mean extra antibiotic cost was US$ 996, mean extra total cost was US$ 2,255 and the extra mortality was 30.3% (Rosenthal et al., 2005).

* Faculty of Nursing, Chiang Mai University Chiang Mai, Thailand (akeau@hotmail. com).
** Faculty of Medicine, Prince of Songkla University, Thailand.
*** The Karolinska Institute, Stockholm, Sweden.

In Thailand VAP is also the most severe infectious complication among critically ill patients receiving mechanical ventilation. Data revealed that the incidence rate of VAP ranged from 10.8-18.8 per 1000 ventilator-days (Thongpiyapoom et al., 2004; Judaeng, 2002). The VAP rate in a general hospital during 3-month period of surveillance was 39.3 per 1,000 ventilator-days (Ruengwattana et al., 2005). Mean antibiotic cost of VAP treatment in a medical ICU of a university hospital per one case of VAP was found to be approximately US$ 560 (Judaeng, 2002).

Prevention of VAP has proven to be difficult and needs high cooperation and strong commitment from multidisciplinary hospital personnel. There is a wealth of knowledge that is simply not being used and broad variation exists in the provision of care. Collaboration is needed to expose these gaps and accelerate the diffusion of existing knowledge into practice (McLaughlin, 1999). Evidence-based knowledge and experience sharing are crucial for the reduction of VAP. Even though many hospitals pay attention and make a concerted effort to reduce VAP, the outcome was not satisfactory and efforts were not sustained.

Following suggestions from the overseas member of our team, a collaborative quality improvement project was to reduce VAP problems among 18 hospitals in Thailand (Øvretveit, 2002). This study aimed to assess the effectiveness of this collaborative in preventing VAP among ICU patients and to explore the attitudes of relevant hospital personnel on the application of the collaborative concept, to determine problems, obstacles and factors determining success in applying the concept in the prevention of VAP.

Methods

Eighteen secondary and tertiary care hospitals, with an Infection Control Committee (ICC), at least one Infection Control Nurse (ICN) per 250 hospital beds and good HAI surveillance system were invited to participate in the project. The number of hospital beds ranged from 150 to 1000. The local research team consisted of a physician as a team leader, ICU staff, ICNs and multidisciplinary professional. The team had their hospital director signed project approval and support for the collaboration. Nine medical, 4 surgical, 7 medical-surgical and 2 neurological ICUs with 8-10 beds of 18 hospitals participated in the project.

The project was conducted for 16 months, between February 2004 and May 2005. The main project activities consisted of 2 national workshops conducted for all 18 hospitals and 6 regional workshops conducted twice for hospitals in each region. Four representatives from each hospital team, consisted of a physician (team leader), an ICN, ICU staff and /or other team member, participated in each project workshop.

In the first national workshop, collaborative method and the project intervention were introduced. Knowledge on CQI was revised and successful CQI projects (e.g. strengthening surveillance system, promoting hand hygiene) were presented, followed by active discussion and agreement on translated definition criteria of pneumonia (CDC. NHSN, 2005) and guideline for prevention of hospital acquired pneumonia (Tablan et al., 2004), VAP surveillance system, VAP data collection form and VAP prevention gap analysis form.

After participating in the first national workshop, all hospitals established their CQI-VAP team consisted of a physician as a team leader, an ICN, ICU staff, and other relevant multidisciplinary professional as they needed, e.g. physiotherapist, nutritionist, pharmacist, chief of central sterile supply department. Teams implemented VAP prevention activities according to their situation and problems.

VAP surveillance system and VAP prevention guideline were revised and strengthened. Education on VAP prevention was conducted for ICU staff and all relevant personnel to fill the gap between actual and evidence-based practices. CDC guideline on Prevention of nosocomial Pneumonia (Tablan et al., 2004) and the Institute for Healthcare Improvement (IHI) ventilator bundle (Resar et al., 2005), were emphasized. Hand hygiene and alcohol-based hand rub were promoted. Oral care practices and management of respiratory care equipment were improved.

The first and second regional workshops were carried out in May and October 2004, three and nine months after the first national workshop, respectively. During the regional workshops, teams reported their interventions and progress, changes in their practices and other aspects, shared and exchanged knowledge and experiences and discussed different strategies and solutions.

In May 2005, the final national workshop was conducted to conclude the project outcome. Team representatives presented the interesting and effective activities, e.g. training of respiratory care ward nurse (RCWN), expansion of collaborative approach to other ICUs and units, effective system for sending and interpreting chest x-rays and oral care for mechanically ventilated patients. Brainstorming session was also conducted to determine the effective VAP prevention interventions and the appropriate collaborative model for hospitals in Thailand.

Results

Data on VAP morbidity, mortality and cost of treatment during May 2004 and February 2005 were used to assess the impact of collaborative method. Data were

aggregated into five 2-month periods for the purpose of comparison. The results were as followed.

Incidence of VAP

During the 10-month period, 358 VAP cases took place with an aggregate of 28,413 ventilator-days. VAP rate of hospitals varied from 0 to 26.6 per 1000 ventilator-days and the overall rate was 12.6 per 1000 ventilator-days, with 95% CI 11.3-13.9. The rate in the first period was 15.9 per 1000 ventilator-days and gradually decreased to 14.3 and 13.8 in period 2 and period 3, respectively. The sharp rate reduction was observed in period 4 but slightly climbed up in period 5. Overall decreasing trend was statistically significant ($p < 0.01$) (Table 1).

Table 1
VAP rate, mortality rate, case-fatality rate and costs
of VAP treatment of 18 hospitals by 2-month period

	MAY-JUN 2004 (PERIOD 1)	JUL-AUG 2004 (PERIOD 2)	SEPT-OCT 2004 (PERIOD 3)	NOV-DEC 2004 (PERIOD 4)	JAN-FEB 2005 (PERIOD 5)	TOTAL	p-VALUES LINEAR FOR TREND
No. of ventilated patients	647	584	603	645	523	3002	–
No. of ventilator-days	4901	5605	6284	6580	5043	28413	–
No. of VAP cases	78	80	87	63	50	358	–
No. of VAP deaths	21	22	17	18	19	97	–
VAP rate (per 1000 ventilator-days)	15.9	14.3	13.8	9.6	9.9	12.6	0.002
Mortality rate (per 100 ventilated pts)	3.2	3.8	2.8	2.8	3.6	3.2	0.86
Case-fatality rate (%)	26.9	27.5	19.5	28.6	38.0	27.1	0.37
Cost of antimicrobial treatment, range (US $)	7.5–3198.8	1.95–10686.6	2.1–2149.2	24–3647.3	12.7–1777.8	1.95–10686.6	–
Total costs of treatment	52270.8	51426.4	39525.8	33662.2	16509.1	193394.2	0.01
Cost per 1 VAP (Mean ± S.D.)	573.7± ±634.3	685.1± ±1294.1	430.1± ±386.4	637.7± ±672.5	416.9± ±430.3	552.6± ±777.2	0.40

When the VAP rate is compared in each period with period 1, the rate gradually decreased approximately 7% and 12% in period 2 and 3, respectively and decreased

significantly, almost 40%, in period 4 (p = 0.006). During these five 2-month periods, the VAP rate had a statistically significant decreasing trend (p = 0.002).

Mortality

Among the 3002 mechanically ventilated patients, 97 died. The overall mortality rate was 3.2 per 100 ventilated patients. Mortality rate in each period ranged from 2.8 to 3.8 per 100 ventilated patients and was not statistically significantly different (p = 0.86).

Overall case fatality rate of VAP was 27.1% with the range of 19.5% to 38%. There was no statistically significant reduction trend of case fatality rate during the study period (p = 0.37) (Table 1).

Cost of VAP treatment

Total costs of treatment reduced significantly during each period from US$ 52,270.8 to 16,509.1 (p = 0.01). Cost per one ventilated patient also decreased significantly (p = 0.02). An average cost of treatment per one VAP case ranged from US$ 416.9 to 685.1. Mean cost during the 10-month period was US$ 552.6 ± 777.2. Costs of one VAP treatment were not in a decreasing trend (p = 0.4) (Table 1).

Opinions about using the collaborative method in preventing VAP

Two hundred and ninety eight self-administered questionnaires (90%) were obtained from 8 team leaders, 140 team members and 150 ICU nursing staff.

87.8% and 75.3% of team members and ICU nursing staff respectively informed that this method could be applied effectively in their hospitals. It was a jump up process which create motivation of administrator and personnel. It helped create team work, conveniently solve VAP problems, obtain meaningful information from expert consultation, knowledge and experience sharing among participants during the project workshop to ask for administrative support and create more compliance among hospital personnel and also helped them save time and resources in prevention of VAP.

95.3% and 87.4% of team members and ICU nursing staff would continue applying this method although the project ended. 84.5% and 72% of these 2 groups

respectively believed that changes resulted from the project would last long and hospital personnel would still comply with the VAP guideline.

84.6% of all respondents suggested that other collaborative projects should be conducted in the future, in order to improve other Infection Control activities, e.g. prevention of catheter-associated urinary tract infection, surgical site infections and so forth.

Information from in-depth interviewed 10 team leaders, 18 Infection Control Nurses, 28 team members and 38 ICU nursing staff using structured interviewing form revealed that participating personnel had a positive attitude towards applying collaborative quality improvement method in prevention of VAP. Their opinions on advantages of collaborative method included.

Firstly, caring of mechanically ventilated patients was rapidly improved. Relevant multidisciplinary personnel realized the severe impact of VAP, gained more knowledge and understood their roles in prevention of VAP. All these created well cooperation, collaboration and strong teamwork among multidisciplinary team and hospital personnel, both intra and inter departments. Relevant personnel were willing to improve and change their practices according to the guideline and worked eagerly. They learnt and consulted each other more closely which led to good relationship and more rapid work. Hospital administrators were interested about the outcome and gave more support. Many participating teams expanded the collaborative concept and method to other units in their hospital. There were many relevant sub projects conducted in many participating hospitals. Some team leaders perceived that the progress after participating in the collaborative project was a great jump forward their past experiences.

Secondly, nursing care of ventilated patients was standardized according to VAP guideline. Nurses gained more knowledge and more confidence in caring ventilated patients. They complied with the VAP guideline and solved problems more systematically. Standardized and effective mechanically ventilated patient care practices led to reduction of VAP rate, shorten length of hospital stay, less cost of antibiotic treatment and quality of life of the patients.

Thirdly, diagnosis of VAP were more reliable and timeliness according to clear VAP definition and effective VAP surveillance system. Clinicians helped diagnose VAP, ICN assessed the completeness of data and efficiency of VAP surveillance of ICU nursing staff.

About the sustainability of their work after the project was ended, many of team members informed that they would continue their work and try to expand this method to other units of their hospital.

Factors influenced the success of using collaborative method

CQI-VAP team members mentioned that factors influenced with the success of using collaborative method were: the cooperation of ICU staff and multidisciplinary professionals, the intention and commitment of ICU head nurse and nursing staff and effective CQI-VAP teams, support of hospital directors, support from Infection Control Nurse and Infection Control Ward Nurse, and compliance of ICU staff on VAP prevention guideline.

Recommendations about and limitations of the collaborative method in Thai hospitals

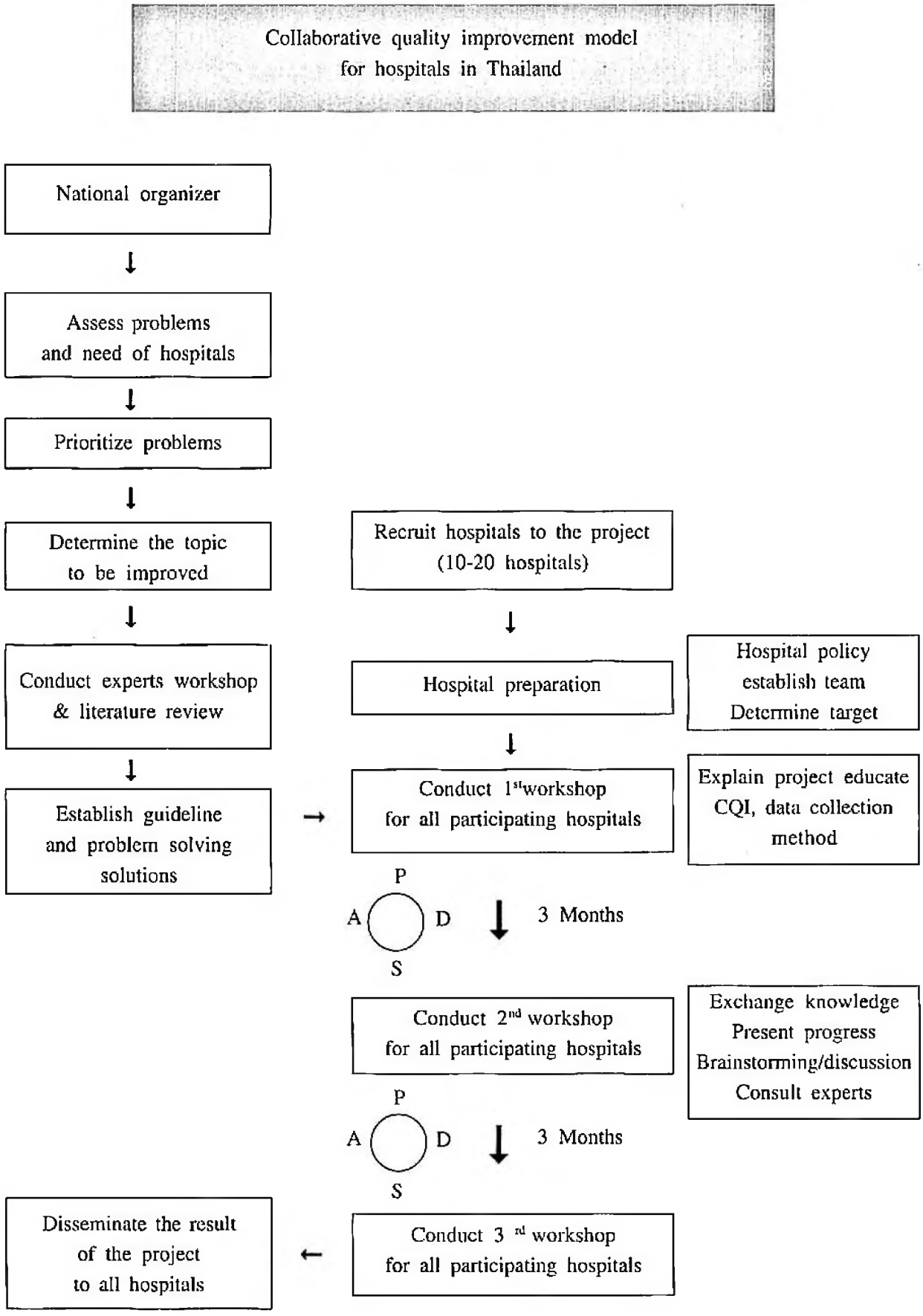
Team members suggested inviting head of all relevant departments to participate in the project workshop, especially in the first national workshop, so that they understand and have opportunity to meet and exchange ideas with other hospitals. Project organizer should manage study visit to successful and best practices hospitals for teams in each region in order that they can observe real situation and directly discuss with hospital personnel. Project organizer should visit and supervise CQI-VAP teams every 3 month. After the project ended, project organizer should conduct scientific conference for hospital teams to present their work, share knowledge and exchange experiences with non-participating hospitals. Communication via project webboard should be strengthened.

Approximately three-forth of CQI-VAP team members and ICU nursing staff indicated over workload as their main obstacle. Team leaders seldom join team meeting owing to their limited time. Other problems included lack of knowledge and understanding on CQI among hospital personnel (55.7%), difficulty in conducting team meeting regularly owing to the different time of work (50%) and no supervisor or facilitator to guide teams (46.6%).

Inconvenience to record VAP surveillance data via project website and to communicate with other teams via electronic mail since there was no internet connection in the offices of some teams. Lack of responsible national organization and direct consultative support and advice to individual hospitals owing to limited resources and inconvenience contact via electronic mail.

New collaborative model for hospitals in Thailand

During final national workshop, 18 teams suggested the future collaborative model for hospitals in Thailand. They recommended one-year collaborative project with 10-


Collaborative quality improvement model
for hospitals in Thailand
National organizer
Assess problems
and need of hospitals
Prioritize problems
Determine the topic
to be improved
Conduct experts workshop
& literature review
Establish guideline
and problem solving
solutions
Recruit hospitals to the project
(10-20 hospitals)
Hospital preparation
Hospital policy
establish team
Determine target
Conduct 1st workshop
for all participating hospitals
Explain project educate
CQI, data collection
method
P
A
D
S
3 Months
Conduct 2nd workshop
for all participating hospitals
Exchange knowledge
Present progress
Brainstorming/discussion
Consult experts
P
A
D
S
3 Months
Conduct 3 rd workshop
for all participating hospitals
Disseminate the result
of the project
to all hospitals

20 participating hospitals. In the first 3 months of the project, the national organizer assess and prioritize hospital problems, review literature and evidence-based practices, conduct expert meeting to establish guideline and determine solution for problems. The second 3-month period is for hospital recruitment and hospital preparation. The rest 6 months consist of 3 national workshops conducted for all participating hospitals, not separate by region.

Discussion

Our collaborative project could help reduce VAP rate in the ICUs. It is not possible to point out which specific prevention measure was a major cause of success, since multiple well known effective measures, e.g. placing ventilated patients in a semirecumbent position unless contraindicated, strengthening hand hygiene, improving oral care practices, and improving management of respiratory care equipment, were simultaneously implemented during the project.

This success contrasts with previous experience in developing countries where VAP prevention program cannot be implemented effectively according to the traditional authority structures and the organization culture (Øvretveit, 2002).

Within the collaborative project, a systematic approach has been accelerated (Jain, Miller, Belt, King and Berwick, 2006) and peers were influenced within a reinforcing environment (Kosseff and Niemeier, 2001). The commitment and involvement of the participating teams provide an opportunity to improve the quality of care (Girouard et al., 2001). Besides, participating organizations can learn from the others' experiences, discuss and modify ways to overcome common barriers to fit their own situation (Plsek, 1999).

The overall VAP rate of this collaborative project was lower than the aggregate VAP rate of 55 ICUs of 8 developing countries reported by the International Nosocomial Infection Control Consortium (Rosenthal, Maki, Salommao, Alvarez-Moreno, 2006). Almost all participating hospitals also had overall rate lower than the above report. One hospital did not have VAP case during the 10-month period since they had only 66 ventilated patients and approximately 75% of the patients were on mechanically ventilators for a short period. The chance of under-report and inaccurate VAP cases in this study was unlikely, since the participating teams gave priority to strengthen VAP surveillance system to obtain high quality of VAP data. All VAP cases were confirmed diagnosis by physicians, strictly adhering to the NNIS definition criteria and ICNs assessed the completeness of the data. Besides, the numbers of patients admitted to the ICUs were limited and changes in patient clinical signs were rapidly investigated since such signs were life-threatening. The

reduction of VAP rate during the study period was certainly not the effect of the under-reporting of VAP cases. There were also no identified outbreaks occurring during the study period.

The characteristics of the patients which contributed to VAP occurrence, e.g. severity of illness and age group, during five 2-month periods were quite stable throughout the study period; thus, these variables did not confound the result. It was not possible to compare VAP data with data from other hospitals, since VAP definition and data collection measure were different.

Mortality rate and case-fatality rate of VAP during five 2-month periods did not significantly decrease. Although VAP is associated with mortality, it is not the sole cause of death among ventilated patients. Measures to reduce case-fatality from other causes had not been implemented. There were also no specific measures implemented to reduce these two rates among participating hospitals. However, the overall VAP case-fatality rate was lower than the report of the International Nosocomial Infection Control Consortium and other developing countries (Rosenthal et al., 2006; Rosenthal, Guzman & Orellano, 2003).

As mortality is the product of incidence and case-fatality rate, the reduction in VAP incidence is usually expected to be followed by the reduction in mortality rate. In our study, mortality was not solely caused by VAP. This non-VAP mortality may explain the lack of effectiveness of intervention on case fatality. Moreover, there was no special measure to improve treatment of VAP. Case-fatality rate was, therefore, not reduced.

Conclusion

Existing evidence on VAP incidence rate and opinion of participating hospital personnel suggest that inter-hospital multidisciplinary collaborative quality improvement is effective and can be applied effectively among secondary and tertiary care hospitals in Thailand. The findings suggest that multidisciplinary collaborative quality improvement has the potential to improve patient outcomes. Collaborative quality improvement method can be effectively applied not only in prevention of ventilator-associated pneumonia but also in human resource development among hospitals in Thailand. However, further study to evaluate the sustainability of the project outcome and the effectiveness of VAP surveillance system are needed.

References

Bergmans, D. C. J. J. & Bonten, M. J. M., 2004. Nosocomial pneumonia. In Mayhall, C. G., ed., Hospital epidemiology and infection control. 3rd ed. Philadelphia: Lippincott Williams & Wilkins. p. 311-339.

CDC. NHSN, 2005 [internet]. Criteria for defining nosocomial pneumonia. Atlanta, GA: National Healthcare Safety Network, Centers for Disease Control and Prevention. Available at: http://www.cdc.gov/ncidod/hip/NNIS/members/pneumonia/Final/PneumoCriteriaV1.pdf [accessed 20 October 2007].

Craven, D. E. & Steger, K. A., 1995. Epidemiology of nosocomial pneumonia: new perspective on an old disease. *Chest,* 108, p. 1-16.

Emmerson, A. M., 1995. The impact of surveys on hospital infection. *Journal of Hospital Infection,* 30 (Suppl.), p. 421-440.

Girouard, S., et al., 2001. Pediatric Prevention Network: a multicenter collaborative to improve health care outcomes. *American Journal of Infection Control,* 29, p. 158-161.

Institute for Healthcare Improvement (IHI), 2007. Intensive care. Available at: http://www.ihi.org/IHI/Topics/CriticalCare/IntensiveCare [accessed 20 October 2007].

Jain, M., et al., 2006. Decline in ICU adverse events, nosocomial infections and cost through a quality improvement initiative focusing on teamwork and culture change. *Quality & Safety in Health Care,* 15, p. 235-239.

Jarvis, W. R., 1996. Selected aspects of the socioeconomic impact of nosocomial infections: morbidity, mortality, cost, and prevention. *Infection Control and Hospital Epidemiology,* 17: 552-557.

Judaeng, T., 2002. Pneumonia in patients admitted to a medical Intensive Care Unit, Siriraj Hospital. *Bulletin of Nosocomial Infection Control Group of Thailand,* 12, p. 2-13.

Kosseff, A. L. & Niemeier, S., 2001. SSM health care clinical collaboratives: improving the value of patient care in a health care system. *Joint Commission Journal on Quality Improvement,* 27, p. 5-19.

Leroyer, A., & et al., 1997. Prolongation of hospital stay and extra costs due to hospital-acquired infection in a neonatal unit. *Journal of Hospital Infection,* 35, p. 7-45.

Mayhall, C. G., 1999. Hospital epidemiology and infection control. 2nd ed. Philadelphia: Lippincott Williams & Wilkins.

McLaughlin, C. P. & Simpson, K. N., 1999. Does TQM/CQI work in health care? In McLaughlin, C. P. & Kaluzny, A. D., 1999. Continuous quality improvement in health care: theory, implementation, and applications. 2nd ed. Gaithersburg: Aspen. p. 34-56.

Øvretveit, J., 2002. Improving health care quality in developing countries. *Quality and Safety in Health Care,* 12, p. 200-201.

Plsek, P. E., 1999. Quality improvement methods in clinical medicine. *Pediatrics,* 103, p. 203-214.

Resar, R., et al., 2005. Using a bundle approach to improve ventilator care processes and reduce ventilator-associated pneumonia. *Joint Commission Journal on Quality and Patient Safety,* 31, p. 243-251, May 2005.

Rosenthal, V. D., Guzman, S. & Orellano, P. W., 2003. Nosocomial infections in medical-surgical intensive care units in Argentina: attributable mortality and length of stay. *American Journal of Infection Control,* 31, p. 291-295.

Rosenthal, V. D., et al., 2005. The attributable cost and length of hospital stay of nosocomial pneumonia in intensive care units in three hospitals of Argentina: a prospective, matched analysis. *American Journal of Infection Control,* 35, p. 157-161.

Rosenthal, V. D., et al., 2006. Device-associated nosocomial infections in 55. intensive care units of 8 developing countries. *Annals of Internal Medicine,* 145, p. 582-591.

Ruengwattana, S., et al., 2005. Situation of pneumonia among mechanically ventilated patients, Lumpun Hospital. *Bulletin of Nosocomial Infection Control Group of Thailand,* 15, p. 44-53.

Tablan, O. C., et al., 2004. Guideline for preventing health-care associated pneumonia 2003. *MMWR,* 53 (RR03), p. 1-36.

Thongpiyapoom, S., et al., 2004. Device-associated infections and patterns of antimicrobial resistance in a medical-surgical intensive care unit in a university hospitals in Thailand. *Journal of the Medical Association of Thailand,* 87, p. 819-824.

Wenzel, R. P. & Pfaller, M. A., 1991. Feasible and desirable future targets for reducing the costs of hospital infections. *Journal of Hospital Infection,* 18 (Suppl. A), p. 94-98.

Wenzel, R. P., 1995. The economics of nosocomial infections. *Journal of Hospital Infection,* 31, p. 79-87.

Perceived Benefits and Challenges of Introducing Accreditation in Pakistan

Sylvia Sax*
Nadeem Ahmed**
Barbara Donaldson***
Martina Bardroff****
Bernd Appelt*****

Introduction

Improving quality of healthcare services to achieve health outcomes is an ongoing topic of discussion (Roberts, M. et al., 2004; Smits et al., 2002; Travis et al., 2004; World Bank, 2006). Most low and middle income countries include in their health policy and strategy documents the improvement of healthcare quality as a priority (Tanzania, 1998, India, 2002, Pakistan, 2001 and many more). But there is little published information on what methods are appropriate for improving quality within these countries, though several authors have contributed to this discussion (Doyle, 1999; Massoud et al., 2001; Nandraj et al., 2001; Catsambas, 2002; Shaw, 2003; Øvretveit, 1997, 2002, 2003 & 2004). There is also a growing awareness of the need for context specific quality improvement programmes, methods and tools (Nandraj et al., 2001; Øvretveit, 2004; Peters, Rao & Fryatt et al., 2003; Potter, 2006). But when management capacity is low and knowledge about quality management is poor, how can key stakeholders in developing countries explore context specific quality improvement methods and tools? We used a method common to Social Marketing research and management (SWOT analysis) with a group of key health sector stakeholders in North West Frontier Province, Pakistan to explore the appropriateness of introducing accreditation within their healthcare setting.

Many authors have indicated the need for quality evaluation methods conducive to resource poor settings (Catsambas, T. et al., 2002; Øvretveit, 2002; Shaw, 2003) but, as Smits et al. (2002) point out, "developing countries rarely monitor the quality of services delivered". In 2003 the WHO and ISQUA published a global review of

* Dept. of Tropical Hygiene and Public Health, University of Heidelberg, Germany.
** Health Regulatory Authority, Peshawar, North West Frontier Province, Pakistan.
*** International Society for Quality in Healthcare, Wellington, New Zealand.
**** EPOS Health Consultants, Bad Homburg, Germany.
***** GTZ (German Technical Cooperation), Health Secretariat, Peshawar, Pakistan.

Quality and Accreditation in Healthcare Services, the purpose of which was to give "examples from around the world of quality structures and processes that might inform local improvement of health services, especially in the developing countries" (Shaw, 2003). Within the methods they reviewed, Accreditation is increasingly used in many counties. Accreditation is an external review of quality with four principal components:

- It is based on written and published standards;
- Reviews are conducted by professional peers;
- The accreditation process is administered by an independent body;
- The aim of accreditation is to encourage organizational development (Montagu, 2003).

Three health system characteristics frequently mentioned as prerequisites for accreditation are:

a) an established health sector regulatory mechanisms;
b) able health facility managers with powers to make resource and process decisions and;
c) health professionals (including paramedical staff) with defined accountability mechanisms (Shaw, 2003; Shaw, 2004; Potter, 2006; Roberts, M. et al., 2004). These characteristics are often weak in low and middle income countries.

In the last year, the Government of North West Frontier Province (NWFP) of Pakistan has invested in ways to improve quality and build capacity for assuring quality in health care delivery within the province. The German development agency (GTZ) has been working with the provincial health department through its Health Sector Research and Reform Unit (HSRRU) on several reform areas including capacity building in quality management and development of health care standards for primary and secondary health care, with the plan to accredit health care facilities against the standards. A Health Regulatory Authority (HRA) was established as an autonomous body in 2002 to regulate the private health sector in the province. In the first phase of its work the HRA has focussed on the registration of private healthcare providers in the province. So far approximately 1100 private providers are registered, representing around 20% of private providers in the province. This registration process has involved the listing of current facilities without any evaluation or evidence of their ability to provide a quality service. The HSRRU has recently commissioned the HRA to introduce Accreditation within the public and private sectors.

Background

Pakistan has a population of 145 million and its people are multiethnic with great linguistic diversity. It has a per capita income of US$ 420-460, with a low expenditure on health of 0.7% GNP, a low literacy rate of 38.9% and a population doubling time of 25 years (Economic Survey of Pakistan, 2001-2). The public Health System is three tiered (federal, provincial and district). Health is constitutionally a provincial matter; but the Federal Ministry sometimes with assistance form international donors implements vertical public health programmes (e.g., Lady Health Workers, various disease control programs e.g. EPI, Malaria, Hepatitis, AIDS, TB). The private sector plays a major role in health service delivery — possibly 70% of all health care takes place in the private sector.

The province of NWFP is the most Northern Province in Pakistan, with Peshawar as its capital. Since 2000 there has been an increasing sense of urgency within the NWFP government to improve the quality of health care services. The key outcomes they have identified are:

- To improve the performance of health care services.
- To increase public safety and reduce risks of poor outcomes.
- To increase public confidence in the quality of health care services.
- To increase accountability of health services to players and the public.

The two key organisations to work towards these outcomes have been the HSRRU and the HRA. The HSRRU was established in 2002 with a focus on Decentralization, Quality Management, Human Resource Management and Health Financing and the HRA with a focus on registration of private health facilities and providers, development of healthcare standards and establishment of Accreditation.

"... in the longer run, economic and political stability can be not just a cause, but a consequence of improvement." (Smits et al., 2002)

Intervention/method

We worked with 42 key stakeholders from both the private (for profit and not for profit) and public sector to introduce healthcare standards and explore the introduction of Accreditation in NWFP. The stakeholders comprised key officials from the Department of Health (DoH), HRA and HSRRU and providers and managers from public and private health care organizations. All key stakeholders had deep

experiential knowledge of the NWFP political, cultural and economic context. Some of the stakeholders had participated in previous workshops to develop a set of draft primary and secondary healthcare standards.

The stakeholders had been invited to a one day Sensitization Workshop on Quality Management. The list of invitees was decided by the HRA and DoH to include key decision-makers at all levels within the health sector from clients to policy makers. The stated purpose of the Workshop was threefold, to:

- introduce key concepts for improving quality,
- create awareness regarding health regulation and various tools for quality improvement and regulation, and
- provide participants with an opportunity to examine if Accreditation would fit into the NWFP health system as a means to evaluate and improve health services.

The morning consisted of presentations on four key topics: (1) the Role of the HRA, (2) an Introduction to International Healthcare Service Standards and (3) the draft NWFP Healthcare Service Standards, and (4) an Introduction to Quality Management, with a special focus on Licensing, Accreditation, and Certification. Approximately half of the participants stayed for the afternoon session where the SWOT analysis tool was used in 4 groups of 9-12 people to explore if Accreditation would be the appropriate tool for the NWFP health system as a means to evaluate and improve quality of health care.

The SWOT Analysis is a planning tool used to examine the *S*trengths, *W*eaknesses, *O*pportunities, and *T*hreats of a project, process, or proposed activity. It involves (1) clarifying for the group using the SWOT the objectives or goals to be achieved by the activity or project and (2) identifying, within that particular context, the internal and external factors that are favourable or unfavourable for the activity. The technique was introduced by Albert Humphrey, in the 1960s and is widely used by businesses and organizations in their decision-making processes, especially in strategic planning (Armstrong, 2006). The SWOT is also a technique commonly used in Social Marketing Research to analyze the social marketing environment with the key aim to clarify the purpose of an intervention, its potential impact or benefit, and challenges that could be faced (Grace-Bishop, K., 2004). The SWOT technique is most successful when those using it have a deep knowledge about the environmental conditions within which the activity or project is being implemented. The SWOT is easy to understand and use, it does not take a lot of time; it is advantageous to use a facilitator, especially in teams or groups who have not worked together before.

The tools used to assist participants to understand the task and work together were:

a) A one page overview of the group work, including a description of the key elements of Accreditation (Standards as the basis of evaluation; self-assessment leading to an action plan for improvement (provider responsibility); the on-site visit for external evaluation (HRA responsibility); Accreditation as the result; 3 years as the length of the result; the Scope focusing on organizational functions; the Approach of accreditation for quality improvement with a client/ patient focus; Accreditation as voluntary & quasi-regulatory; peers as the external evaluators and the final results as a graded score between non-compliance to substantial compliance);
b) A working sheet of the SWOT analysis tool;
c) Verbal and written instructions for undertaking a SWOT analysis; a general introduction to the SWOT analysis was presented to the participants before they were divided into four groups. Facilitators provided clarifications on the SWOT technique to each of the groups upon request.

The division into the groups was made by the three facilitators of the workshop aiming for multi-level and multidisciplinary groups involving personnel from the different organizations within the health system and different occupations. The groupings tried to minimize people participating in groups with their counterparts or superiors, in order to ensure open and transparent discussions. A group leader and scribe were agreed within each group, in some groups this was the same person. The groups visualized their work on large paper using pin boards and we proposed two hours for the exercise. The results of all groups were briefly discussed with the participants in a plenary at the end of the workshop and the overall results were emailed to all group participants within two weeks of the exercise.

The results were coded within each of the four sections of the SWOT tool and then analyzed for emerging themes by one of the facilitators. One other facilitator provided comments and suggested revisions to these themes.

Results

The four groups were very consistent in the results from the SWOT analysis. One group took more time than the others to master the tool but by the end of two and a half hours they had all completed the SWOT analysis.

INTERNAL

Strength sweaknesses

1. Provides for clear goals, boundaries and measures
2. Leads to safer services
3. Acts as a motivator for improvement
4. Provides a means/tool for improvement

1. Relies on accurate measurement, and not everything is measurable
2. Current level of knowledge and education not sufficient to manage and implement the method
3. Requires increased resources
4. Resistance to change, facing the problems could make the situation worse (escalate current conflicts)

EXTERNAL

Opportunities threats

1. Introduces a National agenda to improve Quality
2. Provides research, data on improvements and what works
3. Improved coordination of leadership role for quality
4. Better health outcomes through improvement in provision
5. Change attitude of providers
6. Capacity building
7. Integration of health services (teams, shared ideas, multi-sectoral)
8. Private accredited organizations make more money
9. Save money
10. Supports decentralization
11. Improved inputs, processes and results

1. Stakeholders not supporting the change
2. Government not supporting the change (set bureaucracy)
3. Won't get resources/funding/training needed
4. Change is a threat to many (conflict and confrontation)
5. Lack of change management skills
6. Current issues in the system won't allow change, such as Illiteracy, lack of relevant skills, threat to VIP culture, threat to 'safarish' and the whole system of government,
7. Lack of confidence (tradition)
8. Inflation of user fees of those who are good and accredited

Discussion

Methods

The four groups adjusted to using the SWOT analysis tool and reported that it was easy to use; several people had previously used it in planning exercises. The facilitators observed that all 4 groups were very interactive with excellent participation and discussions.

As a research tool the SWOT technique has several strengths and problems. Probably its greatest strength is that it involves users in the research, this is increasingly emphasized in healthcare research, it is even sometimes required by those sponsoring research. But, because it is based solely on the input of the users, the objectivity of the data can be compromised. A related strength is that it enables the use of the rich experience and knowledge of those involved in day-to-day decision making in the topic area of the research. Thus the data is up to date and relevant, but it also means it can be biased because of the informants. In this workshop the majority of participants were doctors which could have biased the output of the groups; this represents the reality in health management in NWFP and Pakistan which is dominated by this profession and their interests. In addition, the Pakistani culture is hierarchical, and particularly so in NWFP, thus the group process was influenced by the age and positions/perceived power/influence of those within each group.

The sample in our workshop was based on systematic, non-probabilistic sampling. This sampling method enables exploration of a relevant research topic by "a specific group of people who either possess characteristics or live in circumstances relevant to the social phenomenon being studied" (Mays & Pope, 1995). The key themes that emerged can be used as a basis for development of specific research questions and tools. But, the groups were responsible for their own documentation and could have left sensitive or unpopular topics undocumented or their written statements could be biased.

The SWOT technique could have been strengthened in several ways. The reliability of the analysis would have been strengthened if all three facilitators had analyzed the results separately and then compared the outcome. Additionally, an independent assessment of the results by other skilled qualitative researchers would have improved the reliability.

Triangulation, using such methods as interviews with key stakeholders and document reviews, and receiving feedback on the analysis from the participants to see if they regard the results as a reasonable account would have strengthened the study (Mays & Pope, 1995).

Results

The groups identified many challenges and benefits in introducing accreditation into NWFP. Improvements in safety and performance were clearly identified as potential benefits but this will be hindered by resistance to change and issues in resources, both financial and human.

Surprisingly, the lack of basic healthcare regulations and implementation of existing, though scanty, legislation were not mentioned as challenges, though they are prerequisites for introducing Accreditation. The acute lack of regulations, including licensing of healthcare providers against basic requirements in Pakistan has been commented on by many authors (Dodani, S., 2003; Potter, C., 2006; Ismail, Z., 2006). Potter (2006) in his review of progress towards planned autonomy of hospitals in NWFP found that, "There is no regulation and not even a hint of quality assurance" (p. 20). Accreditation cannot replace basic health sector legislation and regulations, though many accreditation programmes often fulfil a quasi-regulatory role. The tension between the proposed voluntary nature of accreditation and the reality of governments legislating that providers have an External Quality Assessment process such as Accreditation is emerging as a point of debate in healthcare quality literature (Shaw, 2004: p. 15). Montagu states: "The effectiveness of accreditation is dependent on its voluntary nature, non-threatening process and interactive process with external reviewers as a means of effecting and ratcheting up quality improvements" (2003: p. 4).

The participants identified Accreditation as a means to introduce a National agenda to improve Quality, but a potential threat to this is that "the DoH is not perceived as a custodian for the well being of its citizens and their access to high quality health services, but only as the agent to provide health services of questionable quality in government facilities". This, together with the government's "VIP culture" and "safarish" way of decision making, were seen as barriers to introducing Accreditation. The problem of the VIP culture has also been identified as a hindrance to progress in Pakistan by Prime Minister Shaukat Aziz. He stated that, "We have to come out of this VIP culture and work collectively to ensure that justice reaches the common people" (Rabi-us-Sani, 2006). It appears that this question of 'safarish' is evident in selection for positions at all levels in the society, as evidenced by the content of many blogs and other communications in the internet.

The participants viewed Accreditation as a means to provide direction on Quality and to obtain evidence on what works. In many developing countries — alas not yet in Pakistan — it is common to rely on standards and other quality mechanisms to provide information and direction (Shaw, C., 2003: p. 15.) But the participants' desire for direction was offset by the fear of change and their expectation of general resistance to change; to counter this they identified the need for training in change management.

They emphasized throughout the SWOT analysis the need for training and capacity building in management and quality concepts. They identified that Accreditation could provide the support, direction, and motivation in training and decentralization. One of the prerequisites for Accreditation is able health facility managers with powers to make resource and process decisions (Shaw as above). The government of NWFP passed in 2002, the NWFP Local Government Ordinance, 2001, which is the enabling legislation for decentralising the management of district services to district level, including "administrative and financial authority for the operation, functioning and management of specified offices to local government" [Chapter 1, paragraph 2 (vii)]. Recent reports and discussions in Peshawar have all indicated that this failure of implementation of the legislation has had a major negative impact on the efficiency, quality and overall performance of health facilities (Potter, C., 2006; Ismail, Z., 2006; Yunis, S. & Minett, C., 2006). The Institute for Healthcare Improvement in the USA, a world leader in research and learning activities for quality, states that, "Patient safety requires commitment from all levels of an organization. While executives foster a safety culture and establish clear goals and metrics for the organization, the day-to-day execution is the responsibility of frontline managers who direct resources at the "sharp end" of care. [...] In order to ensure that quality and safety are a top priority, these managers need the tools and understanding of patient safety to lead the work at their institutions" (Institute for Healthcare Improvement, 2007).

That accreditation would assist in clarifying goals, boundaries and measures was identified, though the lack of precision in measurement in methods such as accreditation was recognized. One of the prerequisites for Accreditation is that all health professionals have defined accountability mechanisms; Charles Shaw states that Accreditation "requires a culture of transparency and acceptance of personal and corporate responsibility among management and clinical staff" (2003: p. 30). But again, Potter, in his review of hospitals in NWFP, outlined the lack of transparency and inability of those responsible for management decisions to undertake this role and implement management methods. He and others have found that the roles of managers in health facilities in NWFP are unclear and they are not provided with decision-making powers over the financial and human resources for which they are held responsible (Potter, 2006; Yunis & Minett, 2006).

Conclusion

Though NWFP faces significant challenges in introducing a Quality Management system they have made some progress and have taken a first step in identifying the challenges they must overcome. Outdated administrative systems within the

government, lack of management and quality training and capacity building, and unclear accountability mechanisms from the individual up to the government level places significant barriers for progress in improving quality.

Context specific multi-faceted strategies are needed to improve the quality of healthcare services and these are best identified by those with experiential knowledge of the local context. But in a busy health sector it is often difficult to get key stakeholders together and even more difficult to gain access to their knowledge when management skills are weak and research methods are complex. The SWOT technique can provide a simple, useful research tool to gain insight into important variables during the formative stage of research. The results of this SWOT analysis can be useful as a basis for research into Accreditation.

References

Armstrong, M., 2006. A handbook of human resource management practice. 10th ed. London: Kogan Page.

Catsambas, T., 2002. The evaluation of quality assurance: developing and testing practical methods for managers. *International Journal for Quality in Health Care*, 14 (Suppl. 1), p. 75-81.

Doyle, V. & Haran D., 2006. Quality assurance in health care. Available at: http:// www.liv.ac.uj/lstm/hsrqapb.html [accessed 20 June 2006].

Government of Pakistan, 2002. *Economic Survey of Pakistan 2001-2002*. Islamabad: Economic Advisor's Wing. Finance Division.

Government of Pakistan. Federal Bureau of Statistics, 2002. Pakistan Integrated Household Survey. Islamabad: Federal Bureau of Statistics.

Grace-Bishop, K., 2004. Social marketing "in action". Available at: http:// www.nhhealthpolicyinstitute.unh.edu/ec/ppt/04_social-marketing.ppt#1 [accessed 20 September 2007].

Institute for Healthcare Improvement, 2007. Leading and implementing safety strategies: role of managers. Available at: http://www.ihi.org/IHI/Programs/ConferencesAndSeminars/ LeadingandImplementingSafetyStrategiesRoleofManagerOctober2007.htm [accessed 10 July 2007].

Massoud, R. Askov, J. & Reinke, L., 2001. *A modern paradigm for improving healthcare quality*. Bethesda, Maryland: QA Project. (QA Monograph Series; 1/1).

Mays, N. & Pope, C., 1995. Qualitative research: rigour and qualitative research. *BMJ*, 311, p. 109-112.

Montagu, D., 2003. *Accreditation and other external quality assessment systems for healthcare*. London: DFID Health Systems Resource Centre.

Nandraj, S., et al., 2001. A stakeholder approach towards hospital accreditation in India. *Health Policy and Planning,* 16 (Suppl. 2), p. 70-79.

Nishtar, S., 2006. *The Gateway Paper: health systems in Pakistan: a way forward.* Islamabad, Pakistan: Pakistan's Health Policy Forum and Heartfile.

Øvretveit, J., 2004. A framework for quality improvement translation: understanding the conditionality of interventions. *Joint Commission Journal on Quality and Safety,* Global Supplement, p. 15-24.

Øvretveit, J., 1997. Would it work for us? Learning from quality improvement in Europe and beyond. *Joint Commission Journal on Quality Improvement,* 23, p. 7-22.

Øvretveit, J. & Gustafson, D., 2003. Improving the quality of health care: using research to inform quality programmes. *British Medical Journal,* 326, p. 759-761.

Øvretveit, J., 2002. Improving the quality of health services in developing countries: lessons for the West. *Quality & Safety in Health Care,* 11, p. 301-302.

Peters, D., Rao, K. S. & Fryatt, R., 2003. Lumping and splitting: the health policy agenda in India. *Health Policy & Planning,* 18 (3), p. 249-260.

Potter, C., 2006. *Mission to evaluate progress on hospital autonomy in NWFP.* Bad Homburg: EPOS Health Consultants.

Rabi-us-Sani, 2006. VIP, safarish culture must end: Shaukat. *Pakistan Observer,* Tuesday, May 09, 2006. Available at: http://archive.pakobserver.net/200605/09/news/topstories11.asp?txt=VIP,%20safarish%20culture%20must%20end:%20Shaukat [accessed July 14, 2007].

Roberts, M., et al., 2004. *Getting health reform right.* Oxford: Oxford University Press.

Shaw, C., 2004. *Toolkit for accreditation programmes.* Melbourne: ISQUA.

Shaw, C., 2003. *Quality and accreditation in health care services: a global review.* Geneva: WHO.

Smits, H., Leatherman, S. & Berwick, D., 2002. Quality improvement in the developing world. *International Journal for Quality in Health Care,* 14 (6), p. 439-440.

Transparency International, 2007. Corruption Perception's Index 2006. Available at: http://www.icgg.org/downloads/CPI_2006.xls [accessed July 14, 2007].

Travis, P., et. al., 2004. Overcoming health-systems constraints to achieve the Millennium Development Goals. *Lancet,* 364, p. 900-906.

World Bank, 2006. World Development Report. New York: World Bank.

Yunis, S. & Minett, C., 2006. Developing a decentralised human resource policy framework for the health sector. Peshawar, Islamic Republic of Pakistan: NWFP (North West Frontier Province). GTZ.

Zafar, H. Ismail, 2006. Reforms in the health sector: Final Report Version 2. Peshawar: United Nations Children's Fund.

Occupational Health and Ergonomics Toward Patient Safety

António Sousa Uva*,**
Florentino Serranheira*,**
Paulo Sousa*
Ema Leite*
João Prista*,**

Introduction

The major challenges facing health care systems include (i) the demographic change and the aging population; (ii) the increasing complexity of health care and technological developments; (iii) the high patient expectations and the growing pressure for accountability (iv) the greater than ever costs (Berwick et al., 2003; Borger et al., 2006). All these challenges have influenced the quality and sustainability of health care services.

In the last years, we've witnessed remarkable advances in the pharmacological field, massive technical developments and new therapeutic strategies, which have allowed treating patient with evermore complex and difficult medical conditions.

Despite improvements in healthcare intervention, the incidence of adverse events constitutes a major contributor to the global burden of disease, and a concern for patient safety. As a matter of fact improving the safety of patient care is now a core issue to health care systems in both developed and developing countries.

Patient safety is a key component of quality in health care and it is considered for several authors as the first and the most essential one (Stevens, 2005; Burroughs et al., 2007). New knowledge leading to improve patient safety intimately contributes to develop the quality of health care.

Recently, some studies, mostly of them in developed countries, have shown that the rate of injured hospital patients ranges among 4% to 16% (Thomas & Brennan, 2000; NHS, 2004). When adverse events happen, the impact can be catastrophic, for patients, their families and for health care system itself, by damaging its credibility and using valuable resources.

Patient safety improvements require better information sharing about the number, types, causes and consequences of errors and adverse events (Den Heed et al., 2006).

* National School of Public Health, New University of Lisbon, 1500-560, Lisbon.
** CIESP – Centro de Investigação e Estudos em Saúde Pública (asuva@ensp.unl.pt).

At the same time, safety in action will spread best practices for implementation of changes in organizational team and also at clinical practices to improve patient safety (WHO, 2006a).

Evidence-from research demonstrates that adverse events are not just a series of random, unconnected one-off incidents (Altman et al., 2004). Such events often demonstrate common root causes provoked by weak systems. There appears to be much that health care can learn from the systematic and sustained attention to improving safety demonstrated by other high-risk industries, like aviation, for instance. Leape (Leape, 1994) compares hospital care to aviation industry, justifying that this industry seems to be safer once only occur four crashes a year in more than 10 million takeoffs and landings each year. Helmreich refers that in aviation has been reported failures of compliance, communication, procedures, proficiency and decision making contributing to errors that may be similar in clinical care (Helmreich, 2000). There is a lack of knowledge linked with error management that is related with the actual culture of professionals, hospitals, patients and other stakeholders. It's urgent to increase efforts to understand how the most current errors happen to act on its prevention. This must be based in more scientific approaches of the fault tree causes.

Improving the safety of patient care requires system-wide action on a broad range of fronts to identify and manage actual and potential risks to patient safety, and implement long-term solutions. This requires actions in performance improvement, environmental safety, and risk assessment and management, including infection control and occupational health and safety, safe use of medicines, safety equipment, safe clinical practice, and safer and healthier environment of care.

Nowadays, there is a large amount of evidence about the consequences of patient safety, or its absence, in health organizations, in patients and their families, and also, in the health care professionals.

One of the most frustrating aspects of patient safety is the apparent inability of healthcare systems to learn from their mistakes (WHO, 2006). Safety cannot be improved without a variety of valid reporting, analytical and investigative tools that identify sources and causes of risk in ways that promote learning and preventative actions. Reporting is vital, but it can never give a complete picture of all sources and causes of risk if we don't make an accurate fault tree analysis. A multi approach is therefore needed incorporating strategies such as audit of medical records, pro-active risk assessment, and observational tools for error identification and quality improvement based on an understating of complexity of system and individuals, namely the health professionals (Lilford, 2002; Donaldson, 2004).

Improving the organization of healthcare can reduce errors and raise safety, but research is needed to show how we can improve the system and help health

professionals to perform the highest standards, that can contribute to a correct risk management (Farquhar et al., 2002; Lewis & Fletcher, 2005). Research is fundamental for developing solutions using scientific and evidence-based approaches and to evaluate the effectiveness of interventions. It is also important for understanding the extent and causes of patient harm and developing appropriate strategies.

It is important to be aware that research for patient safety is not only about increasing knowledge; it is also about translating knowledge into practice. It's also important the bridging between the levels of research, dissemination, and adoption at policy, practice, managerial and consumer.

For instances, improving patient safety is not confined to error related with patient care based on mistakes of health care professionals but requires carefully designed systems of care, based on accurate ergonomic approaches that can determine safer and healthy work which also reduce risks to patients. Complementary actions are needed to prevent adverse events, make them quickly visible when they do occur, mitigate their effects on patients and health care workers and reduce risks to future patients (Meyers & Eisenberg, 2002).

Change is needed at the level of individual health care workers, teams, organization and whole healthcare system that abandon the most prevalent model based on blame, to persuade proper performance (perfectibility model). If health care system equipments needs is needed more patient safety, so the system and health professionals must learn with errors and make the prevention not based on guilt or blame. For instances we can design a system that must be redundant, duplicating "filters" of critical items to reduce failure probability.

Safe care is not an option. It is the right of patients to entrust their care to our healthcare system. That system incorporates competent, conscientious and safety-conscious health workers in frontline services (WHO, 2006) and must develop a culture where errors are not seen as human failures independent of the system processes, but opportunities to quality management developments.

There is a growing body of research on patient safety, but its extent and significance is still limited. The research about the consequences of working conditions and activity on patient safety, for instance, is in its beginning worldwide. The understanding of how the environment of care impacts the ability of providers to improve safety and how interactions with the physical healthcare environment (e.g., facility design, aesthetics, ergonomics conditions, etc.) influence the provision of safe high quality care, have been underestimate by researchers and policy makers (WHO, 2006).

Occupational Health & Safety and Ergonomics studies have demonstrated that the great majority of causes related with error are, most of the times, beyond the control

of each individual. If we want to undertake the prevention approach of that, it will be necessary to understand error related circumstances and factors. For instance:

- How can shift work, timetables and work schedules result in workload?
- What is the relationship between schedules and fatigue?
- Organizational and time pressure may induce errors? How can we manage this kind of demands?
- How can task descriptions, team composition or other work organization issues influence error prevention?
- How can we prevent unsafe environment acting on time pressure and fatigue or acting on working arrangements that minimize those psychological precursors?

The old saying 'to err is human' is certainly true but does not tell the whole story. Another defining characteristic of humans is their ability to work reliably in chaotic circumstances. We need research to understand in what situations people are most prone to errors or rule violations. In particular, how do contextual factors, such as a culture of health and safety environment, the physical milieu, or regulatory mechanisms, that influence people's behaviour. We need to understand better the feasibility of optimal patient care, given the time pressures and other resource constraints that are prevalent in hospitals. And more research is needed about teamwork, particularly multi-disciplinary teamwork and in the development of systems helping decision that is based on fault tree process. This line of research should be drawing on experimental psychology, anthropology/sociology and ergonomics; in addition to more conventional epidemiologic studies namely on Occupational Health and Safety (Pittet & Donaldson, 2006).

The improvement of patient safety is related with the prevention of undesirable events for the patients and the interest on the subject is greater in the last 10 to 20 years.

For instance, at any given time 1,400,000 people all over the world, particularly in the developing countries, are estimated to suffer from a nosocomial or healthcare-associated infection (WHO, 2007).

Mercy Health Systems' leaders identified six impediments to patient safety culture: production demands and time pressures; absent or inadequate processes; failure to focus on process problems; poor teamwork; inadequate communication; fear and pride. They have identified five key elements that should enhance patient safety: improved leadership; reporting systems; measurement; best practices; and a supporting structure (Ballard, 2006).

Risk and uncertainty are very common in healthcare providing and we must undertake efforts to prevent them establishing risk control systems based essentially

on five capital aspects: (1) good reporting; (2) good practices; (3) ergonomics for health care workers; (4) good planning and design of health care facilities; and (5) safer and healthy workers.

1. Good reporting

Today almost half of all preventable adverse events may be related with medication errors (European Commission, 2005). The investment on errors visibility, related or not with prescription is, for certain, the most important step to prevention. With that recognition we may do risk management, mitigate adverse effects for patients and, predominantly reduce risks for future patients. The investigation on development of better risk assessment tools that could make a comprehensive identification of hazards, events and other causes is necessary to learn from the more common mistakes. A good reporting system of errors and other incidents are a priority to the prevention of those errors and there is not enough knowledge about design of such a kind of subject that may contribute to the most appropriate prevention tools (NHS, 2004; Sousa, 2006).

Yet, simply a good detecting system is not enough to improve patient safety but permits responses that can lead to their reduction and prevention. Incorporation of methods such as audit and incident investigation can be more accurate on problems related with diagnosis. Analysis of errors and incidents to identify root causes must be part of a strategy toward patient safety improvements.

So, we must assume that the risk control (or risk management) must be based on a correct risk assessment (Uva, 2006) of the work conditions and practice in a trusting working environment with a culture based on learning from events as opposed to concentrate our efforts on "blame/punishment".

Sundström-Frisk (1993) interviewed intensive care staff about near-accidents associated with errors related with new technical equipment and concluded that creates new patient safety risks. Health professionals report psychological strain related with the fear to provoke an injury to the patient. Few of the near-accidents have been reported and as background factors they identify various such as: (1) deficiencies related with equipment and user guides; (2) stress; or (3) physical layout. Later there has been a complementary study (Sundström-Frisk & Hellstrom, 1995) aiming at obtain answers concerned with:

- What type of human errors has been involved?
- What external and internal stressors interplay in provoking human errors? or
- What are the reactions after a mistake?

They concluded that the risk of making a mistake is a psychological strain for staff and that patient safety is a work environment issue once inadequate introduction and implementation of techniques are, at least, as significant as human reliability.

The culture of errors acceptation may substitute the infallibility culture (Leape, 1991). That kind of approach can promote a quality culture and a continuous excellence search that is decisive on fault prevention. Most mistakes perceived by healthcare workers never become visible, because they are afraid of legal sanctions and of being ostracized by their peers. Feelings of shame are an important cause of underreporting errors too (Sunström-Frisk, 1999).

The other face of the coin is related with the fact that any expenditure on some aspects of patient safety, for instances prevention of health care related infections, is often seen more a cost than an investment. Some studies on nosocomial infection control programs reveal that the cost on infection control teams are 7% of the infection costs (Haley et al., 1985) which brings forth the idea that such intervention is really an investment.

What is important, thus, is reporting, analyzing and studying the different factors that contribute to make mistakes, instead of accusing the health care professional. Identifying their causes allows us to act, for example with technology and aid systems that prevent memory failures and coping of stressed healthcare professionals.

Implementation of interventions designed to improve a hospital's culture of patient safety can, if led by senior hospital executive, show the way to a substantial, profound, and lasting increase in error reporting and improvement in employee perceptions of the organization's safety culture (Cohen et al., 2004).

Feelings of guilt and worry about making mistakes could be responsible for work-related stress and illness in healthcare workers, so mistakes, stress and illness could be causes and/or consequences in their relationships. Error reporting must be anonymous to promote notification once the target is its promotion and not a guilt approach of events. Really, what we want is the route causes analysis that can be returned to health professionals. Obviously we can make a network of health care organizations that could identify local or national priorities on patient safety planning.

2. Good practices

In the past decades there has been an increase on the number of claims against health care professional related with potential malpractice. Nevertheless it is not always easy to distinguish unexpected event, related with the patient disease, from an

adverse event related with negligence caused by medical (or clinical) management. Troyen et al., cited in Brennan et al., define negligence as *"... care that fell bellow the standard expected of physicians in their community..."* (Brennan et al., 1991). They have reviewed 30,121 records from 51 hospitals in New York State in 1984, reporting that adverse events occurred in 3.7% of the hospitalizations, and 27.6% of these were caused by negligence affecting mostly the elderly, concluding that most are result of substandard care.

The introduction of risk management routines, in health care, for example, by developing guidelines and indicators as a part of a quality assessment system in the health care sector is decisive to implement such a practice. For instance, the National Patient Safety Agency "Clean your hands" campaign, developed by World Alliance for Patient Safety and applied by the United Kingdom's National Patient Safety Agency, to increase hand hygiene and reduce the nosocomial infections aimed study the impact of such a measure based on the installation of alcohol-based handrubs at points of patient care (World Alliance for Patient Safety, 2006). This is a risk management action that could even be cost effective.

Better practices are beyond standards and guidelines. Patient safety is assured not only by means of building sophisticated or complex organizational models but also by introducing simple solutions, clear modes of reasoning, being aware that we are dealing with a matter in which reasoning as Byzantines' surely doesn't facilitate an already complex apparatus (Tartaglia et al., 2006).

The study of the 1133 cases related with adverse events involving medical interventions showed that the most of them were *"delayed treatment"* and *"failure to use indicated tests"* associated with memory and attention failures (Leape, 1994). These failures were more related with physical and mental stressors than with professional skills. Better practice is almost always related with better guidelines and not with the application of guidelines (Leape, 1994). It will be interesting to investigate a little bit more that field trying to understand decision making process.

3. Ergonomics for health care workers

According to the International Ergonomics Association, organizational ergonomics is concerned with the organization of work systems, including their governmental structures, policies and processes (Karwowski, 2006). Although organizational ergonomics concerns the design of work systems to fit human nature (Hendrick & Kleiner, 2002), it can gives information about the study of patient safety interventions. In addition to incorporate open systems theory, organizational

ergonomics theory supports the development of work systems that complement human behaviour and account for human error (Kohn et al., 1999; Hendrick & Kleiner, 2002).

Organizational ergonomics proposes that work systems organizations design must consider the interaction of four major socio-technical system elements: (i) the external environment; (ii) the technical subsystem; (iii) the internal environment; and (iv) the personnel subsystem (Hendrick & Kleiner, 2002).

In the context of patient safety, the study of work organization is the health care delivery system of a particular country (Schutz et al., 2007). By applying organizational ergonomics to analyse patient safety improvement interventions, the external environment can be conceived as health care policies and associated health care organizations (those organizations that influence patient safety indirectly rather than through the provision of health care services). Likewise, the technical subsystem is described in the context of patient safety interventions. The internal environment also encompasses organization-level patient safety improvement initiatives, but describes the organizational and management structures rather than how work is performed and who and how it is done. Finally, the personnel subsystem related to patient safety interventions comprises initiatives that affect clinical microsystems (the individual or clinical teams of practitioners who deliver patient care) (Schutz et al., 2007).

This approach has developed over the past 50 years in order to address the complex interactions that occur between a worker, their tools, their colleagues, and their work organisation. More recently a need to look still further and consider the role of regulations, societal and cultural pressures has been recognised (Moray, 2000). For the health care sector, this appears to be a daunting, but necessary, challenge. This sector has specific needs given its complexity, scale, and potential impact on its very diverse user groups, particularly patients. However, similar complex challenges are being met by a number of other safety-critical industries, including both nuclear and aviation. These industries have adopted an ergonomic system's approach precisely because they have realised the dangers of considering only one element of a system, in isolation from others (Department of Health, 2003).

Open systems theory is also applicable to patient safety improvement. According to systems theory, a system exists within an external environment and draws inputs from and supplies outputs to this environment. The limits of the system determine the inputs, transformations and outputs (Hallock et al., 2006). Therefore, from a system's approach, the components of a health care system that impact patient safety, including workers, technology and the environment, are not considered isolated. Moreover, maintaining patient safety is understood as the ultimate goal throughout all components of a health care system and work processes are designed ultimately to

attain the goal of safety (O'Neill, 1998). Nevertheless there is insufficient research in this systemic approach to patient safety goal.

Accordingly to Tessler (2003), the components of an effective ergonomic program include:

- Job analysis for identification of hazards, and risk evaluation to classify high risk tasks and hazardous departments (risk assessment);
- Identification of solutions, including use of appropriate equipment and possible changes in how the work is organized. Some facilities implement "Zero Lift Policies" or utilize "Lifting Teams" (risk control and risk management);
- Accurate and complete methods of collecting incident and injury data related to patient care activities;
- Training and education programs provided to all health team members; education of patients and/or family members regarding also their safety;
- Evaluate how well the solutions are working;
- Implementation of a pro-active and blameless medical management system and injury reporting environment; employees should feel encouraged by their supervisors to report the first signs and symptoms of injury;
- Return-to-work programs should provide support, appropriate referrals, and safe job tasks for injured workers.

The organizational ergonomics theory of socio-technical characteristics of organizational structures go together with the classic quality management principles. The application of organizational ergonomics to patient safety interventions provides an understanding for clarifying the current state of patient safety improvement within health care systems. Moreover, the structural framework makes clear which interventions are appropriately undertaken by policymakers, organizational leaders and work-unit managers. Although human factors research has contributed extensively to understanding the causes of medical error, organizational ergonomics is a field that still has much to offer to patient safety research, particularly in the design of effective and safe health care systems (Schutz et al., 2007).

Ergonomics represents the bridge between quality and safety. It was a methodological approach for establishing connection among different disciplines, useful for incrementing synergy and to provide added value to the continuous quality improvement process (Tartaglia et al., 2006).

A good dose of pragmatism and modesty is required in order to handle patient safety within hospitals if we want to understand its causes and prevent error and other issues related with outcomes to the patients. The systemic approach of ergonomics may come up to a more harmonious environment.

4. Good planning and design of health care facilities

Often a patient's first impression of a hospital is as an "ugly" place, impersonal for recovery and maybe not safe. The importance of environment and functionality in a hospital should not be underestimated. Poor ergonomic design and bad planning, often spoiled by hateful extensions, 'temporary buildings' that look more like factories than hospitals, are frequently unsafe for patients and health professionals.

If the buildings are not attractive and sometimes contribute to healthcare-associated infection, it is imperative that architects, ergonomists, designers and builders be partners with healthcare staff and infection control teams when planning new facilities or renovating older buildings (NHS, 2002).

There is an actual requirement to use ergonomic design principles and create adaptable facilities, in order to meet the pace of clinical and technological development, not only in patient diagnosis and treatment, but also in many other aspects of care and organisation (NHS, 2001) such as hospital planning and design. The quality of a building's layout, safety, security actions and facilities is a significant factor in preventing risks, occupational diseases, absenteeism and sickness (NHS Estates, 1994). Satisfaction in a well-designed building, with good aesthetics and functionality, safety workplaces and staff communication do wonders for worker safety and also for patient safety. Aesthetics and economics should be indivisible, not mutually exclusive.

Aesthetic design and outstanding ergonomics are essential for functionality and patient safety. Like in the industry all hospital equipment should be structured for ease use and practicality for the user. Medical apparatus must integrate fully into the product's appearance — buttons or switches, packings, tablet bottles, labels, for example, must have distinctive colours, forms, contrast, allowing to a correct identification. It is important to counteract these developments by ensuring that a product's safety and innovation are readily recognisable by customers, avoiding errors.

Another example is the increasing demand for diagnostic imaging and interventional radiology treatment procedures and it is anticipated that this trend will continue in the future. Technological advances and innovations in imaging may produce changes in radiological methods, and it is likely that these will significantly affect the patient safety (NHS, 2001). Diagnostic imaging services have a greater role in the total management of the patient, involving consultation, diagnostic procedures, discussion and treatment. Therefore workers at a high tech environment are in risk of error because work conditions assume an important role on high work and mental load (NHS, 2001).

Health care organizations, such as hospitals, must have a feeling of harmony. The suitability of materials and layouts, can contribute to a sense of safety. The use of natural light and ventilation gives a link with life outside, a degree of individual environmental control and is economically preferable to remote controls. Outside views can appear to extend space and provide a secure connection to the world. Patients should not be made to feel cut off from the activity going on around them both inside and outside the building. This sometimes feels like (and are) unsafe places (NHS Estates, 1994).

"Wayfinding" is a relatively new term in Hospitals which covers everything to do with how people (patients or professionals) find their way round environments. All sites will have to allow access to all areas, removing physical barriers or providing reasonable alternative access for people with disabilities. Hospitals also have to provide accessible "wayfinding" information and aids such as large print written directions, good lighting at information desks to enable people who are deaf or hard of hearing to lip-read, and both audio and visual safety information in case of an emergency (Miller & Lewis, 1999). "Wayfinding" is a way to promote patient safety and safer workplaces, without disruptions/invasions caused by "lost" people and, at the same time, keeping all well informed.

Hospital planning and design must bear in mind the reduction of the healthcare-associated infections, not only for patients but also for professionals. It is imperative that planning and work teams be partners with healthcare staff and infection control staff when planning new facilities or renovating older buildings.

As a matter of fact, research re-enforced that the healthcare environment is a reservoir for organisms with the potential for infecting people. If healthcare-associated infection (public, patients and/or professionals) is to be reduced, it is imperative that infection control is carefully designed at the planning and construction stages of a healthcare-facility new building, or renovation project (NHS, 2002). More research is needed for planning new buildings and safer departments for people and healthcare professionals, keeping in mind the contribution of the ergonomic approach.

5. Safer and healthy workers

Human resources are fundamental to treat patients. Obviously, the hospital needs healthy workforce, physically and psychologically. Their work conditions and also the safety environment of the workplace are determinant for their wellbeing. Healthcare workers may be weakened by emotional and physical distress and when demands exceed the workers' capacities the risk of making mistakes increases.

Human mistakes have been associated, most of the times, with individual characteristics of the worker instead of being related with a constellation of causes that determine human behavior. The understanding of the interaction between organizational, technical and human factors is essential to manage the risk of making mistakes (Sundström-Frisk, 1999).

The importance of understanding the causes of errors and the need to undertake a 'systematic analysis of incidents' in the health care sector has been widely referred (Leape et al., 1995; Department of Health, 2001; Audit Commission, 2001). When an adverse incident occurs probably there are elements from physical, technological, psychosocial and cultural environments (Nieva & Sorra, 2003; Kho et al., 2005) that are contributing to the event (Health and Safety Executive, 1999) also those related with occupational health and safety. Inevitably, these factors present a significant challenge for accidents prevention that remains an area requiring urgent research to evaluate the potential benefits for patient safety.

Shift and night work can interfere with work performance and with possible errors and/or accidents (Costa, 1996). People who work rotating shifts are more likely to suffer from sleep disturbances and reduced alertness and performance. One study involving 635 nurses found that the odds of making, or almost making, a medication error, as well as the odds of having an accident or a near miss while commuting, doubled among rotating shift workers (Hughes, 2004). This kind of approach, on potential error causes, is not very common, and should be studied more deeply in the future. Moreover, safer and healthier shift work conditions have a good relationship with better working time arrangements (Kogi, 2004).

In hospitals with high patient-to-nurse ratios, nurses are more likely to experience burnout and job dissatisfaction and surgical patients experience higher risk-adjusted 30-days mortality and failure-to-rescue rates (Aiken et al., 2002).

Factors like shift work, working for a long time, time pressure and bad physical and ergonomics conditions contribute for distress, work accidents and increase the probability of errors. For example, a nurse who was observed in the Occupational Health Department showed psychosomatic and psychological symptoms associated with feelings of guilt, after had administered insulin to a wrong patient. She had worked night shift and the next morning shift, so she was tired and she didn't well understand the instruction of a colleague. In that hospital, the medication's distribution system was not based on unidose system for each patient, so the risk to give the wrong drug to the wrong patient was higher.

On the other hand any occupational injury resulting from an unsafe workplace affects negatively a health care organization by increasing costs and reducing the facility's ability to provide services. A back injury resulting from work after a heavy lifting (patient or equipment) sustained by a health care employee may not appear to

directly affect the wellness of his or her patients. It may result in the employee not being available for work or working at a difficult condition. The employee's manager (medical, nurse or technician) may have to temporarily fill the place with supplemental staff that may not be knowledgeable in departmental policy and procedures or made a different work balance between all staff. This could lead to patient injuries.

Health programs for reducing/preventing professional injuries need hazard identification for a correct risk assessment and management (Uva, 2006). Back injuries frequently associated with lifting's of heavy equipment or heavy patients and Work-Related Musculoskeletal Disorders resulting from repetitive motions, extreme postures, or frequent force exertion need a well structured ergonomic program for effective prevention. For example at a workshop focussing on ergonomics interventions at back and shoulder stress, researchers found a reduction in the prevalence of back problems, lost work days and quality of patient care when assistive devices were introduced as an aid to reduce physical demands of lifting and transferring patients (Owen, 1995). Physical strain at health care workers is mostly related with bending over (10% of their work-day) during patient care (Malchaire, 1992). Research denotes that 16% to 24% of nurses working hours are spent in uncomfortable positions (Estryn-Béhar, 1995).

Another place in hospitals that assume a very import role on patient safety is the intensive care units (ICUs). Their work is stressful in temporal adverse conditions that lead frequently to rapid and skilled decisions. In USA there are about 6,000 of them that care a lot of patients (Angus et al., 2000). Some authors (Donchin, 1995) estimate that there are 1.7 errors per patient per day in ICU's. Patient safety is also decisive at ICU practice because it is filled with high workload situations (Crickmore, 1987; Malacrida et al., 1991; Oates & Oates, 1996). Nurses must continuously respond to the needs of patients and families, and routinely interact with the most intense emotional aspects of life. Research shows that nursing workload is one of the most important determinants of patient safety and quality of care in ICUs.

On other hand, insufficient nursing staff was found to be associated with the occurrence of the following incidents: drug administration or documentation problems, inadequate patient supervision, incorrect ventilator or equipment setup, and self-extubation. Undesirable patient outcomes associated with insufficient nursing staff include major physiological change, patient or relative dissatisfaction, and physical injury (Carayon & Gurses, 2005).

Accordingly Blegen et al. (Blegen et al., 2005) the most important recommendations for change are divided in three categories: (i) technology; (ii) working conditions or organization of work and (iii) culture or climate in health care organization. Safety culture is related not only with reporting mistakes, rules and

procedures or process auditing but also to aspects related with risk perception, fatigue and stress or attitudes and experiences that must be interpreted as part of health care.

Conclusions

Occupational Health Departments could, or should, work together with Quality Services to contribute for health care worker's wellbeing. The route cause analyses have an important role in the reduction and prevention of other near-accidents or even accidents in the future. Adopting safety strategies and promoting patient safety research concerned with Occupational Health and Safety, and Ergonomics could also be crucial to reduce the risk of adverse events inherent to health care delivering system.

Hospital worker's Health and wellbeing are important issues on promoting patient safety. Behavior is not only related with personality but also to each work situation (Zink, 2005) and, therefore successful improvement of patient safety may be related with organizational aspects and social climate that is also influenced by Occupational Health and Safety aspects.

Effects of working conditions on patient safety with the aim to develop an understanding of how the environment care impacts the ability of providers to improve safety (the effect of fatigue, stress, sleep deprivation, and shift work on cognitive ability and the relationship to patient safety) and how interactions with the physical healthcare environment impacts the provision of safe high quality care are needed at hospitals.

It is our belief that the knowledge of incidence and prevalence of adverse events and the study of the multifactor's underlying these occurrences, particularly those related to working conditions, individual characteristics and safety environment will, in the near future, be one of the main areas for action, reflection and research in the public health domain. We must focus our efforts on the whole work system (including the healthcare worker) and not only in the individuals itself in the comprehension of clinical errors once the organizational environmental and other work conditions are so important as the human behavior and it is not certain what is cause or consequence.

References

Aiken, L. H., et al., 2002. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA: the Journal of the American Medical Association,* 288 (16), p. 1987-1993.

Altman, D. E., Clancy, C. & Blendon, R. J., 2004. Improving patient safety: five years after the IOM report. *New England Journal of Medicine,* 351 (20), p. 2041-2043.

Angus, D. C., et al., 2000. Current and projected workforce requirements for the care of the critically ill patients and patients with pulmonary disease: can we meet the requirements of an aging population. *JAMA: the Journal of the American Medical Association,* 284 (21), p. 2762-2770.

Audit Commission, 2001. *A spoonful of sugar.* London: Audit Commission.

Ballard, L., 2006. Putting safety at the core. *Health Progress,* 87 (1), p. 29-34.

Berwick, D. M., James, B. & Coye, M. J., 2003. Connections between quality measurement and improvement. *Medical Care,* 41 (1), p. 130-138.

Blegen, M. A., Pepper, G. A. & Rosse, J., 2005. Safety climate on hospital units: a new measure. In: US Department of Health & Human Services. AHRQ Advances in Patient Safety: from research to implementation. 4. Rockville, MD: Agency for Healthcare Research and Quality. (AHRQ Publication nos. 050021 (1-4) p. 429-443.

Borger, C., et al., 2006. Health spending projections through 2015: changes on the horizon. *Health Affairs,* 25 (2), p. 61-73.

Brennan, T. A., et al., 1991. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. *The New England Journal of Medicine,* 324 (6), p. 370-376.

Burroughs, E. T., et al., 2007. Patients concerns about medical errors during hospitalization. *Joint Commission Journal on Quality and Safety,* 33 (1), p. 5-14.

Carayon, P. & Gurses, A. P., 2005. A human factors engineering conceptual framework of nursing workload and patient safety in intensive care units. *Intensive and Critical Care Nursing,* 21 (5), p. 284-301.

Catchpole, K. R., et al., 2006. Identification of systems failures in successful pediatric cardiac surgery. Ergonomics, 49 (5-6), p. 567-588.

Clarkson, P., et al., 2004. *Designing for patient safety: a scoping study to identify how the effective use of design could help to reduce medical accidents.* Cambridge, Surrey, UK: Engineering Design Centre. University of Cambridge. (ISBN 0-9545243-0-6).

Cohen, M. M., et al., 2004. Implementing a hospitalwide patient safety program for cultural change. *Joint Commission Journal on Quality and Safety,* 30 (8), p. 424-431.

Committee on Quality of Health Care in America. Institute of Medicine, et al., 1999. *To err is human: building a safer health system.* Washington, D.C.: National Academy Press.

Costa, G., 1996. The impact of shift and night work on health. *Applied Ergonomics,* 27 (1), p. 9-16.

Den Heed, K. V., et al., 2006. Adverse outcomes in Belgian acute hospitals: retrospective analysis of the national hospital discharge dataset. *International Journal of Quality in Health Care,* 18 (3), p. 211-219.

Department of Health. Design Council, 2003. Design for patient safety: a system-wide design-led approach to tackling patient safety in the NHS. London: Design Council. Department of Health. (ISBN 1-84182-7657).

Donaldson, L., 2004. Patient safety global momentum builds. *Quality & Safety in Health Care,* 13 (2), p. 86-91.

Donchin, Y., et al., 1995. A look into the nature and causes of human errors in the intensive care unit. *Critical Care Medicine,* 23 (2), p. 294-300.

Estryn-Béhar, M., 1995. *Working conditions and health hazards of health care workers.* In Hagberg, M., Hofmann, F., Stössel, U. & Westlander, G., ed., 1995. International Symposium on Occupational Health for Health Care Workers, 1995. Hospital ergonomics. Landsberg, Lech: Ecomed Verlag.

European Commission, 2005. *Luxembourg Declaration on Patient Safety* [internet] Luxembourg: European Commission. (published 2005) Available at: http://ec.europa.eu/health/ph_overview/Documents/ev_20050405_rd01_en.pdf [accessed 20 October 2007].

Farquhar, C. M., Stryer, D. & Slutsky, J., 2002. Translating research into practice: the future ahead. *International Journal of Quality in Health Care,* 14 (3), p. 233-249.

Grant, M. J., Donaldson, A. E. & Larsen, G. Y., 2006. The safety culture in a children's hospital. *Journal of Nursing Care Quality,* 21 (3), p. 223-229.

Groah, L. & Butler, L., 2006. Is there a relationship between workplace and patient safety? *Association of periOperative Registered Nurses Journal (AORN Journal),* 84 (4), p. 653-654.

Groene, O., ed., 2006. Implementing health promotion in hospital: manual and self-assessment forms. Geneva: WHO Regional Office for Europe.

Haley, R. W., et al., 1985. The efficacy of infection surveillance and control programs in preventing nosocomial infections in U.S. hospitals. *American Journal of Epidemiology,* 121 (2), p. 182-205.

Hallock, M. L., Alper, S. J. & Karsh, B., 2006. A macroergonomic work system analysis of the diagnostic testing process in an outpatient health care facility for process improvement and patient safety. *Ergonomics,* 49 (5-6), p. 544-566.

Helmreich, R. L., 2000. On error management: lessons from aviation. *British Medical Journal (BMJ),* 320 (7237), p. 781-785.

Hendrick, H. W. & Kleiner, B. M., 2002. Macroergonomics: theory, methods and applications. Mahwah, NJ: Lawerence Erlbaum Associates.

Hughes, R., 2004. The perils of shift work: evening shift, night shift, and rotating shifts: are they for you? *The American Journal of Nursing,* 104 (9), p. 60-63.

Karwowski, W., ed., 2006. International encyclopaedia of ergonomics and human factors. 2nd ed. Boca Raton, FL: Taylor and Francis. CRC Press.

Kho, M. E., Carbone, J. M., Lucas, J. & Cook, D. J., 2005. Safety climate survey: reliability of results from a multicenter ICU survey. *Quality & Safety in Health Care,* 14 (4), p. 273-278.

Knauth, P., 1998. Innovative worktime arrangements. *Scandinavian Journal of Work, Environment & Health,* 24 (Suppl. 3), p. 13-17.

Kogi, K., 2004. Linking better shiftwork arrangements with safety and health managements systems. *Revista Saúde Pública,* 38 (Suppl.), p. 72-79.

Leape, L. L., 1994. *The preventability of medical injury*. In M.S. Bogner, ed. Human error in medicine. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1994, p. 13-25.

Leape, L. L., 1994. Error in medicine. *JAMA: the Journal of the American Medical Association,* 272 (23), p. 1851-1857.

Leape, L. L., et al., 1991. The nature of adverse events in hospitalized patients: results of the Harvard Medical Practice Study II. *The New England Journal of Medicine,* 324 (6), p. 377-384.

Leape, L. L., et al., 1995. Systems analysis of adverse drug events: ADE Prevention Study Group. *JAMA: the Journal of the American Medical Association,* 274 (1), p. 35-43.

Lewis, R.Q. & Fletcher, M., 2005. Implementing a national strategy for patient safety: lessons from the National Health Service in England. *Quality & Safety in Health Care,* 14 (2), p. 135-139.

Lilford, R. J., 2002. Patient safety research: does it have legs? *Quality & Safety in Health Care,* 11 (2), p. 113-118.

Malchaire, J., 1992. *Analysis of the workload of nurses*. In. Estryn-Béhar, M. & Gadbois, C. & Pottier, M., 1991. International Symposium on Occupational Health for Health Care Workers, Paris, 1991. Ergonomie à l'Hôpital. Toulouse: Octares Editions.

Meyers, G. S. & Eisenberg, D. R., 2002. The end of the beginning: the strategic approach to patient safety research. *Quality & Safety in Health Care,* 11 (1), p. 3-4.

Miller, C. & Lewis, D., 1999. *Wayfinding*. Norwich: NHS Estates. HMSO.

Moray, N., 2000. Culture, politics and ergonomics. *Ergonomics,* 43 (7), p. 858-868.

NHS Estates. HMSO, 1994. *Better by design: pursuit of excellence in healthcare buildings*. London: Her Majesty's Stationery Office.

NHS Estates. HMSO, 2001. *Facilities for diagnostic imaging and interventional radiology*. London: The Stationery Office (Health Building Note; 6).

NHS. National Patient Safety Agency. 2004. *Seven steps to patient safety: a guide for NHS staff*. London: National Patient Safety Agency.

NHS Estates. HMSO, 2002. *Infection control in the built environment: design and planning.* London: The Stationery Office.

Nieva, V. F. & Sorra, J., 2003. Safety culture assessment: a tool for improving patient safety in health care organizations. *Quality & Safety in Health Care,* 12 (Suppl. 2), p. ii17-ii23.

O'Neill, M., 1998. *Ergonomic design for organizational effectiveness.* Washington, D.C.: Lewis Publishers.

Pelikan, J. M., 1994. *Improving occupational health for health care workers within the health promoting hospital: experiences from the Vienna WHO model project "health and hospital".* In Hagberg, M., Hofmann, F., Stobel, U. & Westlander, G. ed., 1994. 2nd International Congress on Occupational Health for Health Care Workers, "ProHealthCare94", Stockholm, 1994, June 26-29. Stockholm: National Institute of Occupational Health. International Commission on Occupational Health (ICOH). 1.9. p. 55-62.

Pittet, D. & Donaldson, L., 2006. Challenging the world: patient safety and health care-associated infection. *International Journal for Quality in Health Care,* 18 (1), p. 33-39.

Schultz, A. L., Counte, M. A. & Meurer, S., 2007. Assessment of patient safety research from an organizational ergonomics and structural perspective. *Ergonomics,* 50 (9), p. 1451-1484.

Sousa, P., 2006. Patient safety: the need for a national strategy. *Acta Médica Portuguesa,* 19 (4), p. 309-317.

Stevens, D., 2005. Why new guidelines for reporting improvement research? And why not? *Quality & Safety in Health Care,* 14, p. 314-318.

Sundström-Frisk, C., 1993. Near-accidents in the use of advanced medical equipment. In G. Salvendy & M. Smith, ed., 1993. *Human-Computer Interaction.* (Advances in human factors/ergonomics). New York: Elsevier, 1993, p. 821-827.

Sundström-Frisk, C. & Hellstrom, M., 1994. *The risk of making treatment errors: an occupational stressor: a study of the link between technology, human error, patient safety and perceived psychological strain.* In Hagberg, M., Hofmann, F., Stobel, U. & Westlander, G. ed., 1994. 2nd International Congress on Occupational Health for Health Care Workers, "ProHealthCare94", Stockholm, 1994, June 26-29. Stockholm: National Institute of Occupational Health. International Commission on Occupational Health (ICOH). 1.21. p. 56-62.

Sundström-Frisk, C., 1999. The risk of making mistakes: a neglected occupational health problem. In Hasselhorn, H. M., Toomingas, A. & Lagerstrom, M., 1999. Occupational health for healthcare workers: a practical guide. Amsterdam: Elsevier, p. 168-171.

Tartaglia, R., et al., 2006. *Healthcare systems ergonomics and patient safety.* In Tartaglia, R., Bagnara, S., Bel-landi, T. & Albolino, S., ed., 2006. Human factors, a bridge between care and cure. Leiden: Taylor & Francis. p. 3-6.

Ternov, S. & Akselsson, R., 2005. System weaknesses as contributing causes of accidents in health care. *International Journal for Quality in Health Care,* 17 (1), p. 5-13.

Tessler, J., 2003. Patient handling solutions. Albany: Occupational Safety and Health Department of the New York State Public Employees Federation (PEF).

Uva, A., 2006. *Diagnóstico e gestão do risco em saúde ocupacional.* Lisboa: ISHST. (Segurança e Saúde no Trabalho. Estudos; 17).

Thomas, J. S. & Brennan, T. A., 2000. Incidence and types of preventable adverse events in elderly patients: population based review of medical records. *British Medical Journal (BMJ),* 320 (7237), p. 741-744.

Zink, K. J., 2005. From industrial safety to corporate health management. *Ergonomics,* 48 (5), p. 534-546.

WHO, 2006. *WHO guidelines on hand hygiene in health care: advanced draft.* Geneva: WHO.

WHO, 2007. *Patient safety.* [Online] (Updated 2007) Avaliable at: http:// www.who.int/patientsafety/en/ [accessed 10 November 2007]

World Alliance for Patient Safety, 2006. *Forward programme 2006-2007.* Geneva: WHO.

APPENDICES

APPENDIX 1
Format for Presenting Quality Research Methods and Issues

JOHN ØVRETVEIT*

NOTE: you may be at the stage of planning your research, or part way through doing it, or nearly finished or finished. Use the headings below according to the stage you are at and make clear what stage you are at when you write the paper.

Your chapter title.

Your name, title, institution, work address and email.

Structured abstract (max 400 words).

Phase of the research when this was written: *(e.g. between planning and doing a fuller review of the literature, before any data gathering).*

— Main research question
— General design and sample
— Data collection and analysis methods
— Main findings
— Methodological challenges
— Practical challenges
— Main lessons for other researchers
— Main paper
— Your chapter title

Heading 1: The practical problem and gap in knowledge this research addresses. Start with "This chapter was written as the research reached the stage of... *(e.g. beginning to collect the data, after planning and an initial review).*

Heading 2: The general research design and sample chosen and reasons for this choice e.g. *A case study research design was chosen which concentrated on services. The reasons were a balance between convenience and...*

* Director of Research at Medical Management Centre, Karolinska Institutet, Stockholm, Sweden (jovret@aol.com).

A systematic review method following (ref) of published research was chosen because.

Heading 3: Data collection and analysis methods chosen, and why:

Heading 4: Methodological challenges:

Refer to the issues you raised at the QIRN meeting and three questions you asked for help about.

Remember the "Challenges and issues I needed to work through as a researcher with other colleagues (based on your "I would like suggestions or advice about this" and "Three questions raised by my research"). Then work through these.

Which challenges did you face (and to what extent you honestly think you really overcame these) in deciding:

which previous research to review, the perspective you would take, the framework or model to guide you data gathering (did you use one?), how to narrow down onto a research question which could be answered, which was limited but would also be significant if answered (did you find one?), how to define research objectives and how these changed, which general design to use, which sample to use, exactly which data sources, which data gathering methods, which ways to analyse the data, how to present the findings and which ones to present, how to draw conclusions, whether to make practical recommendations, how to relate you findings to previous research. What were the challenges in deciding each of these? Also note other methodological and theoretical challenges you face or expect to face, not list above.

Heading 5: What helped me to resolve some of these methodological challenges [people, books, events (chance), experience of...]

Heading 6: Practical challenges: *(e.g. illness, people left, computer broke)*

Heading 7: What helped me resolve some of these practical challenges *(people, books, events, experience of...)*

Heading 8: Main findings to date (or what I expect)

This is a way of saying "how the story turned out" so far and "where you have reached in your journey"

Heading 9: Main lessons for other researchers:

What would most help them, and any inspiration you can give?

Do not do this because...

Pay particular attention to doing this...

Heading 10: References I found most helpful and inspiring (maximum 5)

Heading 11: Appendix listing of resources which would help other Q&SI researchers

APPENDIX 2
Headings for Explaining your Research and Getting Advice in Earlier Stages

JOHN ØVRETVEIT*

Presenter (20 mins)

1) Working title of my research.
2) Main question the research will answer.
3) These people could make better informed decisions as a result of an answer to this question.
4) Theories and perspectives the research will draw on and contribute to.
5) What excites me most about this research.
6) The research designs I am considering to answer this question *(see next page for a list — pick the designs which seem closes to what you are considering).*

   Design 1: (with strengths and weaknesses of the design).
   Design 2: (with strengths and weaknesses of the design).

Any other designs considered:

7) The design I prefer and why (or why I still can not decide).
8) Challenges I expect which may make it difficult to answer the question on time when I carry-out this research.
9) Ideas for minimising the possible difficulties to ensure the research is successfully completed.
10) What I most need advice, suggestions and help with.

Discussant comments on these items (5 mins)

1) What I think is exciting about this research and what is (or could be) a potentially important contribution;
2) Issues which journal reviewers and examiners may raise when assessing this research;
3) Areas I would suggest need to be developed in the research (and why).

---

* Director of Research at Medical Management Centre, Karolinska Institutet, Stockholm, Sweden (jovret@aol.com).

APPENDIX 3
Eleven Checklist Questions for Improving your Research Plan and Eight-step Research Framework

JOHN ØVRETVEIT*

Eleven checklist question for improving your reseach plan

1) Does the plan/study sufficiently consider *previous research* and draw on this to define the question or the hypothesis?
2) Does the plan/study give a clear and specific *question* which is answerable, or a testable hypothesis?
3) Is the question or hypothesis to be tested *significant* (see 8 below)?
4) Is the *design clearly described* and does the plan/study give a *justification* for using this design rather than another one?
5) Are the *data gathering and analysis methods* clearly described and does the plan/study give a *justification* for using these methods rather than other ones?
6) Are the *findings* clearly presented and explained?
7) Are *limitations* of the methods and of the study described?
8) Does/will the plan/study show how the findings *contribute* to:

   a) knowledge (by filling a gap and/or by reporting findings similar or different to published research)?
   b) practical action (by explaining how people would act differently as a result of knowing these findings)?

9) Does/will the plan/study give *conclusions which follow from* what has been presented, which do not introduce new material, which would not be misleading for an average reader, and link these to other research?
10) Does the plan/study show *clear links* between the question, the chosen design, the data gathering methods (i.e. clearly the best to answer the question), the findings and the conclusions, and links to other research?
11) Does the study give full *references,* especially of the key research in this field, and following reference conventions.

* Director of Research at Medical Management Centre, Karolinska Institutet, Stockholm, Sweden (jovret@aol.com).

Eight-step research framework

Reiterative and interactive not sequential: after working on one step, reconsider and revise other steps. E.g. reviewing research continues and always feeds into revising other steps.

1) *Find a fruitful research subject* Find a subject about which research is needed and where research can help (practical value, empirical knowledge, fills a gap). Make a first "working title" (and keep a list of other possible titles).
2) *Define the constraints* to the research: time target-date, people and time resources available, already-gathered data which could be used, finance.
3) *Reviewing:* finding and summarising the most important previous research on the subject and theories: what is already known? Consider how this links to your work.
4) *Defining the research question and theoretical framework:* any tentative hypotheses (the research objective is to answer this question)? Which theory or framework is suggested by the review of previous research which you will use to guide the data you will collect to answer the question? e.g. will you only look at the economic aspect, and which aspects?
5) *High-level planning:* deciding the main activities which need to be carried out and ordering these in time and level of detail.
6) *Deciding methods: perspective, design and data gathering:* Listing possible perspectives and designs. Choosing the design and data which is most likely to answer the question within the research constraints. Deciding the timing of data collection and which data gathering method and analysis will gather data to answer the question most effectively.
7) *Detailed planning and doing*

   Planning sample, *data gathering, methods and analysis,* report writing and presenting the research.
   Predicting *practical problems.*
   Timetabling (and responsibilities if a team is involved).
   Carrying out the research: Project management.

8) *Writing the report, thesis or publication:* headings, style and other writing requirements. Linking to other research and presenting the results: similarities and differences in the findings, filling gaps, written and verbal presentation.

APPENDIX 4
Future Priorities for Quality and Safety Research Subjects

JOHN ØVRETVEIT*

Over the last 6 years some of the fields of quality improvement research which QIRN presentations have covered include:

— the epidemiology of quality and safety problems (how widespread and serious is this problem, are there large variations?).
— quality costing and the economics of quality.
— development of measurement or indicators (e.g. safety culture surveys).
— evaluation of small scale interventions (e.g. team project).
— evaluation or case studies of large scale complex interventions (e.g. quality programmes in hospitals, health systems or regional or national programmes),
— qualitative research into provider or patient experiences.
— surveys of provider or patient perceptions of quality and safety.
— reviews of research or theories.
— cross case comparisons and syntheses.

Discussions and workshops between John Øvretveit and colleagues in the Nordic and Swedish Patient Safety Research Networks and the Australian Patient Safety Research Network identified the following subjects and issues as fruitful areas for future research:

Safety and inter-professional and inter-service coordination

— Safety problems because of poor integration or lack of coordination.

Local prevalence and variations

— Which treatments/diagnostic groups are the highest for safety problems, and what are the variations between units?

* Director of Research at Medical Management Centre, Karolinska Institutet, Stockholm, Sweden (jovret@aol.com).

Data and reporting

— Routine indicators — validity as indicators of safety performance? Special data bases?

— Personnel reports & increasing reporting. Reporting systems which take little time for practitioners to make a report (e.g. computer based systems).

— Anonymous no blame vs. identifyable accountable (firewall between reporting and disciplinary action).

— Legal discovery and anonimity.

Reporting and culture change

— Is it necessary to change culture in order to increase reporting, or could other changes increase reporting (e.g. easy reporting system, effective action and feedback)? (example of NSW safety experiment).

Patient injury claims

— Are patient injury claims rates a valid proxy for safety performance?

— What is the association between patient claims and other measures of safety and quality (e.g. mortality, other indicators)? Initial data — consistent pattern over time. Same departments similar rates each year. Logic: carefully assessed by clinicians and lay assessors, only awarded when certain injury caused by health care.

— Organisational factors associated with high and low claims.

— Identify 5 high and 5 low safety-performing units; Develop hypotheses from previous research about possible organisational influences; gather data to assess presence and absence of these factors in the units. Consider associations with personnel well being or attitudes.

— Comparison of different patient injury claims data (NZ, AUST, US, UK Sweden, Norway).

How data are used

Assessment systems

— Assessing risk of incident in service — assessing patients and assessing services.

Safety intervention design and evaluation

— Adapt locally and test. Build on guideline implementation research.

IT & automation implementation for safety

— (EMR) (forced function reactions). Evaluating IT systems requires access to control and intervention sites, ideally before change is instituted. This means learning which sites are considering a system and getting research established early.

Safety Theory in health care

— Causes, interventions, implementation theory — problem-intervention-condition combinations.

Leaders role in safety programmes and projects

— What do leaders have to do to lead improvement to safety (see Johns review of leadership research).

Patients role in safety improvement

— Involving patients in assessing safety and designing interventions.

Culture

— Describing what characterizes high and low safety performing organisations.
— Changing culture; does a safety culture follows, preceed or accompany changed behaviour (health promotion, behavoir change).
— Are any interventions effective for creating a greater safety culture?

Outcome > well being > organisation factors

— e.g. workload/staffing and safety.

Social consequences of adverse events

Inter-organisational/inter-professional safety problems and interventions (communications)

Safety in services not often studied

— PHC, psychiatry, nursing homes, early discharge and home support, rural health care.

Economics of safety

— How much are poor quality costing, how much do safety interventions cost, how much is saved?

Simulation and training

— Evaluation of simulation and training methods for improving safety (e.g. different training methods to improve responses to emergencies on a labour ward). There is very little comparative research on educational interventions or use of incentives to improve patient safety. Even outside health care, drill and simulations have not been evaluated in any rigorous way.

Professions & professionals

— Attitudes and responses to safety: fear, trust, identity.
— Use of patent privacy arguments to protect interests in EMR development.

— Staffing levels, qualification, skill-mix and safety. As communication between staff is one cause of safety problems, would merging professional roles be more effective than education on team working?

Other fields of research

— Tragedies, inquiries, failed organisations, turnaround strategies, role of media in safety improvement. A mentoring system to help organisations learn from their mistakes could be evaluated across intervention and control sites.

— Ethnographic studies of practice in OR & A&E.

John Øvretveit is Director of Research and Professor at the Medical Management Centre, Karolinska Institutet, Stockholm. He is also Professor of Health Policy and Management at Bergen University Medical School, Norway.

John's work is based on the belief that organization and management can bring out the best and worst in people, and that the right organization design is critical for effective healthcare: "the largest risk to health is a hidden one - poor health organization and management". A theme underlying his work is how practical research can contribute to healthy work organization and better care for patients. His recent book describes action evaluation methods for giving rapid feedback for service providers and policy-makers to improve their services.

John has published widely on the subjects of health service quality, health management, organization, evaluation, interprofessional cooperation and health reforms. He has undertaken health evaluation and development projects in Africa, Yemen, Indonesia, Thailand, Mexico, New Zealand, Australia, Japan, Sweden, Norway, Estonia, and the USA. Translations of some of his 280 scientific papers and books have been made into eight languages and a number have won publications awards, including twice winner of the European Health Management Publication award and Baxter award for "Action Evaluation", and in 1992 for "Health Service Quality", and the 1999 British Association of Medical Managers publication award for "Evaluating Health Interventions".

Paulo Sousa is Assistant Professor at Lisbon School of Health Technologies (in the field of Health Promotion and Health Administration), and Researcher at National School of Public Health, New University of Lisbon. The main areas of interest are Quality and Patient Safety, and Risk Adjustment in the analysis of outcomes in Healthcare.

In the last years Paulo coordinated two Post-Graduate courses focused on Health Management and Health Administration, promoted by Lisbon School of Health Technologies. He has published some papers on the subjects of Patient Safety, National Registries and Quality Improvement, Risk Adjustment, and Outcomes Research, in national and international Peer-review journals. He is part of the Quality Improvement Research Network (QIRN) since 2005 and was the QIRN Administrator for 2007. Nowadays Paulo is member of the Editorial Board of the "Revista Portuguesa de Cardiologia" (Portuguese Journal of Cardiology - The Official Journal of the Portuguese Society of Cardiology) and is Reviewer of the journal "Leadership in Health Services", from Emerald group.