

# Measuring the Patient Experience of Hospital Quality of Care

Submitted for Fulfilment of the Degree of PhD by Publication

Michelle Beattie

University of Stirling

March 2016



# Acknowledgements

This work has been influenced, directed and supported by many. I would like to thank the following people:

Professor William Lauder, University of West of Scotland (previously University of Stirling), Dr Iain Atherton, Edinburgh Napier (previously University of Stirling) and Dr Gill Hubbard, University of Stirling, who supervised the PhD, providing their time, support and encouragement throughout the journey.

Dr Douglas Murphy, University of Dundee, who shared his knowledge of psychometrics freely and inspired me to develop the CEFIT. His patience and generosity also enabled the time-consuming application of the COSMIN criteria within the systematic review.

Mr Brian James, Senior Teaching Fellow and fellow quality improvement (QI) enthusiast, for his additional share of our QI workload during peak times of the PhD journey. Also for his inspiring discussions, critical reading and helpful grammatical corrections.

Mrs Mariyana Schoultz, Lecturer and friend. The journey was made easier by sharing the highs and lows with a PhD 'buddy'.

Frances Hines, Research and Development Manager, NHS Highland for creating clarity in the early days when formulating my ideas. Frances continued to provide insightful ethical direction and support where necessary.

Staff at Highland Campus. I am grateful to many other colleagues who have listened to more PhD stories than most would tolerate. Also, for the practical 'cover' during my periods of research leave. Special thanks to the library staff for their expertise during literature searching and help in retrieving 'difficult' to find papers. To Kate for the continued IT advice and help. To Liz and Sue for their administrative assistance.

Lorraine Murphy, my best friend and fellow nurse, who helped keep my ideas grounded in the realities of clinical practice. Also, for providing practical help, such as after school pick-ups.

Finally, to Gordon Beattie, my husband, who has listened with interest whilst I shared many of my thoughts of measuring the patient experience of hospital quality of care. He

has also provided much parental care and support for our son, Oscar, now 8 years old. Their unwavering support and understanding was essential.

# Abstract

The primary motivation of this PhD by publication has been the apparent disconnect between the metrics of hospital quality of care at national and board level and patients' experiences. Exploration of the gap led to the realisation of two key points. Firstly, the concept of healthcare quality continually evolves. Secondly, the NHS Scotland Measurement Framework does not include a measure of patient experience at the microsystem level (e.g. hospital ward). This is needed to counterbalance easier to obtain metrics of quality (e.g. waiting times). Resource tends to follow measurement.

Papers 1 and 2 were exploratory, investigating theoretical and practical aspects of measuring quality of hospital care at the clinical microsystem level. With the associated Chapters, they highlighted both the necessity and the possibility of measuring the patient experience at the micro level of the healthcare system. They also drew attention to the inadequacy of "satisfaction" as a metric, leading to closer examination of "experience" as the decisive metric. This required the development of a systematic review protocol (Paper Three), then a systematic review (Paper Four).

The review (Paper Four) examined the utility (validity, reliability, cost efficiency, acceptability and educational impact) of questionnaires to measure the patient experience of hospital quality of care, with a newly devised matrix tool. Findings highlighted a gap for an instrument with high utility for use at the clinical microsystem level of healthcare. Paper Five presents the development and preliminary psychometric testing of such an instrument; the Care Experience Feedback Improvement Tool (CEFIT).

The thesis provides, as well as the matrix tool and CEFIT, theoretical and methodological contributions in the field of healthcare quality. It contributes to an aspiration that the patient's voice can be heard and acknowledged, in order to direct improvements in the quality of hospital care.



# Table of Contents

Chapter 1 .....	1
Introduction: Setting the Scene.....	1
1.1    The Challenge of Measuring Hospital Quality of Care .....	1
1.2    Timely and Relevant Measurement of Hospital Quality of Care.....	1
1.3    Why a Personal Interest in Hospital Quality of Care? .....	6
1.4    Why is the Patient Perspective of Hospital Quality Care Important?.....	8
1.5    Quality of Healthcare: The Governance Structure .....	9
1.6    Aims and objectives of the thesis .....	12
1.7    Summary and Link to Thesis Layout .....	13
Chapter 2 .....	17
What is the Definition of Healthcare Quality?.....	17
2.1    A Contemporary Definition of Healthcare Quality Is Needed .....	17
2.2    The Prominence of the IOM Domains of Quality .....	18
2.3    The Evolving Definitions of Healthcare Quality.....	19
2.4    Summary .....	24
2.5    Aim and linkage to research question.....	25
2.6    Paper One: Do the Institute of Medicines' (IOM) dimensions of quality capture the current meaning of quality in healthcare? – An integrative review .....	26
2.7    Overview of Paper One .....	44
2.7.1    Methods .....	44

2.7.2	Data Sources .....	44
2.7.3	Study Selection .....	45
2.7.4	Data Extraction.....	47
2.7.5	Analysis.....	48
2.7.6	Findings .....	49
2.8	Critical Reflection of Paper One .....	49
2.9	A Revised Model of Healthcare Quality Domains .....	52
2.10	Study Contribution to the Research Question.....	55
Chapter 3	.....	57
Measuring Hospital Quality of Care	.....	57
3.1	Measuring Domains of Quality: From Theory to Practice.....	57
3.2	Aim and Linkage to Research Question .....	57
3.3	Background to Paper Two.....	58
3.4	Paper Two: Compassion or speed? Which is a more accurate indicator of healthcare quality in the ED from the patient's perspective? .....	60
3.5	Overview of Paper Two.....	69
3.5.1	Methods .....	69
3.5.2	Team Preparation .....	69
3.5.3	Data Collection.....	70
3.5.4	Ethical and Legal Considerations .....	72
3.6	Critical Reflection of Paper Two .....	73



3.7	The Patient Perspective of Hospital Quality of Care: Conceptual Meaning and Measurement .....	75
3.8	Study Contribution to the Research Question.....	79
Chapter 4 .....		81
What instruments exist to measure the patient experience of hospital quality of care?81		
4.1	Aim and Linkage to Research Question .....	81
4.2	An Introduction to Psychometrics .....	82
4.2.1	Why are Psychometrics Important? .....	84
4.3	Instrument Utility .....	85
4.3.1	Validity .....	86
4.3.1.1	Content validity (inclusive of face validity) .....	86
4.3.1.2	Criterion validity .....	87
4.3.1.3	Construct validity (inclusive of hypothesis testing, structural and cross-cultural validity).....	88
4.3.1.3a	Hypothesis testing.....	88
4.3.1.3b	Structural validity.....	88
4.3.1.3c	Cross-cultural validity .....	89
4.3.2	Reliability.....	89
4.3.3	Cost efficiency.....	91
4.3.4	Acceptability .....	91
4.3.5	Educational Impact.....	92
4.4	Summary .....	92

4.5	Paper Three: Instruments to measure patient experience of healthcare quality in hospitals: a systematic review protocol.....	93
4.6	Overview of Paper Three .....	103
4.6.1	Clarifying the Concept.....	103
4.6.2	Application of Inclusion Criteria .....	103
4.7	Critical Reflection of Paper Three.....	104
4.8	Study Contribution to the Research Question.....	105
4.9	Paper Four: Instruments to measure patient experience of healthcare quality in hospitals: a systematic review .....	107
4.10	Overview of Paper Four .....	128
4.10.1	The complexity of Instrument Origin .....	128
4.10.2	The Use of COSMIN Criteria .....	128
4.10.3	The Development of Additional Aspects of Instrument Utility .....	129
4.10.4	Beattie and Murphy Instrument Utility Matrix.....	130
4.11	Critical Reflection of Paper Four.....	130
4.11.1	Reflective Versus Formative Models .....	130
4.11.2	Application of COSMIN Checklists .....	131
4.11.3	Application of Quality Criteria for Results of Measurement Properties... .....	132
4.12	Study Contribution to the Research Question.....	133
Chapter 5 .....		135
Developing and testing a measure of patient experience of hospital quality care .....		135
5.1	Aim and Linkage to Research Question .....	135

5.2	Paper Five: Development and Preliminary Psychometric Properties of the Care Experience Feedback Improvement Tool (CEFIT) .....	137
5.3	Overview of Paper Five .....	150
5.3.1	CEFIT Item Construction .....	150
5.3.2	CEFIT Rating Scale.....	151
5.3.3	CEFIT Scoring.....	151
5.3.4	Content Validity Index Procedure .....	152
5.3.5	Queensland Social Survey .....	153
5.3.6	Analysis and Results .....	154
5.3.7	CEFIT Quality Critique .....	155
5.4	Critical Reflection of Paper Five .....	156
5.4.1	Limitation Between Healthcare and Hospital Context .....	156
5.4.2	Limitation of Range .....	157
5.4.3	Statistical Methods .....	157
5.5	Study Contribution to the Research Question.....	158
Chapter 6 .....		159
Discussion.....		159
6.1	Introduction .....	159
6.2	Summary of Findings and Contribution to Improving Hospital Quality of Care from the Patient Experience .....	160
6.2.1	Objective 1: To determine what domains capture the contemporary meaning of healthcare quality to inform the development of a theoretical model of quality of healthcare.....	160

6.2.2	Objective 2: To test whether empathy, as an indicator of caring, can be measured in an acute hospital setting, from a theoretical and practical perspective.....	162
6.2.3	Objective 3: To identify and critique the utility of existing instruments which measure the adult inpatient experience of hospital quality of care. ....	164
6.2.4	Objective 4: To develop a valid, reliable and brief measure of patient experience of hospital quality of care.....	166
6.3	Collective Contribution and Implications.....	168
6.3.1	Practice Implications .....	168
6.3.2	Policy Implications.....	172
6.3.3	Research Implications.....	174
6.4	Limitations.....	176
6.4.1	Testing CEFIT in an Australian Survey.....	176
6.4.2	CEFIT Scoring on Additional Aspects of Instrument Utility.....	176
6.5	In Summary.....	177
Chapter 7 .....		179
Conclusion and Dissemination .....		179
7.1	Final Thoughts .....	179
7.2	Standing of the Journals and Contribution to Published Works .....	179
7.2.1	Paper One: An integrative review of dimensions of quality (Beattie et al 2012).....	179
7.2.2	Paper Two: A cross-sectional study measuring empathy (Beattie et al 2012).....	181

7.2.3	Paper Three: A protocol for systematic review and utility critique (Beattie et al 2014).....	181
7.2.4	Paper Four: A systematic review and instrument utility critique (Beattie et al 2015).....	182
7.2.5	Paper Five: Development and preliminary testing of CEFIT (Beattie et al 2016).....	183
	References.....	185
	Appendices .....	211
	Appendix 1: Local Delivery Plan (LDP) Standards.....	212
	Appendix 2: Patient Information Leaflet .....	213
	Appendix 3: University of Stirling Ethics Approval and Submission.....	214
	Appendix 4: National Health Services Research and Ethics Committee (NHS REC) North of Scotland, Ethics Application and Approval .....	241
	Appendix 5: Care Experience Feedback Improvement Tool (CEFIT).....	281
	Appendix 6: Patient Feedback Tool Ethical Decision Letter.....	282
	Appendix 7: Expert Feedback Form: Content Validity Index .....	284
	Appendix 8: CEFIT Results of COSMIN Checklist for Content Validity .....	285
	Appendix 9: CEFIT Results of COSMIN Checklist for Structural Validity.....	286
	Appendix 10: CEFIT Results of COSMIN Checklist for Internal Consistency ....	287
	Appendix 11: Quality Criteria for Measurement Properties (Terwee et al 2007)	288
	Appendix 12: Additional Aspects of Utility Scoring Criteria and CEFIT Results.	289
	Appendix 13: Approval to Use Published Paper One in Thesis.....	290
	Appendix 14: Approval to use Paper Two in Thesis.....	291

## List of Tables

Table 1: National Institute for Health and Clinical Excellence Hierarchy of Evidence .	48
--	----

## List of Figures

Figure 1: NHS Scotland Quality Measurement Framework.....	2
Figure 2: Beattie’s Model of Healthcare Quality.....	52
Figure 3: Institute of Medicine’s 6 Domains of Quality .....	52
Figure 4: Process Flow Diagram for Data Collection .....	72
Figure 5: COSMIN Domains and Definitions of Measurement Properties .....	84
Figure 6: Framework of Quality Critique of Instrument Utility .....	86
Figure 7: Example of Question and Response Using the COMIN 4-Point Rating Scale to Assess Methodological Quality .....	129

# Chapter 1

## Introduction: Setting the Scene

### 1.1 The Challenge of Measuring Hospital Quality of Care

Improving the quality of hospital care continues to be a global challenge. This thesis and its associated publications represent a series of studies cumulating in the development of a valid, reliable, but brief, instrument to measure the patient experience of hospital quality of care.

How hospital quality of care is measured matters, as limited hospital resources are often directed to what is being measured (Berry et al 2015). If what is measured, or the way in which it is measured is not accurate, there is a real risk that efforts to improve hospital quality of care are at best futile and at worst exacerbating problems. At the clinical microsystem level, resources may be wasted as nurses continue to implement change without establishing whether the intervention is making an improvement, potentially reducing contact time with patients for no real benefit. Meanwhile, the challenge of assuring consistent hospital quality of care would remain.

In this Chapter, the challenge of timely and relevant measurement of hospital quality of care to drive improvements is outlined in Section 1.2, followed by a brief summary of my personal motivation for addressing this challenge (Section 1.3). Section 1.4 critically discusses why it is important to measure hospital quality of care from the patient perspective. Section 1.5 describes the governance arrangements for hospital quality of care in Scotland, which is where the studies for this thesis were conducted. It therefore provides useful contextual information for the studies and articles that comprise the thesis. Section 1.6 outlines the aims and objectives of the thesis and identifies the published articles associated with each specific objective. Finally, Section 1.7 describes the structure of the thesis.

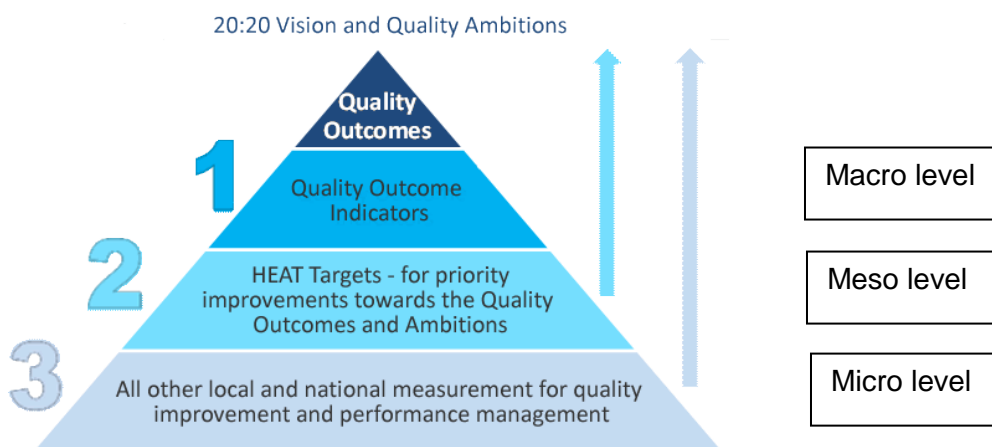
### 1.2 Timely and Relevant Measurement of Hospital Quality of Care

The studies that comprise this thesis were influenced by my increasing awareness of, and discomfort about, the apparent disconnect between the reported metrics of national and board-level hospital quality of care and the experiences of individual patients. That is not to say that all hospital care is of poor quality, rather, the detail of reports of hospital

quality (whether of positive or negative results) do not consistently reflect the experiences of patients at an individual or ward level. This discrepancy required further exploration; first, existing measures of hospital quality of care may not include a measure of the patient perspective of quality of care, which could account for the disparity between hospital and patient reports of hospital quality, and second, measurement at national or board level may not necessarily capture individual or ward-level quality of care from the patient perspective.

The NHS Scotland Quality Measurement Framework demonstrates how quality is measured at different levels within the healthcare system (Information Services Division 2010). The framework is represented as a pyramid demonstrating the interconnected, yet hierarchical, nature of healthcare quality measures (see Figure 1). The framework can also be viewed from a systems-level perspective; containing measures at macro, meso and micro levels. At the top of the pyramid are the Quality Outcomes; or policy ambitions for Scotland’s health service to be safe, person-centred and effective (Scottish Government 2010, Scottish Government 2011a). Level 1 of the framework describes Quality Outcome Indicators which are high level strategic measures set to achieve the quality ambitions (macro level). There are 12 Quality Outcome Indicators. One of these indicators is a measure from the patient perspective, referred to as the care experience indicator. The measure is derived from the Inpatient Patient Experience Survey – a National Annual Survey of inpatient experience. The survey data are used to measure National- and Board-level performance. Data are also available at hospital level within each Board.

**Figure 1: NHS Scotland Quality Measurement Framework**



Level 2 depicts HEAT (Health Improvement, Efficiency, Access, and Treatment) Targets (soon to be replaced by Local Delivery Plan (LDP) Standards) which are used to assure



the Scottish Government of the performance of NHS Boards and can be considered meso level measures. The main change from HEAT to LDP is the integration of health and social care services and their associated measures. There are currently 18 LDP Standards (see Appendix 1). These standards do not include a measure of hospital quality of care from the patient perspective, rather, they are specifically around waiting time, financial management and staff absence. However, the LDP Guidance includes person-centred care as a priority area and requests NHS Boards to demonstrate how they will determine improvement in this area, including how progress will be measured locally (Scottish Government 2016).

It is anticipated that Level 3 measures will feed into the attainment of Level 2 measures, and likewise, Level 2 measures will feed into Level 1. Finally, Level 3 of the pyramid includes all other local and national measures for improvement and performance management. These can be further subdivided into those necessary for compulsory reporting and those driven by local improvement initiatives.

Examples of Level 3 compulsory measures, or micro-level measures, include requirements for all Health Boards in Scotland to submit monthly data on the number of adverse events, complaints and patient safety metrics to Healthcare Improvement Scotland (HIS). There are other compulsory reporting systems which are linked to quality of care: reporting all sudden and unexpected deaths to the local Procurator Fiscal, the necessity for a local significant event reviews, the reporting of all suicides to the Mental Welfare Commission (Scotland), Reporting of all Incidents, Diseases and Dangerous Occurrences (RIDDOR) at work to the Health and Safety Executive and the reporting of all adverse medication reactions using the Yellow Card Scheme (Crown Office 2008, HIS 2016, HSE 2014, Medicine and Healthcare Products Regulatory Agency (MHRA) 2014). The Yellow Card Scheme initially started as a yellow page insert (hence the name) within the British National Formulary (reference guide for prescribing and administering medicines). Practitioners are required to complete a Yellow Card for any adverse event associated with a patient's medication, for example if a patient develops breathlessness after administration of a newly prescribed medicine, and return it to the MHRA. The Yellow Card Scheme still exists as a paper format, but there is now also an online version.

Examples of level three measures for local improvement work include completion of peripheral vascular cannula (PVC) insertion bundles, hand-washing compliance, pressure care bundles, safety briefs and the use of SBAR (Situation, Background,

Assessment and Recommendation) for patient handover within a hospital ward setting. Some hospital wards within NHS Scotland may need to audit and record as many as twenty-five different care process measures per month (Personal Communication 2015).

The purpose of measurement may also differ at each level of the system levels. For example, data collected on quality of healthcare at level one (macro level) are likely being used for judgement and scrutiny to assure quality of services, whereas data collected at ward level (level three) may be used for scrutiny, but are more likely to be used for improvement. Also, ownership of the data at level one is likely to be external to those involved in direct patient care (such as Information Services Division), whilst level three data are more likely to belong to clinicians or the healthcare organisation. This is important to consider when devising a measure of quality from the patient perspective as there are important implications of the robustness and accessibility of data for instrument design (Davies 2006). Measurement arguably becomes more robust as the stakes for data use increase (as discussed later in the thesis).

Whilst quality of care is clearly subject to much scrutiny and measurement there are gaps in relation to measures, specifically from the patient perspective of hospital quality of care, at the micro level of the healthcare system (level 3), such as the hospital ward. Evidence also suggests that there is a focus on aspects of quality which are more amenable to measurement, for example, waiting time (Wiig et al 2014a). The National Framework demonstrates that care is mostly measured and monitored from clinical and managerial perspectives. There is a patient perspective measure at the macro level (Level 1), which provides information on patient experience of hospital care for National and Board comparison of performance. These data are not, however, timely, nor specific enough, to direct or measure local improvement efforts at the clinical microsystem level. For example, the macro measure includes criteria for sampling patients (such as those discharged between the months of January and April). Whilst such criteria are necessary to ensure a robust sampling procedure, the delay between sampling checks, data transfer, postage of survey, data entry and coding and analysis means that the results are released one year following the patient experience of hospital care. Given that much change can occur annually within a hospital, it would be difficult to make recommendations at the clinical microsystem (i.e. ward) level from these data.

Further, the measure is also unlikely to be specific enough to drive quality improvements at local levels. The macro level results from the National Inpatient Survey are available at Board and hospital level. However, if, for example, the hospital level data identified

that there was a statistically significant deterioration in patient experience around privacy and dignity, there is no way of knowing where in the hospital this problem originates. Given that some hospitals can have up to 48 different clinical specialities, there are likely to be many wards and clinical areas within most hospitals in Scotland, thus making identification of areas for improvement difficult (ISD 2015). Similarly, episodes of positive patient experiences cannot be linked to specific wards or teams, thus limiting the receipt of positive clinician feedback and the ability to spread good practice. There is a vast amount of improvement activity at ward or unit level within hospitals, yet these changes are not consistently measured from the patient perspective. For example, a local improvement initiative may be implementing open visiting times on a ward. Anecdotal evidence from patients and families may suggest progress, but there is currently no brief measure that can be routinely collected within clinical practice to measure ongoing improvement, or change.

The disparity in measurement of the patient perspective of hospital quality of care between macro, meso and microsystem levels within the Scottish healthcare system is likely contributing to the disconnect between reports of hospital quality of care and that actually experienced by individual patients. Data from the macro level patient experience survey in Scotland suggest that the overall quality of hospital care is good, from the patient perspective. For example, results from the 2014 survey show that 83% of patients rated their care as good or excellent, a 2% increase from 2012. Similarly, 87% of patients reported their Accident and Emergency care and treatment as good or excellent, a 4% increase since 2012 (Scottish Government 2014a). However, this is not the whole story.

Whilst there are many positive experiences of hospital care, the evidence suggests that the quality of care is variable and often inadequate (Jha et al 2005, Right Care 2011). That is, people in hospital do not receive high quality of care every time. There have been high profile cases where poor care has been endemic (Department of Health 2013a, Francis 2013). Stories of poor patient care appear on a regular basis in the local and national media. There are more hospital complaints and litigation cases than ever before (ISD 2014). The number of complaints reported in NHS Scotland in the year 2013-2014 was 20,364 (an increase of 20% from previous year). The local NHS Board has had a 33% increase in complaints over the same year (ISD 2014). Without an ongoing measure at the clinical microsystem level it is difficult to tell whether these figures are confined to specific areas or teams or whether this is more reflective of a widespread problem of poor hospital quality of care. There is a pressing need for

improvement in hospital quality of care, but a measure from the patient perspective is necessary to direct improvement efforts at the clinical microsystem level. Thus, the challenge of measuring the patient perspective of hospital quality of care, which was relevant and timely at ward or micro level quality improvements, was what this thesis set out to address.

### **1.3 Why a Personal Interest in Hospital Quality of Care?**

From 20 years spent nursing in UK and overseas hospital settings, it has been possible to witness the joy experienced by patients and families when they, or their loved ones, have made remarkable recoveries despite poor odds. Other observations of quality hospital care are more subtle; the nurse who arrived early to give a patient a newspaper not stocked in the hospital shop, the porter who waited past shift time to prevent the patient waiting too long for return transport from the x-ray department, or the nurse who spent time tracing a wound circumference in order to demonstrate that the wound was indeed improving in order to reassure the patient of progress.

For me, acute nursing is synonymous with the quality of hospital care. Nursing has been defined as using clinical judgement to enable people to improve, maintain, or recover health, to cope with health problems, and to achieve the best possible quality of life, whatever the disease or disability, until death (RCN 2014). The day-to-day care of patients in hospital is largely dependent upon nurses, who constitute the largest professional group in healthcare. Nurses are ever-present and highly visible to patients and their families, who are often at their most vulnerable. They are in a unique and privileged position, from which to provide high quality of care, and to detect and intervene when care standards fall short (Carroll 2005). The literature on nursing care supports the synergy between nursing and quality of care. A systematic review and meta-analysis of 130 empirical studies identified the positive patient outcomes associated with high quality nursing care as enhanced emotional well-being, physical healing, trusting relationships and reduced cost (Swanson 1999).

I have experienced the personal satisfaction of feeling needed and valued when patients and families receive and report good quality of care at the ward level. There is an intrinsic reciprocal benefit encompassed in patient/nurse encounters when providing a high quality of care. Whilst nurses can be portrayed as selfless, most nurses would acknowledge their own gratitude and personal benefit when they are directly involved in a person's recovery, positive experience, or even peaceful death. When nursing is associated with good quality of care there is a sense of accomplishment and purpose

and a developing respect for life and death (Watson 2009). Early psychological literature identified that gratitude and perceived need of the recipient are important parameters in the cost/benefit ratio of altruistic behaviour (Trivers 1971). Indeed, it can be argued that nurses are motivated to demonstrate altruism, due to the gratitude they experience and the needs of their patients.

I have, however, observed the negative impact on patients, families and clinical practitioners when care is not of the expected quality. I worked for two years as a Safety, Governance and Risk Co-ordinator (SGRC) for an NHS Board in Scotland. Part of the role was to monitor and report measures of adverse events and near-miss incidents. This highlighted the high frequency of adverse events and wide variation in hospital quality care. The post included facilitating significant event reviews of the most serious adverse events, where the patient outcome had been death or significant harm. Significant event reviews aim to establish, in a non-punitive way, what actions occurred and why, involving all of the participants in the event, and make recommendations for organisational learning and improvements (Gillam and Siriwardena 2013). Indeed, it was the effect of some of these events that has motivated this collection of works to make a contribution to improving the quality of hospital care. The SGRC post afforded me the opportunity to view the quality of hospital care from a wider perspective than many are privy to; it challenged my naive assumption that having good nurses would result in good quality care. Such a linear cause and effect solution could not, and cannot, hold true in highly complex environments, such as those found in acute hospitals.

I have experienced the impact of poor care from the perspective of a bereaved relative. My relative died after a short illness and a four-week hospital admission. As their condition deteriorated they were transferred to a higher level of care where clinical staff worked tirelessly and the technical care was excellent. However, there were many aspects of the hospitalisation which lacked safe and compassionate care: the nurse who blamed his breathlessness on non-adherence of instructions to sit upright (this was in fact a symptom of undiagnosed renal failure); relatives being told to move out of the family room in intensive care as another patient was 'more sick'. The lack of compassion and dignity was, at times, difficult to comprehend. Yet, at the same time, the hospital was publicised as a top performer in patient safety metrics. This difficult experience reaffirmed for me the gap between the patient perspective of hospital quality of care and the data used to measure the quality of hospital care.

In summary, nursing is inextricably linked to the quality of hospital care, but it has been possible to identify a chasm between the quality of care experienced by patients and that which is reported in hospital quality metrics. Current methods of measuring the quality of hospital care may not be valid from the patient perspective, therefore limiting the potential to improve patient care.

#### **1.4 Why is the Patient Perspective of Hospital Quality Care Important?**

Measuring and acting on issues of quality raised by patients can be a partial solution to this persistent problem of poor hospital quality of care (Rathert et al 2011). Patients, through their unique experiences, can offer insights into hospital quality, which would be unseen from other perspectives, such as the way a treatment, process or interaction has made them feel and, subsequently, behave. Due to the complexity of hospital systems, with many care transitions and multiple providers, patients are often the only people to view the quality of hospital care holistically (Rathert et al 2011).

Patients who report poor hospital quality of care are often found to have poor clinical outcomes and an increased length of stay, which leads to psychological distress for families and staff members and an overall reduction in public trust (Aiken et al 2008, Doyle et al 2013, Health Foundation 2011). There are reputational and financial costs to health services from litigation cases when patients report poor quality of hospital care with associated increased costs from longer and more expensive periods of hospitalisation (Gailey and Cachia 2010). There is increasing evidence that patients who have positive healthcare experiences have improved outcomes, resulting in a more efficient healthcare system (Department of Health 2013b, Sofaer and Firminger 2005). The necessity of hearing the patient perspective is not a new concept. However, recent aspirations towards 'person-centred' care and 'mutual' healthcare services have reaffirmed the imperative for clinicians and healthcare managers to listen to the patient perspective and act accordingly to direct improvement efforts. There is a need to gather data on the patient perspective of hospital quality of care in a robust and timely way.

Measuring the patient perspective is now an important aspect of hospital quality monitoring and reporting. As previously mentioned, the Scottish Inpatient Patient Experience Survey (SIPES) analysed data for 21,127 patients from 14 NHS Health Boards (Scottish Government 2014a). The data are primarily used as a national performance indicator of quality of hospital care from the patient perspective in Health Boards in Scotland. Similarly, the National Health Service Inpatient (NHSIP) Survey for England has been operating annually since 2002 (Picker Institute Europe 2012). For

both surveys the data are collected annually and used to benchmark health providers and enable year-on-year comparisons. Whilst the data are useful to determine variations between health providers and sub-groups of patients, it does not adequately capture the views of patients within individual wards or units within hospitals. For example, the sampling strategy is across a whole hospital over a specific time period, therefore the final sample may only include one or two patients from a particular ward. The surveys are also lengthy, which limits their use by hard-pressed clinical teams as an ongoing method of measurement for improvement within clinical areas.

## **1.5 Quality of Healthcare: The Governance Structure**

Much of what happens operationally around measuring the quality of hospital care is determined by existing governance structures, therefore these structures are explained to help set the scene for the studies and associated articles included in the thesis. The collection of studies comprising this programme of work for the PhD was conducted mostly in Scotland; thus the governance structure described below is for Scotland. Other similar structures for healthcare quality, including England and the United States (US), are described for comparison. NHS Scotland and England were governed, until recently, by the Westminster Government and the Department of Health and it is only since devolution in 1998 that healthcare governance structures within Scotland have changed. What happens in England's health service is regularly reported in Scotland. Further, key professional governing bodies, such as the Nursing and Midwifery Council (NMC) governs nursing in both Scotland and England. Thus England's health service continues to inform the Scottish public's understanding of hospital quality of care. The US has been a highly influential country with regards to healthcare quality improvement. Indeed, as discussed in detail in Chapter 2, a United States of America (US) organisation, the Institute of Medicine (IOM), has influenced definitions and understandings of healthcare quality for over a decade. Thus the governance structures of healthcare in England and the US are also briefly presented in this section.

The National Health Service (Scotland) Act 1978 states that it is the duty of all Health Boards to monitor and improve healthcare. Statutory duties for quality were devolved to Healthcare Improvement Scotland (HIS) in 2011, including a general duty to further the quality of healthcare and a duty to provide public information about the quality of care within Health Boards (Scottish Government 2014b). Responsibility for reporting via these compulsory systems is usually devolved from the Chief Executive of each Health Board to others working in the healthcare system. However, ultimately, the Chief

Executive remains accountable for the governance structure within their Health Board. HIS is the independent body, in Scotland, for healthcare governance, but it is also responsible for supporting Health Boards in improvement activities. There has been much debate as to whether one organisation should be responsible for healthcare quality improvement, as well as inspections; with concerns over the need for independent scrutiny (Davies et al 2002, RCN 2009, Scottish Government 2007a).

The NHS England governance structure for quality of hospital care differs from that of NHS Scotland. NHS England has two separate organisations; one for healthcare quality governance, and one for improvement. The Care Quality Commission (CQC), with statutory functions enshrined by the Health and Social Care Act (2008) and the Care Act (2014), routinely conducts audits and inspections similar to those made by HIS (CQC 2015, DoH 2010a). Quality improvement activity, however, is supported by NHS Improving Quality (NHS IQ). NHS IQ is part of NHS England and is accountable to the Department of Health. This is important when measuring the patient perspective of hospital quality of care as the type of data used is dependent upon the purpose for which the data will be used (this will be further discussed later in the thesis).

For comparison, in the US, healthcare governance is often derived from within the private healthcare organisations which deliver care. However, the Institute of Medicine (IOM), a non-profit organisation, provides the US Government and other industries with non-biased information on healthcare quality. In 2000, the Health Quality Alliance (HQA) was formed to encourage voluntary reporting of various quality indicators by hospitals across the US. The HQA is a consortium of organisations involved in quality of care, including the Centers for Medicare and Medicaid Services (CMS), the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), the American Hospital Association, and the American Association of Retired Persons (Jha et al 2005). Through the HQA system, hospitals across the US report to CMS on indicators of hospital quality of care. In 2008, the reporting became tied to the Annual Payment Update (APU) for the Inpatient Prospective Payment Systems (IPPS). This means that hospitals who do not submit their data on quality of hospital care may be subject to a 2% reduction in their APU (Giordano et al 2010). Using the patient perspective as a measure of hospital quality of care has been adopted widely in the US by means of associated financial incentives. These incentives are not used in the Scottish healthcare hospital quality reporting system. It is, therefore, important to consider intrinsic motivational factors, such as 'making the right thing, the easy thing to do' when designing a measure from



the patient perspective of hospital quality of care for use in Scotland. For example, ensuring the instrument is brief and easy to use.

Nurses also have a professional responsibility to ensure care is of the expected standard. The Nursing and Midwifery Council (NMC) sets professional standards of practice and behaviour for all nurses and midwives, commonly known as The Code (NMC 2015). The Code defines what good nursing care looks like and sets professional standards for public protection within the UK. Revised in 2015, the code is designed around four key themes; prioritising people (putting patients first), practising effectively (using and documenting best evidence), preserving safety (identifying and reporting risk) and promoting professionalism and trust (upholding public confidence). The principles and statements are reflective of good quality of care from a public and professional perspective. The NMC exists to protect the public, therefore nurses failing to meet the standards of The Code are subject to a fitness to practise review. Any member of the public or healthcare professional can report a nurse to the NMC if the quality of care he or she delivers to inpatients is of poor quality. The NMC can remove any nurse or midwife from the register, thereby preventing them from practising. Nurses, therefore, do have some external influences which necessitate improving hospital quality of care from the patient perspective. Such improvement of patient perspective can only be determined by measurement (Scales and Schulman 2014).

In summary, hospitals in the UK have a legal duty to provide, monitor and improve quality of care. Quality of care becomes the business of every employee but overall governance remains the responsibility of the Chief Executive of each NHS Board in Scotland, or Clinical Commissioning Group in England. Other developed countries have similar systems, often with financially linked incentives. Nurses also have a moral and professional responsibility to provide high quality healthcare. There are some external drivers (for example, policy and law) influencing the need to measure the patient perspective of hospital quality of care, but intrinsic factors to motivate teams will likely remain important. To determine whether or not these obligations are being achieved and the patient perspective of hospital quality care is indeed improving necessitates the employment of an instrument to measure hospital quality care with high utility.

## 1.6 Aims and objectives of the thesis

The overall aim of this collection of studies was to address the challenge of timely and relevant measurement of hospital quality of care to drive improvements in care at ward level. The collection contributes to the evidence base of measuring the quality of healthcare and provides an instrument to assist the patient voice to be heard in efforts to improve the patient experience of hospital quality of care. This was attained through the following five objectives:

### **Objective One:**

To devise a theoretical model of quality of healthcare informed by an integrative review of the literature.

### **Publication One**

**Beattie M**, Shepherd A, Howieson B (2012) Do the Institute of Medicine's (IOM's) dimensions of quality capture the current meaning of quality of healthcare? – An integrative review. *Journal of Research in Nursing*. 18 (4), 288-304.

### **Objective Two:**

To test whether empathy, as an indicator of caring, can be measured in an acute hospital setting, from a theoretical and practical perspective.

### **Publication Two**

**Beattie M**, Atherton I, McLennan B, Lauder W (2012) Compassion or speed, which is a more accurate indicator of healthcare quality in the emergency department from the patients' perspective? *International Journal of Person Centered Medicine*, 2 (4), 647-655.

### **Objective Three:**

To identify and critique the utility of existing instruments which measure the adult inpatient experience of hospital quality of care.

### **Publication Three**

**Beattie M**, Lauder W, Atherton I, Murphy D. (2014) Instruments to measure patient experience of healthcare quality in hospitals: a systematic review protocol. *Systematic Reviews Journal*, 3; 4. **Highly accessed.**

### **Publication Four**

**Beattie M**, Murphy D, Atherton I, Lauder W. (2015) Instruments to measure patient experience of healthcare quality in hospitals: a systematic review. *Systematic Reviews Journal*, 4; 97. **Highly accessed.**

### **Objective Four:**

To develop a brief measure of patient experience of hospital quality of care which is structurally valid and reliable.

### **Publication Five**

**Beattie M**, Shepherd A, Lauder W, Atherton I, Cowie J, Murphy D. (2016) Development and preliminary psychometric properties of the Care Experience Feedback Improvement Tool (CEFIT). *BMJ Open*:6:e010101. doi:10.1136/bmjopen-2015- 010101.

## **1.7 Summary and Link to Thesis Layout**

This collection of publications and narrative explains the journey of this doctoral work and the contribution to the field of healthcare quality, specifically in relation to the patient perspective of quality of care and how it can be measured at the clinical microsystem level within hospitals. The studies collectively address the challenge of timely and relevant measurement of hospital quality of care. Whilst the contribution remains in the field of healthcare quality, the work begins as exploratory in nature to further refine and define the research objectives. Therefore, although the final objective was to develop a measure of patient experience of hospital quality of care with high utility, this objective only arose in light of an accumulation of knowledge and investigation from the first two studies. These papers informed the direction and development of the final contribution to address the challenge of timely and relevant measurement of hospital quality of care within this thesis, namely the Care Experience Feedback Improvement Tool (CEFIT). The remaining three publications are linked, as the necessity to devise a measure of

patient experience of hospital quality of care for quality improvement purposes became clear as the work progressed.

Chapter 2 begins by exploring the complexity of defining quality of healthcare and the necessity to represent the concept of healthcare quality as multiple domains. The predominance of the Institute of Medicines (IOM) domains of quality are identified and the evolving nature of domains becomes apparent through a review of historical contributions. This leads to the need to question the currency and relevance of the IOM domains of healthcare quality through an integrative review (Paper One). A key argument throughout this thesis is that any model or measure of hospital quality of care must be relevant and timely. Thus, in Chapter 2, a new model of quality of care is presented, which is arguably more relevant to current UK hospital quality of care than, for example, IOM's model, which was developed over a decade ago. The findings are then used to develop a revised model of the domains of healthcare quality. The model is used later in Chapter 5 to inform the development of CEFIT.

Chapter 3 examines whether domains of healthcare quality, which are arguably less amenable to measurement, can be quantified in a hospital setting. Specifically, the Consultation and Relational Empathy (CARE) measure is used to determine whether empathy can be measured in the Emergency Department, and to establish whether empathy and/or waiting time are important indicators of hospital quality of care from the patient perspective (Paper Two). The paper makes three important contributions to the direction of the thesis. Firstly, the study confirms that domains of healthcare quality less amenable to measurement can be quantified at the clinical microsystem level. Secondly, the limitations of using patient satisfaction to capture the patient perspective are highlighted; subsequently directing efforts to measure 'experience' as opposed to 'satisfaction.' Thirdly, the CARE measure was designed for individual practitioner feedback and covers one domain of what constitutes quality of healthcare. Therefore, there remained a need to identify a measure of patient experience which captured all domains of healthcare quality and was suitable for ongoing quality improvement measurement.

Chapter 4 set out to establish what instruments (questionnaires) already exist to measure the patient experience of hospital quality of care. In order to conduct a critique of existing measures there was a need to understand psychometrics. The Chapter therefore begins with a brief explanation of the categories of validity and reliability used and justifies the necessity to take a holistic view of instrument utility. The methods

planned for a systematic review are presented in a protocol (Paper Three). The full systematic review is also embedded within this Chapter (Paper Four). The results of the systematic review found no instrument to measure the patient experience of hospital quality care which was suitable as a measure for quality improvement at the ward level of a hospital.

Chapter 5 describes the development and preliminary psychometric testing of the Care Experience Feedback Improvement Tool (Paper Five). The primary purpose of CEFIT is to use patient experience for quality improvement purposes at the ward level within the healthcare system. The tool is brief enough that it could be routinely employed to collect data for improvement within clinical areas.

Within the Chapters, each paper is followed by an overview, critical reflection and a detailed contribution to the thesis as a whole. The overview enables additional information to be shared which was omitted from the publication due to word count limits. Chapters have different subheadings for the overview as the relevant additional information is different for each publication. Similarly, each paper includes a critical reflection, which critiques the methods and personal learning, enabling demonstration of my research development.

Chapter 6 (Discussion) considers the contribution of the thesis and papers as a collective contribution in the field of healthcare quality. Limitations are acknowledged before considering the wider implications of the collective thesis for practice, policy and research in healthcare quality.

Finally, Chapter 7 (Dissemination) details my individual contribution for each publication and the standing of the journals in which the papers were published. Other mechanisms for dissemination and impact are also highlighted.



# Chapter 2

## What is the Definition of Healthcare Quality?

### 2.1 A Contemporary Definition of Healthcare Quality Is Needed

The focus of the thesis is hospital quality of care, which fits within the broader concept of healthcare quality. Thus, the first study and its associated published article was about contemporary definitions of healthcare quality. A key objective of the research was to devise a theoretical model of quality of healthcare informed by an integrative review of the literature. This is because before hospital quality of care can be measured, it is first necessary to define healthcare quality. There are two reasons which support the need for conceptual clarity. Firstly, an important step in instrument development is to define and conceptualise the construct of interest, in this case, quality of healthcare (De Vet et al 2011). Secondly, what constitutes healthcare quality may change as society changes, so it is possible that defining healthcare quality also evolves. Therefore, there is a need to establish a *current* conceptualisation of healthcare quality and appreciate the potential impact of evolution of the development of a measure of hospital quality of care from the patient perspective. A contemporary conceptualisation of healthcare quality can be used to devise a model of healthcare quality, which will provide the foundations for a measure of patient experience of hospital quality of care.

This Chapter has two main parts. Part one presents a historical overview of the concept of healthcare quality which highlights two important points. Firstly, the predominance of the Institute of Medicine (IOM) dimensions (Safe, Timely, Efficient, Effective, Equitable and Patient-centred), which provide a basis to critique current conceptions of healthcare quality. Secondly, that healthcare quality is ever evolving and therefore dependent on context and time. This has important implications for developing a measure of hospital quality of care from the patient perspective.

Part two of the Chapter presents a theoretical model of healthcare quality, which was informed by an integrative review of the literature (Paper One). A critical reflection of the paper is presented to highlight the methodological limitations of the study (Section 2.8). This is followed by a discussion of the substantive contribution of this specific paper to the main aim of the research, that is, to provide a timely and relevant measurement of hospital quality of care, from the patient perspective, to drive improvements in care, at a ward level. In particular, the review highlights additional domains of healthcare quality

to those proposed by the IOM, as well as highlighting the foundational nature of person-centred care for healthcare quality (Section 2.8). A revised model of the IOM domains of healthcare quality was subsequently developed and is presented in this Chapter (Section 2.9). This model was used later (Chapter 5) in the development of a measure of hospital quality of care from the patient perspective.

## **2.2 The Prominence of the IOM Domains of Quality**

Whilst the healthcare policy context in the US differs to Scotland, some of their policy has had a significant influence on the definitions and conceptions of healthcare quality in Scotland and beyond (Barelds et al 2009a, Department of Health 2008, Haggerty et al 2007, Heenan et al 2010, Scottish Government 2010). The IOM provides the US Government and the private healthcare industry in the United States with non-biased information on healthcare quality. They have produced seminal texts on quality, which have influenced approaches to healthcare quality across the world (IOM 1999, IOM 2001). The IOM also instigated the formation of the Health Quality Alliance (HQA) in the US (mentioned in Chapter 1), which is a consortium of organisations with an interest in healthcare quality, and which incentivised the reporting of healthcare quality data by linking it to the Inpatient Payment System (Giordano et al 2010).

At the turn of the new millennium, the IOM made a considerable contribution to the understanding of quality in healthcare in the publications “To Err is Human” (IOM 1999) and “Crossing the Quality Chasm” (IOM 2001), both of which have influenced UK healthcare policy and beyond (DoH 2010b, DoH 2013a, Scottish Government 2007a). The first of these publications, “To Err is Human” (IOM 1999), exposed the risks associated with being a patient in hospital and the consequent high rates of adverse events. It drew on literature from other high risk industries to recognise the role of human factors and systems thinking when things go wrong. It created a step change in healthcare quality, from a ‘blame’ to a ‘just’ culture, advocating openness to enable individual and organisational learning. This changed the approach to the management of adverse events in healthcare internationally (Stelfox et al 2006). Responses to events are now more focused on system changes as opposed to individual reprimand, or at least, moving in that direction (Stelfox et al 2006).

“Crossing the Quality Chasm” (IOM 2001) exposed the variations in quality of care and called for the need to take action to ensure more equitable healthcare provision. The IOM acknowledged that, although healthcare outcomes for some were improving, the gap in health inequalities was widening. According to the IOM, healthcare quality was



largely dependent on social class, location and ethnicity (IOM 2001). It was in “Crossing the Quality Chasm” that the STEEEP acronym first appeared as shorthand for the domains of quality of healthcare (Safe, Timely, Efficient, Effective, Equitable and Patient-centred). The two IOM publications contribute two things. Firstly, the IOM highlight the importance of improving the quality of healthcare and the implicit need to measure patient perspective (identified in the designation of patient-centred care as a domain of healthcare quality). Secondly, the IOM provide a framework for exploring the current meaning of quality of healthcare which is necessary to understand quality of hospital care from the patient perspective.

The IOM dimensions of quality were formulated through expert consensus at a ‘round table discussion’. The group was composed of clinicians and researchers with expertise in quality of healthcare, with no patient involvement (Personal Communication, 2012). Despite the fact that similar domains have been proposed by others, the STEEEP acronym has had international acceptance and use (Allen et al 2014, Haggerty et al 2007, Heenan et al 2010, Scottish Government 2010, Sipkoff 2004, Sofaer and Firminger 2005, Wiig et al 2014b). This is probably due to the fact that the IOM is a prestigious and powerful organisation, which is held in high regard.

Before the IOM dimensions are re-examined to establish their currency (Paper One), the following paragraphs provide an historical overview of key contributions to defining and understanding what constitutes quality of healthcare. Demonstrating the evolving nature of the concept of healthcare quality serves to highlight how definitions of quality of healthcare change over time as well as illuminating key influences on the IOM dimensions. Moreover, showing that understandings of healthcare quality are dependent on context and time highlights the need to re-examine the STEEEP dimensions, which were established over a decade ago, to ensure that the concept of healthcare quality used in this thesis is reflective of the current discourse of healthcare quality today.

### **2.3 The Evolving Definitions of Healthcare Quality**

Quality has a long history and heritage in healthcare – from the Hippocratic Oath of ‘doing no harm’ in the 4<sup>th</sup> Century B.C. to the work of Florence Nightingale in the 19<sup>th</sup> century on quality management and measurement (Meyer and Bishop 2007). Her contribution is discussed further below. Theoretical concepts focusing on healthcare quality have often emerged from definitions of quality in general industry and include Juran (1967), Pirsig (1974), Crosby (1979), Kano (1984), Deming (1986), Taguchi (1992)

and Chowdhury (2003). Many aspects of their theoretical contributions have influenced conceptualisations of modern healthcare today, for example Plan, Do, Study, Act cycles and the necessity to measure quality (Deming 1986). There have been fewer, although significant, contributors emerging directly from healthcare, including Donabedian (1980), Maxwell (1984), Ovretveit (1992), IOM (1990) and Blumenthal (1996). Their contributions are also discussed below.

Florence Nightingale (1820–1910) is known for her contribution to nursing, and is often referred to as “The Lady with the Lamp” due to the need for lamplight in night-time ‘care rounds.’ She is acknowledged for setting up the first formal training school for nurses in St Thomas’s Hospital, London in 1860 (Dingwall et al 1988). Less is known about her significant contribution in the field of quality of healthcare. It was during the Crimean War that she investigated many care processes, such as the procedure for washing linen and serving food, in an attempt to reduce mortality rates among soldiers due to infection. Campaigning to improve the standards of hospital care, she wrote to senior military figures requesting additional supplies and suggesting logistical changes to the supply chain. Nightingale was adept at mathematics and used statistical analyses to record and compare pre- and post-war infection rates. She created the Nightingale rose diagram, similar to the circular histogram used today, to present infection control and other data visually. Aspects of Nightingale’s work can be seen in healthcare quality today. For example, the Peripheral Vascular Cannula (PVC) bundle is a defined quality of care process used today, the reliability of which is established by regular measurement of implementation and audit of outcome; Florence would recognise the method, if not the equipment. Nightingale’s challenge, to improve and assure infection control in hospitals, remains a concern today. Incidents of low infection control standards hit the headlines at regular intervals, for example, the *Clostridium difficile* outbreaks at the Vale of Leven Hospital in Scotland (MacLean 2014). However, general standards are likely to have improved from the Nightingale era.

Although Nightingale did not offer an explicit definition of healthcare quality, aspects of her work highlight some of the STEEEP domains. For example, her meticulous infection control work could be aligned to the domain of safety. The domain of efficiency can also be seen in her unceasing efforts to improve the procedures for laundering bed linen and serving food. Similarly, her work on audit and measuring mortality rates could be reflective of the domain of effectiveness.

Yet aspects such as 'person-centred' and 'equitable' care are absent. For example, it is unlikely that the quality of 19<sup>th</sup> Century healthcare would have been questioned from the patient perspective as any care provided was valued and not necessarily expected. Prior to the National Health Service (NHS) formation in 1948, most healthcare was provided by family members, aides for the wealthy, or poor houses for those with neither (Dingwall et al 1988). For the army personnel in the Nightingale era, the hierarchical structure of the military and poor health would have prevented soldiers from articulating any concerns over hospital quality; rather, patients were likely grateful for any care given.

Interestingly, much of Nightingale's aspiration to improve quality appears to have been driven by humanitarian goals, as opposed to external drivers, such as policy, performance targets or cost. A similar moral drive can be seen in Deming's work. Deming's (1900–1993) work on quality was in the field of industry and his motivation appears to have been driven from the altruistic notion of improving conditions for the workforce. For example, he referred to poor management 'robbing' employees of their pride of workmanship; he is also remembered for his kindness and consideration for others (Deming 1986). This concept of intrinsic motivation is well recognised in those who work to understand and improve the quality of healthcare in present day (Parry 2014). Deming's theoretical work, referred to as the 'System of Profound Knowledge', promotes the concept that the quality of the system in which people work cannot be transformed or improved without individuals changing in ways that bring new meaning to their lives and interactions with others (Deming 1986). For Deming, perhaps the domains of quality were less important than the way in which they were operationalised in practice.

His influence can be seen in the present day, for instance, aspects of systems thinking are currently in use in healthcare quality. For example, past management of medication errors in hospital would most likely have resulted in the individual, usually the nurse, being punished in some way, for example, being sent for retraining with future promotion prospects damaged. Today, nurses who inadvertently give a patient a wrong medication are more likely to be involved in a root cause analysis, helping to explore aspects of the system which could be improved to reduce the likelihood of the same error reoccurring. For example, storing similarly packaged items separately. The focus has shifted to improving the quality of the 'system,' as opposed to blaming the 'individual,' except where deliberate harm is suspected (Reason 2000).

Avedis Donabedian was a physician and healthcare researcher who contributed to healthcare quality research in the 1950s and early 1960s. His work built on Deming's view of systems thinking. The theory of systems thinking acknowledges the complexity of the systems in which we live and work, explaining the interrelationship and effect one part of a system has on another part, and how systems interact (Laszlo 1991, Plesk and Greenhalgh 2001). Donabedian devised a model of quality; claiming that structure and process equalled outcome(s) (Donabedian 1980). His work has largely informed the whole systems approach used in healthcare quality measurement plans today. In later work he acknowledged limitations in a systems approach – "They are enabling mechanisms only. It is the ethical dimension of individuals that is essential to a system's success" (Mullen 2001, p. 140). Similar to Nightingale and Deming, Donabedian is thereby acknowledging the necessity of altruistic motivation to drive improvements in quality of care. The emotional engagement with each individual's deeply held beliefs is seen as necessary for the continued effort in quality improvement activity (Bate et al 2008, Robert et al 2011).

Donabedian (1998) also acknowledged the multi-dimensional nature of quality and suggested that quality could be divided into technical and interpersonal divisions, whilst acknowledging the interrelationship between them. For example, deciding on the most appropriate treatment (technical) for a patient is often dependent on how well the treatment options are explained (interpersonal) to the patient. He defines quality as an attribute of, and judgement upon, a process of care. His definition also depends, therefore, on who the judges are of the care. There is currently widespread acknowledgement that focusing on technical aspects alone will not improve quality (Wiig et al 2014b). The IOM domains capture technical aspects, that is to say, effectiveness and safety, as well as interpersonal aspects of quality, for example, person-centred care.

Robert Maxwell revisited the multi-dimensional nature of quality in healthcare throughout the 1980s and early 1990s. He expanded the understanding of quality in healthcare by proposing six domains of quality, namely; effectiveness, efficiency, access, equity and relevance (Maxwell 1984). Maxwell argued that these domains captured the multi-dimensional nature of quality when considered as a whole, rather than fragmented parts. Although the multi-dimensional nature of quality had been described in earlier work, Maxwell advanced understanding by articulating these domains and attempting to apply them to quality in an Intensive Care Unit, using Donabedian's model of structure, process and outcome (Maxwell 1992). There are many similarities between Maxwell's and the

IOM domains of healthcare quality, with both having domains of 'effectiveness' and 'efficiency', yet the IOM domains continued and continue to dominate.

In the 1990s there was an increase in public awareness of safety and quality issues following some high profile cases of systemic failures in health and social care, for example, the Bristol Heart Enquiry, the Alder Hay retention of organs scandal, and the death of Victoria Climbié (Smith 1998). The presumption that the doctor, or other healthcare professional, 'knows best' was being questioned. The long held privileged position of assumed quality in healthcare was under scrutiny. By the late 1990s the presumption of quality was no longer automatic, and professionals and health services saw the emergence of clinical governance (Johnston et al 2000). Clinical governance was an umbrella term used to describe a monitoring system which healthcare providers were required to have in place in order to assure quality (Scally and Donaldson 1998). Clinical audit became particularly popular as a means of measuring and reporting the quality of services and care. On occasion, audits were conducted by auditors and the process was more akin to scrutiny and judgement, rather than learning (Johnston et al 2000). Safety, and other aspects of quality amenable to measurement dominated the conceptualisation of healthcare quality during the 1990s.

It was at this time that Ovretveit defined quality care as that which was "fully meeting the needs of those who need the service most, at the lowest cost to the organisation, within limits and directives set by higher authorities and purchasers" (Ovretveit 1992, p. 2). This definition is reflective of the consumerist discourse in the early 1990s, where customer satisfaction versus cost featured heavily in healthcare. Ovretveit's definition highlights the importance of an equitable service (equity is an IOM dimension) and suggests that 'needs' are defined differently depending on which perspective is being considered; the client, the professional, or management. Yet, quality of hospital care remained defined by those delivering care, as opposed to those receiving care.

Blumenthal (1996) offered a definition of quality of healthcare from a clinician's perspective - "doing the right things right." His definition presumes the right thing is always known by the healthcare professional and his perspective was likely to have been influenced by the evidence-based medicine movement of the 1990s. His view could be aligned to the IOM domain of effectiveness, as quality appears to be possible when science or evidence is used to manage people's healthcare problems. There are challenges with this view as there are many clinical situations which make the use of evidence challenging, for example, treatment for a particular condition when the patient

has multiple co-morbidities. A more recent view would be that 'right' may be both transitory and negotiated (Greenhalgh et al 2004).

Finally, the IOM define quality as "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge" (IOM 2001, p. 4). This definition signifies the inclusiveness of populations as well as individuals and takes cognisance of public health and evidence-based medicine. There is an assumption here that the definition of quality evolves as professional knowledge advances. As previously mentioned, the IOM also articulated healthcare quality as six domains; safety, timeliness, effectiveness, efficiency, equity and person-centred care, commonly referred to as the STEEEP acronym (IOM 2001). Their domains remain prevalent and uncontested in healthcare today.

## **2.4 Summary**

There have been a few key contributions to conceptualising the meaning of quality of healthcare, many of which have been influenced by definitions of quality from general industry. Whilst these meanings vary, there is consensus that quality of healthcare is complex and multi-dimensional and contingent upon which stakeholder is being asked to evaluate it (e.g. patient, healthcare professional, manager). There is also an acknowledgement that maintaining and improving the quality of healthcare requires altruism (i.e. caring), as well as technical expertise. The most widely accepted, and used, conception of healthcare quality is the IOM STEEEP acronym. This historical critique establishes that, as society evolves, so too does the conceptualisation and subsequent domains of healthcare quality. There is no reason to believe that such an evolution has halted.

The evolving nature of healthcare quality is important in the development of a measure of hospital quality of care from the patient perspective. If a measure is to remain valid (measuring what it purports to measure), there needs to be a clear process for checking the ongoing validity of the tool. The evolutionary nature of healthcare quality is an important factor to consider during instrument development, for example, designing key domains with prompts which can be adapted to suit context. This will be explored further in Chapter 5.

What constitutes quality is important as domains of quality are usually transformed into measurement plans at all levels of the healthcare system, and healthcare resources

aligned accordingly. Without a clear articulation of domains of quality there will be no consistency or validity in measuring the quality of hospital care from the patient perspective. There is a risk, then, that if current domains do not reflect or measure the healthcare quality, measurement becomes an end in itself and improving frontline care somehow disappears as the true target. It is essential, therefore, that the domains remain reflective of current definitions of quality. What constituted quality over a decade ago may not capture the concept today. There needs to be a re-exploration of the IOM domains to ensure they remain fit for purpose for healthcare today and in the future.

## **2.5 Aim and linkage to research question**

Before a measure of hospital quality of care from the patient perspective can be designed there needs to be a clear articulation of the concept of healthcare quality. There has been much criticism in the literature of instruments which are not derived from theoretical models, bringing into question their very foundation (Polit and Yang 2016, Strauss and Smith 2009). The first Paper in this thesis conceptualises the current meaning of quality of healthcare by conducting an integrative review of the literature about quality of healthcare. This contemporary conceptualisation of healthcare quality was used to develop a model of healthcare quality domains; as an essential step to devise an instrument measuring the quality of hospital care from the patient perspective. The content of an instrument needs to adequately reflect the construct of interest if it is to achieve its measurement purpose (Polit and Yang 2016). The Paper, therefore, lays the foundations necessary to develop an instrument measuring hospital quality of care from the patient perspective. The model is used later in Chapter 5 to inform the development of the Care Experience Feedback Improvement Tool (CEFIT).

### **Objective One:**

To devise a theoretical model of quality of healthcare informed by an integrative review of the literature.

### **Associated Publication**

**Beattie M**, Shepherd A, Howieson B. (2013) Do the Institute of Medicine's (IOM's) dimensions of quality capture the current meaning of quality of healthcare? – An integrative review. *Journal of Research in Nursing*, 18 (4), 288-304.

2.6 Paper One: Do the Institute of Medicines' (IOM) dimensions of quality capture the current meaning of quality in healthcare? – An integrative review

## Journal of Research in Nursing

<http://jrn.sagepub.com/>

---

**Do the Institute of Medicine's (IOM's) dimensions of quality capture the current meaning of quality in health care? – An integrative review**

Michelle Beattie, Ashley Shepherd and Brian Howieson

*Journal of Research in Nursing* published online 25 April 2012

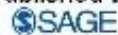
DOI: 10.1177/1744987112440568

The online version of this article can be found at:

<http://jrn.sagepub.com/content/early/2012/04/25/1744987112440568>

---

Published by:



<http://www.sagepublications.com>

Additional services and information for *Journal of Research in Nursing* can be found at:

**Email Alerts:** <http://jrn.sagepub.com/cgi/alerts>

**Subscriptions:** <http://jrn.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

>> **OnlineFirst Version of Record - Apr 25, 2012**

**What is This?**



# Do the Institute of Medicine's (IOM's) dimensions of quality capture the current meaning of quality in health care? – An integrative review

**Michelle Beattie**

University of Stirling, UK

**Ashley Shepherd**

School of Nursing, Midwifery and Health, University of Stirling, UK

**Brian Howieson**

Institute for People-Centred Healthcare Management, Stirling Management School, University of Stirling, UK

## Abstract

**Aims:** The aim of this study was to determine whether the widely adopted Institute of Medicine's dimensions of quality capture the current meaning of quality in health care literature.

**Design:** An integrative review was utilised as there has been a multitude of published papers defining quality in relation to health care, therefore collective analysis may provide new insight and understanding.

**Method:** Papers offering a definition or conceptual understanding of quality in relation to health care were identified by searching relevant databases. Papers were excluded according to predefined criteria. An integrative review was conducted and the Institute of Medicine's dimensions were used as a framework for data extraction and analysis.

**Findings:** The review identified two important additional dimensions of quality; namely *caring* and *navigating the health care system* and argues that they require recognition as dimensions in their own right.

**Conclusion:** In the current climate of constrained finances there is a risk that the allocation of resources is directed to current explicit dimensions to the detriment of others. The result may be a reduction in health care quality, rather than improvement.

---

## Corresponding author:

Michelle Beattie, School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Centre for Health Science, Old Perth Road, Inverness IV2 3JH, UK

Email: michelle.beattie@stir.ac.uk

**Keywords**

Clinical governance, compassionate care, health and social care policy, integrative review, organisation and service

**Introduction**

The challenge of ensuring quality of health care remains high on the public and political agenda internationally (Institute of Medicine, 2001; Department of Health, 2008; Scottish Government, 2010). Despite a growing need to improve the quality of health care, there is still a plurality of perspectives of the actual meaning of quality. In order to determine appropriate measures for health care quality improvement, there needs to be a shared understanding of the elusive concept. A common understanding would also enable clarity for teaching and research within the field.

Historically, there have been many key contributions to understanding the meaning of quality in health care. It has been 150 years since Florence Nightingale advocated that caring attitudes and behaviour were fundamental to quality of care (Meyer and Bishop, 2007). Donabedian (1980) suggested that quality can be separated into technical and interpersonal divisions, whilst acknowledging the interrelationship of both. Maxwell (1984) provided dimensions of quality as accessibility, relevance, effectiveness, equity, acceptability and efficiency, which continue to resonate with current thinking, whilst Ovretveit (1992) defined quality as 'fully meeting the needs of those who need the service most, at the lowest cost to the organisation, within limits and directives set by higher authorities and purchasers'. This definition reflects the consumerist discourse in the early 1990s, when customer satisfaction versus cost featured heavily in health care. Blumenthal's (1996) opinion was that quality meant 'doing the right things right', which it is likely was influenced by the evidence-based medicine movement.

Recently, the Institute of Medicine (IOM) in America has made a considerable contribution to the understanding of quality in health care. The IOM is a non-profit organisation which aims to provide government and private industries with non-biased information about health care. They produced a seminal report in 2001 *Crossing the Quality Chasm* which made specific recommendations to enable improvement in health care quality for all Americans. This key text also conceptualised quality as six dimensions: *safety, timeliness, effectiveness, efficiency, equity and patient centredness* sometimes referred to as the STEEP acronym (IOM, 2001). Despite the American context, these dimensions have been accepted internationally and appear in policy context world wide (Sipkoff, 2004; Sofaer & Firminger, 2005; Haggerty et al., 2007; Heenan et al., 2010, Scottish Government, 2010). The IOM dimensions of quality probably remain seminal due to the accolade and contribution the organisation has made to the field of health care quality particularly as, to date, there have been no further significant contributions to the evidence from such a credible resource. The IOM are independent advisors to the American Government and have subsequently influenced health care policy internationally.

Whilst these key contributions provide an insightful understanding of quality in health care, it is apparent that the meaning of quality is influenced by the discourse of society. What quality meant even 10 years ago would probably not have captured the true meaning of quality in health care today. There has been an increase in public expectations of health care,

changing demographics, the additional risks inherent with new technologies, as well as the need to deliver health care with fewer resources.

Given the current emphasis of person-centred and mutual health care, quality as perceived by the patient/client is of timely significance. Whilst patient satisfaction has been a major component of quality of care, there has been a re-focus from the outcome of patient satisfaction or dissatisfaction to understanding patient expectations and experiences (Sixma et al., 1998). Recipients of health care services are more likely to expect quality from many perspectives, driven by their changing needs. For example, an acutely unwell patient may rate the dimension of effectiveness highly, but during convalescence may rate person-centredness as the most important dimension of health care quality. The current discourse of health care from the public and professional perspective has almost certainly influenced the elusive meaning of quality. There needs to be some assurance that the plurality of quality has been captured in the IOM dimensions.

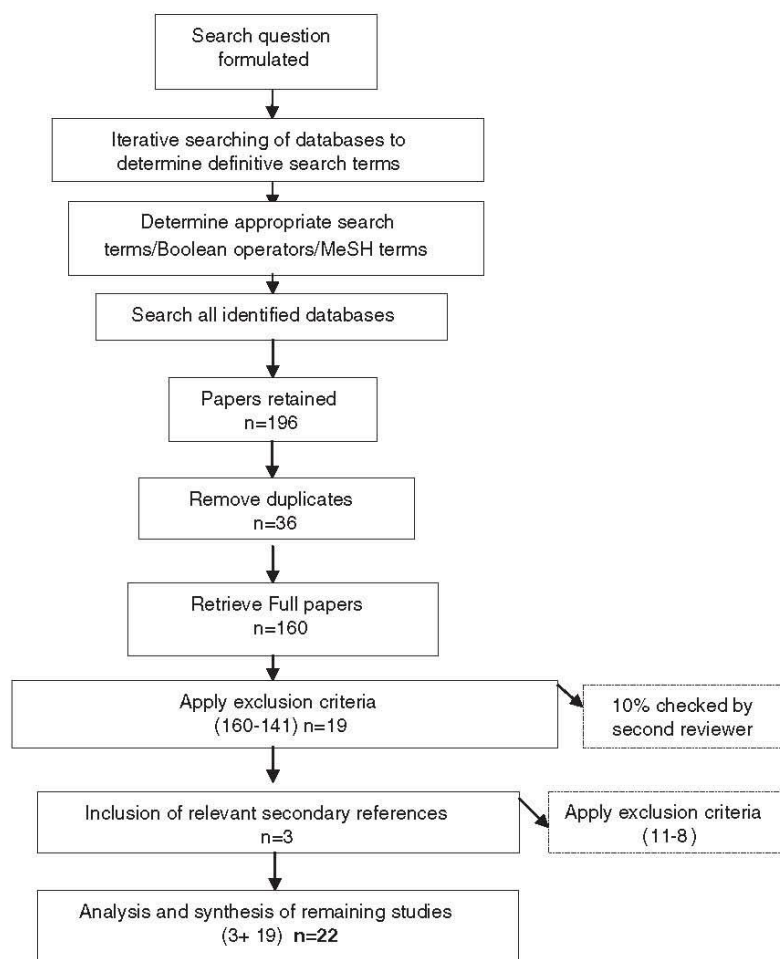
## Methodology

This review aims to determine whether the highly regarded IOM dimensions (IOM, 2001) capture the *current* meaning of quality in health care. An integrative review was selected to determine how the concept of quality was being defined in contemporary (previous 10 years) literature. An integrative review enables synthesis and reinterpretation of specific concepts or content. Importantly for this study, this methodology enables the integration of both theoretical and empirical literature (Whittemore and Knafl, 2005). The aim of the integrative review is to abstract new findings or phenomena from an original starting point. There were multiple publications in relation to defining quality in health care, and therefore it was important to determine the collective contribution.

As the review was specifically focused around the IOM's six dimensions of quality, these were used to extract the themes from the literature. Additional data was also extracted and integrated into themes. A flow diagram (see Figure 1) provides an overview of the review process. The flow diagram was adapted from the preferred reporting items for systematic reviews and meta-analysis (PRISMA) standards (Moher et al., 2009).

Medical Literature Analysis and Retrieval System Online (MEDLINE) and Current Index to Nursing and Allied Health Literature (CINAHL) databases were searched. Exploring the layout of Medical Index Subject Headings (MeSH) terms within each search engine and observing results enabled a specific search strategy to be formulated. Given that the research question aims to clarify the current understanding of quality in health care and to ensure the study is feasible, the search was inclusive of papers from the years 2000 to 2010. Full details of the search strategy are detailed in Table 1. In total 196 papers were yielded from the search strategy, and the removal of duplicate papers resulted in the retrieval of 160 papers.

Inclusion and exclusion criteria were predetermined to target relevant papers and reduce bias. Papers were included if the main focus of the paper was *quality* in relation to health care as defined or utilised by the authors or study participants of the papers. The population was specifically in relation to patient, service user or any other term used to describe those accessing or providing health care. Exclusion criteria were classified into subject, population and context (see Table 2). Subject papers in relation to specific treatment or disease processes were excluded as they often detailed specific care or treatment pathways relevant to a condition, which did not highlight a specific dimension or conceptual definition of quality, for example the paper by Castilla et al. (2008). Population exclusions included animal,



**Figure 1.** Procedural flowchart: modified review.  
MeSH: Medical Index Subject Headings.

in-vitro or laboratory, for example the Sirota (2006) paper on error in anatomic pathology, as they examined quality of testing and procedures rather than quality of care provision. Papers on the context of Eastern health care, such as Hyder (2002), were excluded as there are significant variations in culture and health care service when compared with western health care.

The exclusions were applied to all 160 papers, and those papers which were not rejected following application of the exclusion criteria were retained ( $n=19$ ). Ten percent of all papers were reviewed independently by a second reviewer (AS) using the predetermined exclusion criteria. The 19 papers for duplicate review were selected using the random function within Microsoft Excel™. Both reviewers independently decided whether the paper should be retained or rejected, and reached the same decision on 18 out of

**Table 1.** Search strategy

Database	MeSH terms	Result
Search one Ovid MEDLINE(R) 1950 to July Week 2 2010	*Quality of Health Care/or *Quality Assurance, Health Care/or *Quality Indicators, Health Care/  defin*.ti.	48398   36937
Search two EBSCO Host CINAHLwithFullText Accessed July Week 3 2010	Combine 1 and 2 Limit to English Language and yr= 2000-Current (MM "Quality Assessment") or (MM "Quality Assurance") or (MM "Quality Improvement") or (MM "Quality Management, Organizational") or (MM "Quality of Care Research") or (MM "Quality of Health Care") or (MM "Quality of Nursing Care") or (MM "Clinical Indicators") TI defin*	259 119 30369  5686
	Combine both above Limit to English Language and yr= "2000-Current"	125 77
	TOTAL	196

MeSH: Medical Index Subject Headings; MEDLINE: Medical Literature Analysis and Retrieval System Online; EBSCO is a publishing company who host other search engines.

19 occasions (95%). Consensus was reached on the remaining paper following discussion between both reviewers. Eleven papers were retrieved from secondary references of retained papers. Following application of exclusion criteria eight were rejected and three retained. Twenty two papers in total were included for further analysis.

Studies were graded according to the hierarchy of evidence developed by the National Institute for Health and Clinical Excellence (NICE) Framework (NICE 2006). The NICE (2006) hierarchy ranks studies from high-quality meta-analyses, systematic reviews of randomised controlled trials (RCT), or RCT with a very low risk of bias (graded as 1++) to expert opinion or formal consensus (graded as a 4). A second reviewer (AS) also independently determined the levels of evidence for 50% of the papers. The papers generally scored between -2 and 4 on the hierarchy reflecting the qualitative nature of the papers and the inclusion of guidance and opinion papers. The integrative review aimed to capture conceptual definitions or dimensions of quality, and therefore these papers provided insightful narratives. It should also be acknowledged that whilst guidance and opinion papers were retained, these were usually written by experts in the field who review the evidence prior to developing guidance.

To aid data extraction, characteristics of retained papers were extracted for location, year, population or perspective studied, number of participants and the context (see Table 3). To aid the synthesis, all definitions or attributes of quality within retained papers were mapped to the Institute of Medicine's six dimensions of quality (IOM, 2001). Where papers used words or phrases which clearly described one of the dimensions this was accepted as containing the appropriate IOM dimension. An extraction table was formulated with headed sections for each of the IOM dimensions (see Table 4). The dimension/s utilised by each of the papers were acknowledged by ticking the corresponding box for

**Table 2.** Exclusion criteria

Exclusion	Criteria for rejection	Total numbers excluded
Subject	<ul style="list-style-type: none"> <li>• Performance management/improvement, quality improvement methodology, service re-design, clinical pathways or indicators</li> <li>• Risks/interventions specific to disease/illness/procedure/diagnosis or prognosis</li> <li>• Irrelevant papers not defining quality of health care</li> </ul>	123
Population	<ul style="list-style-type: none"> <li>• Animal</li> <li>• in-vitro or laboratory</li> </ul>	4
Context	<ul style="list-style-type: none"> <li>• Eastern health care – namely Asia, India sub continental, Far East, Middle East, Near East</li> <li>• Dentistry</li> <li>• Nursing home or residential care</li> <li>• End of life or terminal care</li> <li>• Social care</li> </ul>	14
		141

Note: Full details of papers excluded are available from the author.

that dimension. For example Haggerty et al. (2007) identified an attribute of quality as 'Technical quality of clinical care: the degree to which clinical procedures reflect current research evidence and/or meet commonly accepted standards for technical content or skill'. This attribute was easily aligned to the dimension of effectiveness. Where the mapping was less obvious, or indeed did not fit the IOM's dimension, the words or phrases were listed under the 'other' section within Table 4. This reduced the risk of potential misinterpretation. For example Larrabee and Bolden (2001) utilised phrases and words such as 'treating me pleasantly' and 'caring'. Although these aspects could reflect aspects of person-centredness, this dimension did not capture the wholeness of these words and phrases.

All words and phrases captured under 'other' were later categorised using thematic analysis, which led to the creation of two additional dimensions. This analysis was conducted by hand by having the individual word or phrases on individual pieces of paper. Individual words and phrases were then scanned for similar themes before being grouped into the additional dimensions identified. For example, terms such as 'courtesy', 'emotional support', 'holism', 'treating me pleasantly' and 'empathy' were themed under an additional dimension of *caring*, whereas ideas such as 'continuity of care', 'accessibility', 'availability', 'flexibility', 'seamless transitions' and 'co-ordinated' were themed under an additional dimension of *navigating the system*.

## Results

Twenty-two papers were included for analysis, offering either an explicit or implicit definition or conception of quality. The review identified two important additional dimensions of quality categorised as *caring* and *navigating the health care system*. All IOM dimensions were prevalent in the literature, but not necessarily sufficient to capture the wholeness of quality in current health care. Patient-centredness and effectiveness were

Table 3. Data extraction

Author	Location/year	Evidence level and design	Population or perspective	Participants	Context	Explicit definition of quality
Attree (2001)	England (2001)	2 Grounded theory	Patients and relatives	34 patients and 7 relatives	Acute medical	No
Baker (2001)	England (2001)	4 Expert opinion	Manager	Not applicable	Healthcare generally	Yes
Barelds et al. (2009a)	Netherlands (2009a)	3 Focus groups	Service users, parents or relatives	21 parents or relatives	Intellectual disabilities service	No
Barelds et al. (2009b)	Netherlands (2009b)	3 Literature review	Service users, parents or relatives	Not applicable	Intellectual disabilities service	No
Bassett (2010)	England (2010)	4 Discussion	Nursing	Not applicable	Healthcare generally	No
Brook et al. (2000)	USA (2000)	4 Discussion	Medical/Researcher	Not applicable	Healthcare generally	No
Campbell et al. (2000)	England (2000)	3/4 Expert opinion	Researcher	Not applicable	Healthcare generally	Yes
Chilgren (2008)	USA (2008)	4 Discussion	Manager	Not applicable	Healthcare generally	Yes
English (2002)	USA (2002)	4 Discussion	Medical	Not applicable	Primary care	Yes
Frist (2000)	USA (2000)	4 Discussion	Medical	Not applicable	Healthcare generally	Yes
Haggerty et al. (2007)	Canada (2007)	3 Delphi	Clinicians, academics and decision-makers	20 experts	Primary Care	No
Heenan et al. (2010)	USA (2010)	3/4 Expert opinion	Quality committee and board membership	Not stipulated	In hospital	Yes
Hickman (2001)	USA (2001)	4 Expert opinion	Quality manager	Not applicable	Healthcare generally	Yes
Howie et al. (2004)	Scotland (2004)	4 Expert opinion	Medical	Not applicable	Primary care	No
Jones (2010)	England (2010)	4 Expert opinion	Medical	Not applicable	In hospital	No
Larrabee and Bolden (2001)	USA (2001)	3 Content analysis	Adult patients	199 patients	In hospital	No
Manning (2006)	USA (2006)	4 Expert opinion	Medical	Not applicable	Healthcare generally	No
O'Reilly (2007)	Ireland (2007)	-2 Focus groups	Staff and parents	Not stipulated	Physical disabilities service	Yes
Russell (2007)	USA (2007)	4 Expert opinion	Medical	Not applicable	Primary care	No
Sipkoff (2004)	USA (2004)	4 Expert opinion	Medical	Not applicable	Healthcare generally	Yes
Sofaer and Firminger (2005)	USA (2005)	4 Expert opinion	Patient perceptions	Not applicable	Healthcare generally	No
Williams (2000)	USA (2000)	4 Expert opinion	Medical	None	Healthcare generally	Yes

**Table 4.** Quality attributes or dimensions

Author	Effectiveness	Timeliness	Safety	Efficiency	Equitable	Patient-centred	Others
Attree (2001)		X				X	Nature of care provided, anticipation of need, involvement and anticipation Nature of the relationship Humanistically caring relationships, involvement, commitment and concern Adherence to implicit and explicit codes Access to support Provision – food Friendship Reliable Support Continuity Co-operation Bureaucracy Capacities/motivation care providers Continuity of care Accessibility Availability Flexibility Seamless transitions Excellence the first time every time Co-ordinated Informed choices consistent with their values Accessible Attitude and body language of the provider Dissemination of information
Baker (2001)	X					X	
Barelds et al. (2009a)	X	X				X	
Barelds et al. (2009b)						X	
Bassett (2010)	X		X			X	
Brook et al. (2000)	X			X		X	
Campbell et al. (2000)	X			X	X	X	
Chilgren (2008)		X				X	
English (2002)	X	X				X	
Frist (2000)	X					X	

(continued)



Table 4. Continued

Author	Effectiveness	Timeliness	Safety	Efficiency	Equitable	Patient-centred	Others
Haggerty (2007)	X	X		X	X	X	Continuity, family centredness, advocacy, cultural sensitivity Accessible, interpersonal communication, community orientation, comprehensiveness of services, team, responsiveness, integration Kind
Heenan et al. (2010)	X		X			X	
Hickman (2001)	X		X	X		X	
Howie et al. (2004)							Inter-personal effectiveness, as well as technical/clinical effectiveness, Holism, Patient-centredness divided into doctors behaviour and patient participation
Jones (2010)	X		X			X	Range of services Independent audit – transparency Treating me pleasantly, caring
Larrabee and Bolden (2001)	X	X				X	
Manning (2006)	X			X	X	X	
O'Reilly (2007)	X	X	X	X	X	X	Empathy
Russell (2007)	X		X	X	X	X	Co-ordination of care, improved access
Sipkoff (2004)	X	X	X	X	X	X	Access
Sofaer and Firminger (2005)	X			X		X	Communication and information Courtesy, emotional support
Williams (2000)	X					X	

most commonly used as descriptors of quality, whereas the attributes of equity and safety were least used as descriptors of quality. The synthesis of results in relation to each of the IOM dimensions and two additional dimensions are presented below from most to least prevalent.

### *Prevalence of dimensions*

**Patient-centredness.** The IOM explains patient-centred care as care that is respectful of an individual's preferences, needs and values and incorporates the notion of 'nothing about me without me' (IOM, 2001). Patient-centred care was the only dimension to be captured in all of the papers. The concept of patient-centredness is derived from the notion of mutuality in health care where the patient and practitioner work together to attain the best health outcome for the patient. Many of the papers included terms that could be interpreted as patient-centred, although this is dependent on the definition of person-centredness. For example if the dimension is viewed as *nothing about me without me* this suggests that the patient is involved in their own health care decisions in partnership with the health care provider. Although a relationship between both providers and recipients of health care is recognised, the value and nature of the relationship is not explicit. For example a patient in Attree's (2001) study commented on how the nurse demonstrated compassion by holding her hand and not walking off whilst still talking. There were multiple examples in the literature which alluded to the essence of caring behaviours not necessarily captured in the dimensions of person-centredness as defined by the IOM. Those aspects not relating to shared decision-making were categorised in the 'other' section of Table 4, which required the creation of an additional dimension of caring to ensure they were explicitly represented.

**Effectiveness.** Effectiveness appeared frequently in the retained papers (20 out of 22). As the IOM defines effectiveness as matching science to care, this dimension is closely linked to the adoption in western health care of evidence-based medicine (IOM 2001). Of the papers which did not clearly articulate this dimension, both considered quality from the perspective of service users or their relatives. The majority of papers which identified effectiveness as an important dimension explained the dimension from a technical or scientific standpoint, which fits well with the IOM's description. However Howie et al. (2004) divided effectiveness into technical and interpersonal domains, suggesting that effective communication was as important as clinical and technical competence. This division is reflective of Donabedian's earlier work of technical and interpersonal divisions (Donabedian 1980). Interpersonal attributes were commonly identified as key components of health care quality, and these will be discussed in the *other* section.

**Efficiency.** The IOM defines efficiency as care that is not wasteful in terms of duplication of effort and unnecessary treatment, but also includes making full use of all resources, such as enabling staff to be innovative (IOM 2001). Efficiency was identified as an attribute of quality in under half of the papers. The majority of these papers were published between 2000 and 2005, which may reflect the consumerist drive in UK health care at this time, i.e. waiting list targets. The dimension of efficiency appeared randomly in the data and was not more or less prevalent according to location, perspective, participant or context. There may well be an assumption that the health care focus on quality improvement is predominantly in relation to efficiency, as it is in general industry.

*Timeliness.* This dimension was primarily concerned with the avoidance of unnecessary delays (IOM, 2001). Timeliness was identified in approximately one-third of the retained papers. The interpretation by the IOM is in relation to reducing unnecessary waiting or delay within the health care system, such as waiting for surgery etc. However, much of the timeliness identified within the retained papers was in relation to processes of care and specifically about aspects of health workers' behaviour within the system. For example, the study by Attree (2001) in relation to patients' and relatives' perceptions of quality gave examples which identified timeliness in relation to the behaviour of nurses who 'came when called, came back when they said and were there when needed/wanted', or in relation to not so good care stated 'they're too busy'. Within the community context, timeliness was associated with reliable behaviour of care providers, such as 'keeping appointments' as well as giving 'time and attention' (Barelds et al., 2009a). Safety of those papers which identified safety as a dimension of quality, only one paper was from a service user perspective (O'Reilly, 2007). The IOM suggests that safety is causing no harm by care that is intended to help (IOM, 2001). Systems theory from high reliability organisations suggests that reliable, standardised care will reduce error, resulting in safer health care systems. O'Reilly (2007) identified a 'reliability dimension focusing on the services ability to provide the service accurately and dependably'. The aspect of reliability could easily be integrated into other dimensions however, such as timeliness and effectiveness.

*Equity.* This dimension was reported in only six of the retained papers. The IOM describes equity as closing the gap between justice and health care, in which care should not be influenced by individuals' personal characteristics (IOM, 2001). Interestingly, only two papers from the UK identified this dimension, the rest were from the USA. These differences are likely to be reflective of the different health care systems in the UK and USA. Issues of access were raised in the retained papers and some would argue that these could be categorised under the dimension of equity. Often the system appears to only be accessible to people who have the necessary skills and abilities to articulate their health needs. However, there were other phrases extracted from the retained papers which were more akin to challenges once the system had been accessed, and were more readily categorised under the additional heading of system navigation. Interestingly Maxwell (1984) also identified access as an important dimension of health care quality, which the IOM have subsumed under the category of equity.

#### *Additional dimensions identified*

*Caring.* Concepts of caring were extracted from many papers and were often integrated within other dimensions. For example timeliness was often discussed in relation to the behaviour of the health care provider, rather than the system. The nature of the relationship between those accessing health care and those providing care was a recurring theme in the papers. For example, behavioural aspects of those providing care were identified as demonstrating the ability to anticipate needs, displaying empathy and concern, treating patients pleasantly and with courtesy. Service users seemed acutely aware of whether or not the care was given in a compassionate manner by identifying characteristics of the care giver's attitude and body language. It could be argued that these attributes are components of patient-centred care, although the IOM description appears to reflect patient involvement rather than the attitudes and behaviours displayed

by the care provider. By subsuming these under the dimension of patient-centredness there is a risk that these important aspects of quality could become less explicit. Caring remains fundamental to health care provision. Patients, service users, or indeed people in general, still want those who work in health care to provide the 'art' as well as science. Although aspects of holistic and intuitive care are difficult to measure, they remain the foundations on which service users perceive health care quality.

Navigating the system. Accessibility was highlighted in several of the retained papers as an important attribute of quality. Although accessibility could be integrated under the IOM's dimension of equity, this dimension did not accurately capture the meaning as communicated in the papers. Interpretation of the equity dimension from the IOM's perspective is more focused around equal service provision for vulnerable groups, rather than specific challenges of accessing and finding ways round complex health care systems. For example, in the studies by Barelds et al. (2009a,b) parents and carers identified hurdles in accessing appropriate services for their child or relative with intellectual disabilities. These challenges appeared to be exacerbated at times of transition, such as when the child reached school age or moved from child to adult services. The importance of the patient/service user journey is well documented, yet it is not entirely captured in the IOM's dimensions of quality. There are also issues relating to the availability of services, which will continue to challenge services in the current financial climate. These issues were not accurately, nor explicitly, captured within the IOM's dimension of equity. The ability to navigate the health care system is an essential quality dimension. If the health service cannot be accessed and navigated through then it would be impossible to measure other dimensions, such as effectiveness. Health care systems need to be designed to ensure individuals are empowered to access services, and routes through various health care journeys are seamless.

Another important aspect of the system identified was in relation to how care was co-ordinated. Co-ordination was relevant on all levels from the individual working within a team, as well as the inter-relationship within and between service provisions. Part of the integration of services identified was in relation to the responsiveness of the team and system when changes or transitions occurred. How well or otherwise components of the system interact remains critical, and the issue of speciality silos remains a challenge. Closely associated with co-ordination was the attribute of continuity, which was expressed in the literature with regard to the relationship between user and provider, rather than the system in which they worked. For example, continuity was expressed as seeing the same person repeatedly and having the opportunity to forge relationships with those providing care.

## Discussion

Whilst the IOM dimensions of quality have offered some mutual understanding of quality in health care, it remains important to consider these dimensions critically, particularly in an economically challenging time. This review of topical literature has identified two important dimensions of *caring* and *navigating the system*, which risk being marginalised if not made explicit. Whilst the findings need to be interpreted with caution due to the empirical limitations of an integrative review, the paper offers a critical view of the widely accepted dimensions of health care. The high number of publications within the field of health care quality posed significant challenges to the study. Pragmatic restrictions were necessary whilst formulating the search strategy to ensure the study was feasible, which increases the

potential that important contributions have not been captured. Despite these limitations, the search strategy did yield very relevant results and tracing secondary references identified only an additional three papers. During the application of exclusion criteria the risk of bias was reduced by having a second reviewer independently apply the exclusion criteria. Ten percent of papers were checked by a second reviewer due to limitations of time and resources; however, both reviewers concurred on 95% of occasions.

The retained papers were fairly distributed across the USA, UK and other western countries. However, given the geographical size and population of the USA it could be implied that papers exploring the concept of quality are more prevalent within the UK. More papers were evident from the UK between 2006 and 2010, which is indicative of the current discourse in quality and quality improvement within the UK. Retained papers generally scored low on the hierarchy of evidence, which identified the limitations of applying hierarchies to non-intervention questions. This finding also highlighted the fact that conceptual definitions or understandings of quality have, to date, been derived mainly from expert opinion rather than from a sound evidence base.

The fact that all papers made an implicit or explicit reference to the dimension of patient-centredness reflects the current patient/public involvement and mutual discourse in health care over the last 10 years. This dimension is the first quality ambition within the NHS Scotland Quality Strategy (Scottish Government, 2010). The Scottish strategy has coined the dimension 'person-centred', as opposed to 'patient-centred'. The change in terminology may be a deliberate attempt to deflect the negative connotations sometimes associated with patients or service users. Indeed, 'person' and 'personhood' seem more appropriately aligned with the premise of valuing individuals' needs and preferences with the aim of creating mutually beneficial relationships between service user and service provider. It could also be viewed that 'person-centred' is a more inclusive term, which encompasses the value of those who deliver health care, as well as those who access services.

Effectiveness was the second most reported dimension of quality. This was unsurprising, as quality is often measured in terms of outcomes of treatment. Interestingly, effectiveness was more likely to be recognised as a dimension by those providing services than those receiving services. This may be due to the general public's assumption that the health care they receive is indeed effective. This finding and assumption can also be applied to the dimension of safety, which was underreported from a patient/client perspective.

The additional dimension of caring was a predominant theme identified in the literature. Although many facets of this concept could arguably be integrated into the current IOM dimensions, there is a real risk that the art of health care will be lost in its submersion in science. There were tangible aspects of health care practitioners' behaviour which were identified as imperative to the quality of care received. The Healthcare Quality Strategy for Scotland (Scottish Government, 2010) has captured 'caring' under the auspices of the seven Cs – namely *caring, compassion, communication, collaboration, clean, continuity and clinical excellence*. Many of these facets of caring were evident within the IOM's *Crossing the Quality Chasm*, although not recognised as an explicit dimension (IOM, 2001). There are, however, real benefits to ensuring the caring dimension is made more explicit. If caring remains subsumed under the current dimension of patient-centredness there is a risk that imperative components of behaviour, attitudes and therapeutic relationships will be marginalised.

The dimension of navigating the system is also integral to health care quality. There is recognition that health care systems are increasingly complex and this was reflected by

patients, service users and service providers in the retained papers. The ability to move through the complexity of inter-related but not connected systems remains a challenge and a threat to health care quality. There are worldwide efforts to streamline complex health systems. If the dimension of health care navigation is not recognised, resources will probably be focused on achieving other dimensions. If the systems cannot be accessed or the patient's journey is slowed by bottlenecks or other system challenges, the quality of health care will continue to be impeded.

## Conclusion

This integrative review identified the inclusive nature of the IOM's dimensions of quality and suggests an additional two dimensions should be considered for an inclusive definition of quality in health care. There is a clear balance to be found between necessity and sufficiency. There is also an acknowledgement of the plurality of quality. Interpretations will continue to be contested. As Donabedian (1980) identified, definitions of quality will continue to be dependent upon who the judges are and the context in which the enquiry is placed. What can and should be attainable is the sharing of our mental models of quality – untangling espoused mental models to examine and challenge their ideological assumptions. By challenging different assumptions of quality and sharing these mental models, health care professionals and the public can work together to achieve common goals. This would enable varying public and professional perspectives to be considered when planning improvement initiatives and measurement plans in health care.

Although the review aimed to explore the current understanding of quality in relation to the IOM's dimensions, it would appear that the concept of quality bears similarities to that of Florence Nightingale's work on quality 150 years ago, where caring attitudes and behaviour were identified as fundamental to quality care (Meyer and Bishop, 2009). This integrative review has reinforced the fact that caring remains central to quality health care – ensuring the science does not hide the art is imperative. This finding is a reminder that despite the passing of years, an essential need of people as individuals in health care is the establishment of caring relationships. The real challenge now is the development of care measures that accurately capture the elusive dimension of caring and how, if at all, the health care system can be re-designed to ensure accessibility and streamlined movements of those entering and navigating the system.

Caring and system navigation are key elements to health care quality and need to be visible in policy and corporate, as well as clinical, decision making. If these fundamentals are ignored there could be serious consequences for services and recipients of health care, especially as the allocation of resources becomes more restrained. Those dimensions of quality that are easier to measure should not take precedence over difficult to measure, but essential, components of health care quality. Marginalisation of *caring* and *system navigation* could have a negative impact on the very aspect we are trying to improve – health care quality. Maxwell (1984) acknowledges the multidimensional nature of quality – considering the different dimensions of quality for analytical purposes should not disguise its true nature. Quality exists as wholeness, not as fragmented parts. We need to ensure we have captured all the dimensions of quality to ensure the whole is complete.

### Key points

- There has been international acceptance of the Institute of Medicine's quality dimensions; namely *safety, timeliness, effectiveness, efficiency, equity* and *patient-centredness*.
- It is imperative that key attributes of quality are explicit to ensure they are not marginalised in the health care providers' priorities to improve health care.
- There are two essential dimensions of quality, namely *caring* and *navigating the health care system*, which are not explicit in the Institute of Medicine's current dimensions. Navigating the health care system encompasses not only accessing the system, but the ability to move seamlessly throughout the system.
- Given the plurality of definitions of quality it is difficult to balance dimensions which are necessary to ensure the multi-dimensional nature of quality is retained, whilst ensuring they are essential to the overall purpose.

### Acknowledgements

Many thanks go to Kathleen Irvine, Subject Librarian, University of Stirling, for her assistance and support during database searching.

### Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

### Conflict of interest

None declared.

### References

- Attree M (2001) Patients' relatives' experiences and perspectives of 'good' and 'not so good' quality care. *Journal of Advanced Nursing* 33(4): 456-460.
- Baker R (2001) Principles of quality improvement: Part one defining quality. *Journal of Clinical Governance* 9(2): 89-91.
- Barelds A, Van de Goor I, Van Heck G, et al. (2009a) Quality of care and service trajectories for people with intellectual disabilities: Defining the aspects of quality from the client's perspective. *Scandinavian Journal of Caring Science* 24(1): 164-174.
- Barelds A, Van de Goor I, Bos M, et al. (2009b) Care and some trajectories for people with intellectual disabilities: Defining course and quality determinants from the client's perspective. *Journal of Policy and Practice Intellectual Disabilities* 6: 163-172.
- Bassett S (2010) What defines good nursing? *Nursing Management (Harrow)* 16(10): 13.
- Blumenthal D (1996) Quality of health care Part 4: The origins of the quality of care debate. *New England Journal of Medicine* 335(15): 1146-1149.
- Braunstein J (2003) Defining quality diabetes care in the new health system. *Current Diabetes Report* 3(4): 269-271.
- Brook R (1973) Critical issues in the assessment of quality of care and their relationships to HMOs. *Journal of Medical Education* 48: 114-134.
- Brook R, McGlynn E and Shekelle P (2000) Defining and measuring quality of care: A perspective from US researchers. *International Journal for Quality in Health Care* 12(4): 281-295.
- Campbell S, Roland M and Buetow S (2000) Defining quality of care. *Social Science and Medicine* 51(11): 1611-1625.
- Castilla J, Hernandez J, Cabello Y, et al. (2008) Defining poor and optimum performance in an IVF programme. *Human Reproduction* 23(1): 85-90.

- Centre for Reviews and Dissemination (2009) *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. University of York. January: pp.10,24.
- Chilgren A (2008) Managers and the new definition of quality. *Journal of Healthcare Management* 53(4): 221-229.
- Collignon P, Looke D, Ferguson J, et al. (2002) Surveillance definitions for multi-resistant organisms (MROs). *Australian Infection Control* 7(3): i-iv.
- Deming E (1986) *Out of the Crisis*. Cambridge, MA: MIT Press.
- Department of Health (2008) *High Quality Care for All*. June. Gateway reference: 10106.
- Donabedian A (1980) *Explorations in Quality Assessment and Monitoring, Vol.1, The Definition of Quality and Approaches to its Assessment*. Ann Arbor, MI: Health Admin Press.
- English E (2002) Quality care: All want it; let's define it. *Minnesota Medicine* 85(5): 7-44.
- Frist B (2000) Letters to the editor...definition of "quality healthcare". *Journal for Healthcare Quality: Promoting Excellence in Healthcare* 22(1): 7.
- Haggerty J, Burge F, Levesque J, et al. (2007) Operational definitions of attributes of primary health care: Consensus among Canadian experts. *Annals of Family Medicine* 5(4): 336-344.
- Heenan M, Khan H and Binkley D (2010) From boardroom to bedside: How to define and measure hospital quality. *Healthcare Quarterly* 13(1): 55-60.
- Hickman K (2001) Defining quality: Stanford's Quality Management Program achieving quality. *Stanford Nurse* 23(1): 3.
- Howie J, Heaney D and Maxwell M (2004) Quality, core values and the general practice consultation: Issues of definition, measurement and delivery. *Family Practice* 21(4): 458-468.
- Hyder A (2002) Structural adjustment in health in Pakistan: Defining the questions. *International Journal of Epidemiology* 31(2): 509-514.
- Institute of Medicine (1990) *Medicare: A Strategy for Quality Assurance*. Washington DC: National Academy Press.
- Institute of Medicine (2001) *Crossing the Quality Chasm: A New Health System for 21st Century*. Washington DC: National Academy Press. Available at: [http://books.nap.edu/openbook.php?record\\_id=10027&page=R1](http://books.nap.edu/openbook.php?record_id=10027&page=R1) (accessed 10 November 2010).
- Jones A (2010) Defining quality. *Health Service Journal* 120(6200): 14.
- Larrabee J and Bolden L (2001) Defining patient-perceived quality of nursing care. *Journal of Nursing Care Quality* 16(1): 34-60.
- Manning B (2006) Wisconsin Medical Society embarks on campaign to define high-quality care. *Wisconsin Medical Journal* 105(6): 11.
- Maxwell RJ (1984) Quality assessment in health. *British Medical Journal* 288: 1470-1472.
- Meyer B and Bishop D (2007) Florence Nightingale: Nineteenth century apostle of quality. *Journal of Management History* 13(3): 240-254.
- Moher D, Liberati A, Tetzlaff J, et al. (2009) Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 6(7): e1000097.
- National Institute for Health and Clinical Excellence (2006) *Reviewing and Grading the Evidence in NICE Guideline Development*, Chapter 7. Available at: <http://www.nice.org.uk/niceMedia/pdf/2006GuidelinesManualChapter7.pdf> (accessed 26 January 2011).
- Nelson J, Puntillo K, Pronovost P, et al. (2010) In their own words: Patients and families define high quality palliative care in the intensive care unit. *Critical Care Medicine* 38(3): 808-818.
- O'Reilly P (2007) Involving service users in defining and evaluating the service quality of a disability service. *International Journal of Health Care Quality* 20(2-3): 116-129.
- Ovretveit J (1992) *Health Service Quality: An Introduction to Quality Methods for Health Services*. Oxford: Blackwell Scientific.
- Russell T (2007) From my perspective: The fragmented and labyrinthine schematics that define our current health care system regularly frustrate surgeons' efforts to deliver, and patients' attempts to receive, optimal care. *Bulletin of the American College of Surgeons* 92(11): 4-5.
- Scottish Government (2010) *The Healthcare Quality Strategy for NHS Scotland*, May. <http://www.scotland.gov.uk/Resource/Doc/311667/0098354.pdf> (accessed 15 October 2010).
- Sipkoff M (2004) The new consensus favouring IOM's definition of quality. *Managed Care* 13(6): 18-21.
- Sirota R (2006) Defining error in anatomic pathology. *Archives of Pathology and Laboratory Medicine* 103(5): 604-606.
- Sixma H, Kersters J, Campen C, et al. (1998) Quality of care from the patients' perspective: From theoretical concept to new measuring instrument. *Health Expectations* 1: 82-95. Available at: <http://onlinelibrary.wiley.com/doi/10.1046/j.1369-6513.1998.00004.x/pdf> (accessed 26 January 2011).
- Sofaer S and Firminger K (2005) Patient perceptions of the quality of health services. *Annual Review Public Health* 26: 513-559.
- Whittemore R and Knaf K (2005) The integrative review: An updated methodology. *Journal of Advanced Nursing* 52(5): 546-553.
- Williams J (2000) Information defines quality. *Texas Medicine* 96(10): 9-11.

**Michelle Beattie's** current research and educational interests are captured in the theme of health care quality improvement. Her interest in improving health care quality has been built on a diverse clinical background, project management experience aiming to reduce error in clinical skills and a co-ordinating role for safety, governance and risk within the UK health service. As a lecturer she is now considering the meaning of quality and the implications the definition has for policy and practice, as well as considering whether or not we are measuring the right aspects of quality by conducting multiple small research studies. She is also module



co-ordinator for the MSc Module: Quality Improvement Projects in Health and Social Care, and supervises students who are undertaking a Quality Improvement Dissertation.

Email: michelle.beattie@stir.ac.uk

**Ashley Shepherd** is a lecturer in the School of Nursing, Midwifery and Health. Her current research interests include leadership in nursing, maternal and child health and quality improvement in health care. She has published over 40 peer-reviewed research papers and supervises PhD and Clinical Doctorate students in a variety of fields. She currently co-ordinates two quality improvement modules in the Masters in Advances Practice. Email: ashley.shepherd@stir.ac.uk

**Brian Howieson** is a Senior Foundation for Management Education Fellow at Stirling Management School, the University of Stirling. His previous role was Head of Training, Education and Professional Development for the Royal College of Physicians and Surgeons. Prior to this post, he was a commissioned officer in the British Armed Forces for 19 years. His main intellectual and research interest is in leadership and executive education and he is a member of the Editorial Advisory Board of Human Resource Management International Digest. He is co-author of the St Andrews University MLitt 'Leadership in Organisations' module and is Academic Leadership Partner to the MSc Degree in Advanced Leadership Practice at the Edinburgh Institute. Brian is Programme Director for the NHS Scotland Management Training Scheme MSc Degree in Leadership and Healthcare Management, the University of Stirling MBA in Finance with Foreign Trade University in Vietnam, and the Royal College of Physicians and Surgeons MSc Degree in Leadership. He is a Council Member for the Foundation for Management Education and is a Non-Executive (Lay Member) of the Medical Directorate within NHS Scotland.

Email: brian.howieson@stir.ac.uk

## **2.7 Overview of Paper One**

The following paragraphs provide additional details of the research process undertaken for the integrative review not covered in the associated publication, due to the limitations of the journal's word count. This is followed by a critical reflection of the work not included in the original article.

### **2.7.1 Methods**

An integrative review was selected as the most appropriate methodology as it enabled the bringing together, or collective synthesis, of many data sources (Whittemore and Knafel 2005). The integrative review enabled the collection and synthesis of multiple perspectives (233 patients, 75 healthcare professionals and 57 relatives) within a relatively short time frame. Others have interviewed individual patients or conducted focus groups to gather similar data, but the integrative review enabled the findings from all of these approaches to be synthesised (Attree 2001, Barelds et al 2009a, O'Reilly 2007). An integrative review is a recognised research methodology, which requires a clear research question, prior identification of inclusion criteria, a quality critique of relevant studies and synthesis of findings (Whittemore and Knafel 2005). It differs from other systematic literature reviews as it allows the integration of non-empirical data, which was a key source in the literature defining and conceptualising quality in healthcare (Whittemore and Knafel 2005). In addition, an integrative review enables synthesis and reinterpretation of specific concepts, so was an appropriate method to explore the current meaning of quality in healthcare (Broome 1993). All stages must be adequately documented to enable replication by others.

The research question was refined from "What is the current definition of quality in healthcare?" to "Do the Institute of Medicine's (IOM) dimensions of quality capture the current meaning of quality in healthcare?" before the review commenced. Integrating the IOM domains as part the research question did two things; firstly it enabled the review to be taken from the most recent domains of healthcare quality cited in healthcare quality policy (Scottish Government, 2010), and, secondly, it provided a framework and focus for the potentially unmanageable scope and number of papers relating to healthcare quality.

### **2.7.2 Data Sources**

Devising search criteria which produced specific, yet manageable, results was challenging. There were large volumes of papers categorised under the Medical Subject

Headings (MeSH) used. For example, a Medical Literature Analysis and Retrieval System Online (MEDLINE) search of the term 'quality of healthcare' produced 48,393 results. There was a necessity to move backwards and forwards through a process of trial and learning to arrive at a definitive search strategy which yielded specific and reliable results. As described in the Paper, to ensure the data were manageable, the search strategy was limited to papers that had the key search terms within their title. Whilst this may have missed some relevant papers, the limitation was necessary to ensure the review was feasible within a given time frame and resource (further discussed under section 2.8). Secondary references were checked from all retained papers and only three additional papers were retained following application of the inclusion criteria, which provided some reassurance of a sufficient search strategy.

### **2.7.3 Study Selection**

Inclusion criteria, used to determine whether or not studies are included in the systematic review, influence the study results and findings (McDonagh et al 2013). Therefore, ensuring a consistent approach to inclusion decision-making is important to reduce study bias. The aim of a consistent approach is to ensure that each study has an equal chance of being selected in accordance with the pre-defined criteria. Bias was therefore minimised by pre-determining exclusion criteria. As described in the Paper, the terms Subject, Population, and Context were used to structure the inclusion and exclusion criteria, which were derived from an adaptation of the Population, Intervention, Comparison and Outcome (PICO) acronym (Lang 2004). PICO could not be used in its current format as the enquiry was not made in relation to an intervention. The following paragraphs provide an explanation of the exclusion and inclusion criteria applied (also briefly outlined in Table 2 within the Paper).

#### **Subject**

The inclusion criteria focused on the subject of *quality of healthcare* as defined or utilised by the authors of the papers. There were many papers discussing quality in relation to a specific treatment or disease process, such as Castilla et al (2008), who defined quality in relation to IVF treatment, or Braunstein (2003), who defined quality in relation to diabetic care. Papers such as these were excluded as they were too speciality-specific to help establish what the current meaning of quality of care was for general healthcare. Likewise, multiple papers exploring performance improvement or measurement, such as the paper by Heenan et al (2010), explored governance of quality from a Health Board perspective, but did not provide any detail on defining or conceptualising the term

healthcare quality and hence were excluded. Papers were also excluded where the primary focus was a specific aspect of healthcare quality, as opposed to a global conception of healthcare quality. For example, the paper by Collignon et al (2002) discussed surveillance definitions for multi-resistant organisms for infection control quality monitoring, rather than healthcare quality definitions or domains.

## **Population**

The population was patients, service-users and healthcare professionals or any other term used to describe those accessing or providing healthcare. Population exclusions included animal, in-vitro or laboratory, such as the Sirota (2006) paper on error in anatomic pathology, for obvious reasons. Populations included were determined by the papers, for example, whether or not subjects were healthcare users (patients, service users or families) or providers (nurses, doctors, managers).

## **Context**

The context was healthcare. Papers on the context of Eastern healthcare, such as by Hyder (2002), were excluded as the focus of inquiry was Western healthcare. There is evidence to suggest that there are significant variations in definitions of quality across less affluent healthcare systems, therefore including these papers would have threatened the validity and transferability of the findings (Al-Zaru et al 2013). Nursing home and residential settings were eliminated as subsequent definitions of quality would be specific to the context of long-term and private care facilities (Bradshaw et al 2012). Papers on dentistry, such as Barjenbruch et al (2002), were excluded as the focus of the study was in relation to general healthcare. Likewise, papers in relation to end-of-life or terminal care were eliminated as their healthcare needs are different to those requiring general acute care, rehabilitation or health promotion (Brook 1973, Nelson et al 2010).

All decision-making for inclusion and exclusion were documented, including a 10% ( $n=19$ ) sample, which was independently checked by a second reviewer. The 19 papers for second checking were selected using a random number generator to reduce bias of selecting papers where decisions might have been less ambiguous. As highlighted in the paper, although only 10% of papers were scrutinised independently for inclusion, there was 95% agreement between both reviewers' decisions, therefore demonstrating a highly reliable process. Where ambiguity had arisen, both reviewers met to discuss the paper. Discussion of the paper where reviewers disagreed (O'Reilly 2007) revealed that the first part of the paper did contain content on how service users defined quality.

The consensus decision was to retain the paper as it made a valuable contribution to defining healthcare quality.

#### **2.7.4 Data Extraction**

The following characteristics were extracted from included papers to describe and analyse the data; author, year, location, population or perspective, participants, context and whether or not an explicit definition of quality was given. Pre-determined information for data extraction enables standardisation of the information utilised from each paper.

The National Institute for Health and Clinical Excellence (NICE) framework (NICE 2006) was used to structure the quality critique of all retained papers (see Table 1). The NICE framework was used as their evidence criteria for grading the study quality was inclusive of a wide range of study types, that is to say, from randomised control trials to expert opinion, which suited the breadth of literature on definitions of healthcare quality and was also suitable for an integrative review that would include different types of articles. The quality grading of included papers ranged from -2 to 4, indicating the potential for bias, largely due to the research methods used. An independent researcher graded 50% ( $n=10$ ) of papers in an attempt to minimise bias. Both reviewers agreed on all grades of quality.

**Table 1: National Institute for Health and Clinical Excellence Hierarchy of Evidence**

Type and quality of evidence	
1++	High quality meta-analyses, systematic reviews of RCTs, or RCTs (including cluster RCTs) with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews of RCTs, or RCTs (including cluster RCTs) with a very low risk of bias
1-	Meta-analyses, systematic reviews of RCTs, or RCTs (including cluster RCTs) with a high risk of bias
2++	High quality systematic reviews of these types of studies, or individual, non-RCTs, case-control studies, cohort studies, Controlled Before and After Study (CBA) studies, Interrupted Time Series (ITS), and correlation studies with a very low risk of confounding, bias or chance and a high probability that the relationship is causal
2+	Well conducted non-RCTs, case-control studies, cohort studies, CBA studies, ITS, and correlation studies with a low risk of confounding, bias or chance and a moderate probability that the relationship is causal
2-	Non-RCTs, case-control studies, cohort studies, CBA studies, ITS and correlation studies with a high risk – or chance – of confounding bias, and a significant risk that the relationship is not causal
3	Non-analytical studies (for example, case reports, case series)
4	Expert opinion, formal consensus
NB: for policy interventions, then CBA can be awarded level 1 evidence	

National Institute for Health and Clinical Excellence (2006) *Reviewing Evidence in Methods for Development of NICE Public Health Guidance*, Chapter 4, p. 25.

### 2.7.5 Analysis

All definitions, domains or other conceptualisations of quality from individual papers were mapped to the IOM's six domains of quality using a data extraction table (see published Paper One). For example, Haggerty et al (2007) identified an attribute of quality as "Technical quality of clinical care: the degree to which clinical procedures reflect current research evidence and/or meet commonly accepted standards for technical content or skill" (p. 340). The attribute was easily aligned to the domain of effectiveness as the IOM describe effectiveness as matching science to care (IOM 2001). Where the mapping was less obvious, or indeed did not fit the IOM domain, the words or phrases were listed under an 'other' section. This reduced the risk of potential misinterpretation. Data were read and reduced from the other section by grouping similar data together.

### **2.7.6 Findings**

An adaptation of the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) flow diagram (Moher et al 2009) was used to document results of each stage (see Figure 1 in Beattie et al 2012). Following removal of duplicates, the search strategy identified 160 papers. Application of inclusion criteria resulted in the retention of 19 papers. Screening of secondary references and application of inclusion criteria identified a further three papers. In total there were 22 papers analysed and synthesised. Papers exploring the meaning of quality in healthcare are mostly derived from expert opinion and consensus. Only 10 out of 24 papers offered an explicit definition of quality. Person-centred care was the most frequently found domain in the data and appeared prevalent irrespective of who (patient, clinician or manager) was defining healthcare quality or the context (hospital ward, clinic or home care) of the enquiry.

Exploring the data which did not align with the definition of the IOM domains resulted in the identification of two additional domains, namely, *system navigation* and *caring*. As defined in the Paper, system navigation describes not only the need to access healthcare services, but also the ability to move seamlessly throughout a complex healthcare system. Caring describes the observed behaviour of people working within the healthcare system which signifies to patients that healthcare practitioners have their best interests at the core of their daily business and tasks. Person-centred care was highly prevalent and embedded within all other quality domains. Rather than person-centred care being a separate domain, it was fundamental to the enactment of all other domains (further discussed in section 2.9).

## **2.8 Critical Reflection of Paper One**

On reflection, the review re-explores the IOM domains, providing a renewed understanding of what constitutes quality in healthcare. The research aim was over-ambitious; highlighting a novice level of understanding about the systematic nature of an integrative review. The review would have benefited from a more focused approach, such as exploring inpatients' definition of quality of healthcare, as opposed to the multiple perspectives of patients, managers, healthcare workers and families. This would have enabled a more manageable review, with results more specific to the patients' definition of healthcare quality, as opposed to a more general understanding. An exploration of the patients' definition would have been more useful for the development of an instrument to measure hospital quality of care from the patient perspective, as it is known

that views differ between those providing and receiving healthcare (Health Foundation 2013).

Interestingly, the literature describes an integrative review in linear stages (Whittemore and Knaf 2005). Systematic reviews are also explained in a linear fashion (Khan et al 2003, Centre for Reviews and Dissemination (CRD) 2009). Yet, I found there was a necessity to go backward and forward with the literature and subsequent search strategy to find papers which were relevant and specific to the research question. I found the process more complex than simply following a logical, step-by-step procedure. Whilst I acknowledge the need for certain steps to occur before others to reduce bias, such as being clear about the inclusion criteria before applying the criteria, there is a necessity to revisit stages as the process becomes more detailed. For example, initial literature searches followed conventions for search strategies, but needed to be revisited when it was realised that some important papers were missed due to how they had been filed within the hierarchical trees within certain databases. It appeared to be more of an art than a science to finally achieve search strategies which yielded specific, inclusive and manageable papers. Perhaps there is a need for texts to describe the integrative review process in stages for the purpose of simplification for learning. I would argue that it is also important to acknowledge the non-linear process required, such as the need to revisit and refine strategies as an understanding of the literature evolves. Other types of systematic reviews, such as meta-ethnographic reviews, acknowledge the necessity of an iterative process (Noblit and Hare 1988).

The risk of bias in paper selection could have been minimised by an additional researcher applying the inclusion criteria to all search results, as opposed to an additional researcher scrutinising 10% only (CRD 2009). Bias, at this stage, refers to the risk of individual pre-formed opinions affecting the chance of a study being included or excluded (Oxman and Guyatt 1993, Slavin 1995). There was also a risk of human error when data were extracted from retained papers, which could have been reduced by having all data extraction checked by another researcher (CRD 2009). Both of these solutions have resource implications. There was no funding available for the review, but alternative strategies could have been employed, such as establishing who could participate in various checking roles and, potentially, offering authorship on the associated publication.

Also, study methods with a higher risk of bias are potentially less credible than RCTs, due to the inability to control and reduce error. The NICE (2006) hierarchy of evidence



was designed to critique studies of clinical interventions and health promotion guidance, as opposed to qualitative studies and theoretical contributions. With hindsight, it would have been more appropriate to use an alternative tool to critique the quality of the included studies, such as those designed by the Critical Appraisal Skills Programme (CASP). CASP have produced eight quality critique questionnaires which are specific to the type of methodology used, for example, there are separate checklists for cohort, qualitative, and other study types (CASP 2013). The most appropriate checklist could have been applied to each included paper to assess the trustworthiness and relevance of the findings. Nevertheless, the advantage of a broad assessment tool was its speed and relevance in this integrative review.

To date, evidence of the various definitions and domains of quality of healthcare has been derived from qualitative studies and expert opinion. This type of evidence needs to be valued for its worth, particularly in theoretical contributions, whilst being mindful of the risk of bias when quality is compromised. The widely accepted IOM domains of quality were formulated through expert consensus at a 'round table discussion' (Personal Communication 2012). These have enabled consensus and a way forward, but their currency and relevance to evolving healthcare required to be reviewed.

Despite the limitations acknowledged in the integrative review, the study has a valid contribution to make. The review highlights the necessity for two additional domains; *system navigation* and *caring*, as well as recognising the foundational nature of *person-centred care*. It is imperative that domains of quality are explicit in policy and strategic measurement plans. If the important domains of 'caring' and 'system navigation' are not explicit within all levels of healthcare measurement plans, there is a risk that resources will not be allocated to these fundamental aspects of quality in healthcare. Diverting resources to current domains of quality to the detriment of these additional domains could diminish, as oppose to improve healthcare quality. There is a necessity to include these domains to develop a measure of hospital quality of care, from the patient perspective, which is to be used for improvement purposes. Otherwise, it is likely that aspects of care will not be improved from the patient perspective. To ensure these domains impact on patient care they need to be measurable. Professor Carr provided an editorial critique of the integrative review and reiterated the importance of measuring these elusive concepts. "The translation of these additional domains into measurable criteria will be important for the delivery of quality healthcare" (Carr 2013, p. 306).

## 2.9 A Revised Model of Healthcare Quality Domains

Findings from the integrative review suggest a re-conceptualisation of quality of healthcare is required. This section provides an explanation of how the domains identified within the integrative review were used to devise a model of healthcare quality domains, which was not included in Paper One. The establishment of a model enables a diagrammatic representation of the theoretical/conceptual framework. A theoretical framework is constructed to explain or predict phenomena (in this case quality of healthcare) and provides the foundation to make generalisations or predications about relationships between variables or domains (Egbert and Sanden 2014). The model is a necessary step to devise a valid measure of hospital quality of care from the patient perspective.



**Figure 2: Beattie's Model of Healthcare Quality**



**Figure 3: Institute of Medicine's 6 Domains of Quality**

The following paragraphs provide a comparison between the IOM model of healthcare quality and the model proposed from this work, hence proposing and articulating Beattie's Model of Healthcare Quality. Figure 2 represents Beattie's Model of Healthcare Quality, namely, care which is; safe, effective, timely, caring, enables system navigation and is person-centred. The model was devised from extensive reading and the results of the integrative review. Figure 3 is a graphical representation of the existing IOM domains of quality in healthcare, used for direct comparison. Note that in the IOM model all six STEEEP domains are equally represented. However, Figure 2 demonstrates that

person-centred care is fundamental to all other domains of quality. Person-centred care means that patients get the care they need, when they need it, and in the way they need it (Berwick 2009). Person-centred care was prevalent in every retained paper within the integrative review. It was not only the prevalence of person-centred care which was evident; how other domains were described included person-centred care in such a manner as to highlight its fundamental nature.

There are numerous examples of poor quality of care, where failure to see the patient's perspective is evident. One example is the case where an 80-year-old lady was 'starved of care' as her condition fell between two silos of medical and psychiatric care (Mental Welfare Commission for Scotland 2011). She was transferred from psychiatric care to medical care as she needed an intravenous drip (medical care required), despite most of her needs being as a result of her vascular dementia (psychiatric care required). She had five hospital or ward moves during her stay. No one individual took responsibility for her nutritional and medication needs and there was no co-ordination or shared understanding of how to manage her physical and mental care. She died on an acute medical ward. Including person-centred care within the domains of healthcare quality might be a small step towards reducing similar failings. The premise that person-centred care is needed is not new (Berwick 2009); however, the fundamental nature of its existence appears to have been overlooked. Without acknowledging the foundational importance of person-centred care, attempts to improve healthcare quality may continue to be limited. Person-centred care is represented in Beattie's Model of Healthcare Quality as being fundamental to all other domains of healthcare quality.

Consider, as an example, the domain of *safety* using an example of medicine administration. The standard dosage of most medication requires adjustment where individuals have underlying liver disease (most medications are metabolised by the liver). Therefore, the adaptation of medication dosage to suit the individual needs of the patient (*person-centred*) is essential if patient *safety* is to be assured. What safety means in specific instances only makes sense in the context of the care for that individual. Applying a person-centred approach to enact all the domains of healthcare quality should result in a quality experience for patients.

Similarly, the domain of *effectiveness* requires practitioners to consider how the research findings of a particular treatment or intervention apply to the uniqueness of a given individual. For example, Warfarin reduces the risk of stroke by approximately 66% in patients who have atrial fibrillation (Hart et al 2007). However, patients who have had a

recent (not clearly defined) gastrointestinal bleed would not be given Warfarin as this would exacerbate the risk of haemorrhage (Go et al 2000). Again, quality of care can only be achieved if the domain of effectiveness is enacted in a person-centred way.

The domain of *timeliness* is also dependent upon person-centredness, if quality of care is the intent. There are standard waiting time targets within the National Health Service, such as the 12-week wait for surgery within Urology services. Currently, many patients wait longer than the 12-week target, referred to as a 'breach.' Once patients have breached the 12-week wait, they are then reprioritised according to the seriousness of the condition, ability to work or perform their primary role, and the associated impact on their mental health. When these individual factors are considered in relation to waiting, time is largely dependent on the domain of person-centeredness – how will the wait affect this individual in particular? The *Timely* domain is dependent upon a *person-centred* approach to care. The waiting time target thus becomes an important domain of quality, from the patient perspective, when articulated in a person-centred way.

Again, the domain of *caring* is largely dependent upon the individual's interpretation of caring behaviour. Studies within the integrative review described the ability of practitioners to anticipate their needs as a predominant example of caring (Attree et al 2001). Although there are many shared values of what constitutes caring, how caring is perceived and interpreted is influenced by individual factors (Watson 2009). Examples of caring found within the integrative review were largely around care being personalised and demonstrations of being treated as persons (Haggerty et al 2007, Howie et al 2004, Larrabee and Bolden 2001). Beattie's model recognises caring as a domain of healthcare quality; which has equal status alongside the other four domains (safe, effective, timely and system navigation).

Finally, *system navigation* is largely dependent on putting the person/patient at the centre of decision-making and service design. Quality was described as poor when different parts of the healthcare system did not communicate with each other to consider the unique requirements of the patient, for example, patients and relatives being unable to comprehend what services were available and how care needed to be coordinated to meet the unique needs of individuals (Haggerty et al 2007, Russell 2007). Families also described a 'chasm' in care when adolescents were transitioning from child to adult services (Barelds 2009b).

Every domain is dependent upon person-centredness being present to attain quality of healthcare. Person-centredness is not just another domain, but fundamental to the

attainment of all other domains. Two domains within the IOM model of *efficiency* and *equity* are subsumed into other domains within Beattie's model. Efficiency is concerned with reducing duplication of effort and making full use of resources (IOM, 2001). Whilst this is of particular importance in the current financial climate, it was not a predominant feature identified in the integrative review. Efficiency may have been a more dominant domain if the perspectives of managers only were represented, due to their duty to manage financial budgets. Efficiency is likely the end product, or outcome, when the other domains of quality have been achieved. The *efficiency* of services is necessary to design services around individual patient needs, as opposed to service design around speciality silos, and these aspects are captured under the domain of system navigation. The necessity of balancing finance with quality will always be a challenge and potential opportunity to how quality is achieved, but this in itself does not make it a domain by which people define healthcare quality. Similarly, equity has been integrated within the domain of system navigation, as the IOM definition is around equal access to service provision for vulnerable groups. Beattie's model considers access to services as a component part of *system navigation*.

Domains have been explored individually to justify and explain the necessary revisions to the IOM model contained within Beattie's Model of Healthcare Quality. However, whilst multiple domains of quality are essential to conceptualise and define the concept, they need to be considered as a whole. Domains of quality do not operationalise in the real world as distinct aspects, but rather, continually interact with each other. For example, in order to receive the most *effective* care, people need to be able to access (navigate) the right care (effective) at the right time (timely). Once accessed, the care needs to be communicated in a considerate way (caring) and actions adapted (safety) to the needs of the individual person (person-centred). Whilst the wholeness of quality is deconstructed for analysis, it is necessary to reconstruct it as a unified whole. The conceptual clarity of what constitutes quality of healthcare is necessary for the development of an instrument to measure the patient experience of hospital quality of care.

## **2.10 Study Contribution to the Research Question**

Whilst acknowledging the methodological limitations of an integrative review, the Paper makes a valuable contribution to conceptualising a current definition of healthcare quality. Identifying two additional domains; namely, *caring* and *system navigation*, and ensuring these are explicit, is an important contribution to defining healthcare quality. All

domains of quality require a person-centred core to attain healthcare quality and, as highlighted in the integrative review, many definitions and models include the domain of person-centredness. Critically challenging the existing domains of quality also helped to inform a revised model of healthcare quality domains. The new model is an important theoretical foundation on which to develop an instrument to measure healthcare quality from the patient perspective, which is valid. There are criticisms in the literature of instruments which have not been derived from any explicit theory (Polit and Yang 2016, Strauss and Smith 2009).

The evolving nature of what constitutes healthcare quality has important implications for designing a measure of hospital quality of care from the patient perspective. Although a current conceptualisation of healthcare quality domains has now been identified, this will soon become time limited. An instrument needs to be designed with core domains which can be easily adapted to context and time (as discussed further later in the thesis) to enable ongoing evolution whilst maintaining a degree of validity and reliability.

The disparity between national and board-level reporting of hospital quality of care and the experiences of individual patients may, in part, be due to how quality is being defined and subsequently measured. Measurement arguably becomes more robust as the stakes for data use become higher, that is to say, encompassing judgement and scrutiny as opposed to only improvement (as highlighted in Chapter 1). If this is so, then it could also be why existing domains (the IOMs) are currently used; they are more amenable to measurement. For example, the domain of time is easily measured through waiting time metrics, but measuring the additional domain of caring, as identified in Beattie's Model of Healthcare Quality, will likely be more challenging. There is a risk that if caring is not made an explicit dimension it will be marginalised in favour of easier to measure domains of healthcare quality. This argument seems particularly relevant given the recent policy directives and supporting literature for 'compassionate care' (Dewar and Nolan 2013, Firth-Cozens and Cornwell 2009, Scottish Government 2010). It is important to establish whether a healthcare quality domain, which is more challenging to quantify, can be measured in practice. Chapter 3 will explore whether empathy, as an indicator of caring, can be measured from a theoretical and practical perspective at the micro level of the hospital system (specifically, the Emergency Department within a hospital).

# Chapter 3

## Measuring Hospital Quality of Care

### 3.1 Measuring Domains of Quality: From Theory to Practice

Previous Chapters highlighted the gap of measuring the patient perspective of hospital quality of care at the clinical microsystem level. Chapter 2 proposed a revised model of healthcare quality, which included domains less amenable to measurement, which could be contributing to the gap between reported metrics of national and board-level hospital quality of care and the experiences of individual patients at the clinical interface. Given that the aim of the thesis is to develop a measure of hospital quality of care from the patient perspective, it is necessary to establish whether a domain (potentially more difficult to quantify) can be measured from the patient perspective in practice. Investigating measurement of the patient perspective of hospital quality of care (Paper Two, embedded in this Chapter) revealed the limitations of measuring patient satisfaction as an indicator of quality of care, hence the need for further exploration of quantifying the patient perspective of hospital quality of care.

The first part of the Chapter embeds a cross-sectional study in the Emergency Department (ED) (Paper Two) which confirmed that measurement of an elusive domain of Beattie's Model of Healthcare Quality (empathy, as an indicator of caring) is possible in a complex hospital environment (ED). The study also confirmed the limitation of using satisfaction as a reliable measure and redirected efforts to measure patient 'experience' as opposed to 'satisfaction' to measure hospital quality of care, from the patient perspective. The Paper is followed by an overview to provide additional study detail, not covered in the word limitations of the publication, and a critical reflection of the Paper. The second part of the Chapter explores the conceptual implications of measuring the patient perspective of quality of care, highlighting the necessity to measure patient experience.

### 3.2 Aim and Linkage to Research Question

This study (Paper Two) aimed to determine whether patients' perceptions of clinicians' empathy, using the Consultation and Relational Empathy (CARE) measure, were a more accurate indicator of satisfaction of quality of care in comparison to waiting time within the ED. In doing so, the study makes three important contributions to the direction of

the thesis. Firstly, the study explores whether a domain of healthcare quality less amenable to measurement can be quantified at the clinical microsystem level. Secondly, the robustness of using patient satisfaction measures to capture the patient perspective is investigated. Thirdly, it enabled consideration of whether the entirety of quality of care, from the patient perspective, is being measured.

### **3.3 Background to Paper Two**

Caring is arguably the most difficult domain of Beattie's Model of Healthcare Quality to quantify; it is therefore important to test how amenable it is to measurement in a hospital environment. Caring has been, and continues to be, a fundamental component of healthcare quality (Attree et al 2001, Haggerty et al 2007, Howie et al 2004, Larrabee and Bolden 2001, Scottish Government 2010, Watson 2009,). Whilst there is debate around whether the wholeness of caring can be measured, there is some consensus that indicators of caring can indeed be quantified (Reynolds and Scott 2000, Watson 2009). Empathy, as an observable and tangible construct, offers a potential indicator of caring and therefore offers a proxy measure of an aspect of the concept (Watson 2009). Empathy has been defined as the ability to understand and respond appropriately to patients' fears and concerns (Mercer et al 2004). Empathy was also chosen as there was a brief and valid instrument available for use, which enabled immediate testing of the concept in clinical practice – the Consultation and Relational Empathy (CARE) Measure (Mercer et al 2004, Mercer et al 2008).

The ED was chosen as it is arguably one of the most complex environments within hospital care. The ED has additional complexity in comparison to other hospital wards and departments as there is wide variation in patient presentations and associated unpredictability (Perez-Carceles et al 2010). There is also a greater degree of uncertainty and anxiety from the patient perspective due to contact with unknown staff, the 'emergency' situation the patients find themselves in, and a constantly changing and unpredictable environment (Cameron et al 2011). There is also limited time in which to obtain data on the patient perception of hospital quality of care, as most patients leave the department within four hours (Scottish Government 2011b). Thus, if caring can be measured in this, often chaotic, environment it is more likely that it can be measured within other clinical areas.



**Objective Two:**

To test the whether empathy, as an indicator of caring, can be measured in an acute hospital setting, from a theoretical and practical perspective.

### 3.4 Paper Two: Compassion or speed? Which is a more accurate indicator of healthcare quality in the ED from the patient's perspective?

The International Journal of Person Centered Medicine Vol 2 Issue 4 pp 647-655

#### ARTICLE

### Compassion or speed? Which is a more accurate indicator of healthcare quality in the emergency department from the patient's perspective?

Michelle Beattie RN (A) BSc (Hons) MSc<sup>a</sup>, Iain Atherton RGN BA (Hons) MSc PhD<sup>b</sup>, Beverley McLennan BA RGN<sup>c</sup> and William Lauder RMN MEd PhD<sup>d</sup>

a Lecturer, School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Inverness, UK

b Lecturer in Nursing and Health, School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Inverness, UK

c Clinical Educator, Emergency Department, Raigmore Hospital, Inverness, UK

d Professor & Head of School, School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Inverness, UK

#### Abstract

**Rationale, aims and objectives:** Devising indicators to measure quality of care is challenging in Emergency Departments (ED). It is difficult to measure aspects of quality which are less amenable to measurement; hence waiting time has often been relied on. This study aimed to determine whether patients' perceptions of empathy are a measurable indicator of quality of care in comparison to waiting time within the ED.

**Method:** A cross sectional survey of patients who attended an ED during a 10 day period was conducted to assess correlation between a measure of empathy (the CARE measure), waiting times and perception of care quality. Data other than waiting times were obtained using a questionnaire completed by patients immediately on completion of treatment. Waiting times were obtained from an existing database. Both waiting times and CARE scores were correlated with responses to a patient satisfaction question using Spearman's rho.

**Results:** Of the 81 patients who participated the majority reported care to be good (21%) or very good (75%). Waiting times varied between 11 minutes and 5 hours 17 minutes. CARE scores ranged from 12 to 50 (mean 41.1). Analysis showed a statistically significant relationship ( $p < 0.001$ ) between ratings of patient satisfaction and CARE measure scores with a moderate correlation (Spearman's rho = 0.55), whereas no correlation was found between satisfaction and waiting time (Spearman's rho = -0.07,  $p = 0.56$ ).

**Conclusions:** Length of time was not associated with patients' perceptions of care quality and hence would have been of limited value as a current measure of quality in the ED. Conversely, empathy was associated with care quality and thus should be considered as a means for assessing quality from the patient's perspective in the context of ED departments.

#### Keywords

Emergency department, empathy, quality measurement, quality of healthcare, waiting time

#### Correspondence address

Mrs Michelle Beattie, School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Centre for Health Science, Old Perth Rd, Inverness, IV2 3JH, UK. E-mail: michelle.beattie@stir.ac.uk

Accepted for publication: 6 September 2012

#### Introduction

The quality and safety of healthcare is variable worldwide, despite an array of quality improvement programmes [1-3]. Current thinking suggests that an appropriate family of measures helps drive quality improvement within healthcare systems [4,5]. Emergency Departments (ED) pose unique challenges in developing improvement measures due to the complex variation of patients and the impact of other system components upon its functioning.

Despite these challenges significant work has been undertaken to develop quality indicators within the ED setting internationally [6,7,8]. The UK waiting time target

is a maximum length of stay in the ED of 4 hours [9,10]. The target was devised in response to public and political concerns reporting long and inappropriate 'trolley' waits in ED [11]. Results were under great scrutiny with financial and other penalties for hospitals failing to reach the target. There were concerns that dysfunctional activity was apparent in order to meet the pressure of the target [12]. Such activity has been termed 'effort substitution' and 'gaming' [13]. Effort substitution is the reduction of effort in other activities which were not being measured. For example, reducing the clinical care patients receive to enable patients to be seen quicker. Gaming refers to an activity which represents the data as better than they

actually are; for example, ambulances waiting outside busy ED until 'ready' to receive the patient. Despite these concerns, there is evidence that the 4 hour wait dramatically improved wait time performance between 2003 and 2006 in England, UK [13]. However, a systematic review found that there was no evidence to suggest that the target had resulted in consistent improvements in care [12].

Most work to date focuses on dimensions of quality which are easily measured, that is, time. Yet, quality is multi-dimensional and necessitates measuring different aspects or indicators, to ensure a more comprehensive analysis. The Institute of Medicine, which provides advice on matters of healthcare quality to the American Congress, have devised 6 dimensions of healthcare quality; namely safety, timeliness, effectiveness, efficiency, equity and person-centredness. Despite originating in America, these dimensions are accepted worldwide [1]. A more recent integrative review elicited the dimensions as safety, timeliness, effectiveness, caring, system navigation and person-centredness [14]. A recent paper devising a framework for measuring quality within the ED also highlights the need to develop indicators for a range of quality dimensions [5].

There have also been concerns that current measures are an inaccurate reflection of quality from the patient's perspective [15,16]. These concerns have become increasingly important as healthcare systems in the UK and other higher income countries attempt to implement systems which are more reflective of patient and public views. Despite an array of literature acknowledging the plurality of perspectives of what constitutes quality, current indicators are devised from the view of providers, rather than recipients of healthcare. Definitions of healthcare quality vary between clinicians, managers, policymakers and those in receipt of health services [17]. Indicators which are representative of patient/public perspectives of quality need therefore to be devised.

This study aims to build on the integrative review by Beattie *et al.* (2012) [14] to determine whether caring can be measured as an indicator of healthcare quality from the patient's perspective within the ED. Although caring may be perceived as an element of person-centredness, we believe that caring should be an explicit and conceptually separate dimension. Creating caring as an explicit dimension would increase the likelihood of measures and targeted interventions to maintain and improve this fundamental dimension of healthcare quality. Otherwise, there is a risk that caring will become marginalised in favour of dimensions more amenable to measurement.

While the elusive nature of caring remains complex, there is some consensus that elements or indicators of caring can indeed be measured [18]. Many definitions or conceptions of caring capture the notion of empathy or "the ability to communicate an understanding of the patient's world" [19]. Empathy may be a useful element or indicator of healthcare quality from the patient's perspective. Most caring theories focus on what people 'say they do', rather than 'what is actually done' [20]. While healthcare has changed considerably from a technical perspective, basic care needs (such as good

communication and caring behaviour) remain central to quality healthcare. As measuring the wholeness of caring remains elusive, empathy, as an observable and tangible construct, may offer a measurable indicator of care (as discussed later in relation to the CARE measure).

This study thus sets out to assess if a measurement of empathy could be effectively used as a measure of quality in an ED setting by assessing: (a) if it correlates with a measure of care quality from patients' perspectives and (b) whether this correlation is greater than any found for a measure of waiting time.

## Method

We hypothesised that a measure of empathy (namely the CARE measure) would correlate with responses to a question assessing perceptions of quality amongst attendees in an ED and that the correlation would be greater than would occur between the indicator of quality and waiting time.

To assess these hypotheses, we conducted a cross sectional survey of all adult patients who attended an ED during a 10 day period in December 2011. Such data were not routinely collected and thus necessitated primary collection. Cross-sectional data collected at the point of time where people were in the ED enabled participants to record their experience of care immediately and thus reduced the likelihood of recall bias. Furthermore, the data enabled analysis to explore associations between different measures of care quality. Ethical approval for the study was sought from and granted by the University of Stirling and the National Health Services Research and Ethics Committee (North of Scotland).

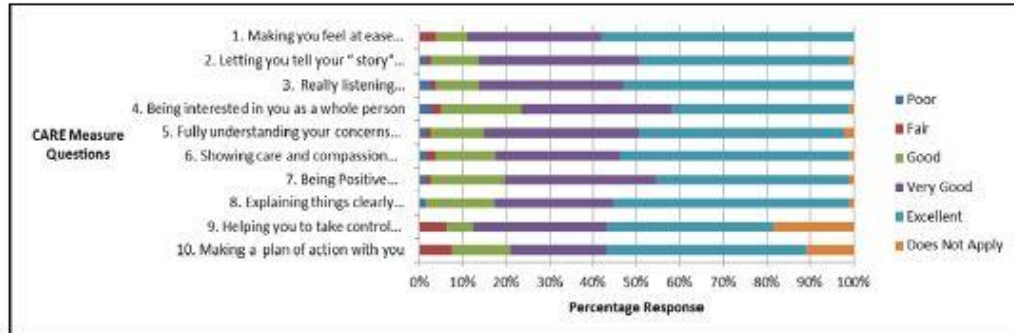
## Setting

The ED is typical of departments across higher income countries. It is located within a 577 bed general hospital with a catchment area including urban and rural areas and sees approximately 33,000 patients per year.

## Data collection tool

The questionnaire contained the Consultation and Relational Empathy (CARE) measure, socio-demographic questions and a rating scale of patient satisfaction. The CARE measure quantifies patients' perceptions of healthcare practitioners' empathy. The measure consists of statements in relation to the healthcare practitioners' ability to understand and respond to patients' fears and concerns, termed 'relational empathy'. The CARE measure is simple and quick to complete making the tool attractive for an ED setting. Other tools were available. We decided against using these, however, given their limited utility in practice, their length and their complexity [21], the need to measure staff and patient's perceptions [22] and their inappro-

Figure 1 CARE Measure Responses



priateness to alternative settings [19]. The CARE measure has been demonstrated to have a high degree of validity in measuring the elusive notion of empathy [23,24].

Some studies have effectively utilised the CARE measure within a secondary care setting [25,26]. The CARE measure represents a modern conception of caring - 'collaboration', rather than 'doing', which is appropriate in current healthcare practice. The theory explicit within the CARE measure explains components of empathy in emotive and behavioural stages. Stage 1 involves the practitioner understanding the patient's perspective and feelings, Stage 2 requires the practitioner to communicate their understanding of the patients perspective and Stage 3 requires the practitioner to act appropriately [27]. These stages are helpful as they clarify that measuring empathy requires more than what practitioners say they do; rather, empathy needs to be demonstrated by the practitioner and perceived by the patient. These behavioural aspects of care were identified as key components of healthcare quality in the literature review previously conducted by Beattie *et al.* 2012 [14].

The CARE measure requires patients to respond to 10 questions using a 6-point rating scale ranging from 'poor' to 'excellent' or to select "does not apply" (see Figure 1). Each response totals to provide an overall score of relational empathy ranging between 10 and 50. The CARE measure has usually been utilised to provide overall empathy score for consultations with practitioners [24-26]. As the purpose of this study was to determine the relationship between empathy scores and satisfaction ratings, individual scores were not calculated; rather, scores were correlated with patients overall ratings of quality of care.

The reverse of the questionnaire included sociodemographic details of the sample to determine whether these characteristics influenced satisfaction of quality of care ratings (see Table 1). Age, gender and type of practitioner have been identified in the literature as factors which may influence satisfaction with healthcare [28-30].

Table 1 Sociodemographic details of the sample

Variable	% of Patients n=81
<b>Gender</b>	
Male	50.6
Female	49.4
<b>Age (Years)</b>	
18-29	28.4
30-39	12.3
40-49	19.8
50-59	8.6
60-69	14.8
70-79	9.9
80-89	4.9
89-99	1.2
<b>Seen by</b>	
Doctor	45.7
Nurse	22.2
Both	18.5
Don't know	6.2
Missing	7.4

The main outcome of interest was how patients perceived the quality of care they had received during their visit to the ED. Patient satisfaction measures are commonly used to determine quality of care from the patient's perspective [31]. A patient satisfaction rating scale was used to capture patients' perception of quality of care (see Table 2). A 5-point scale ranging from 'very good' to 'very poor' was devised to encourage a response

which was reflective of the patients' perception of quality of care [32].

**Table 2 Patient satisfaction ratings**

Satisfaction	% of Patients n=81
Very Good	75.3
Good	21.0
Fair	2.5
Poor	1.2
Very Poor	0

### Data collection

Patients were recruited as they presented at the ED. Clinical staff distributed Study Information Leaflets to all patients on arrival who were 18 years of age or over and who were considered to have the capacity to give informed consent. Patients who agreed were then seen by members of the study team who addressed any questions and, for those who remained agreeable to participate, consent was obtained. Patients in Scottish hospitals have a unique number used for administrative purposes. This number was inserted into the questionnaire to enable the study team to link information on the length of time spent in the ED (information routinely recorded by ED staff). Patients completed the questionnaires after their assessment and treatment and immediately before leaving the department for discharge, transfer or admission. Completed questionnaires were returned in envelopes and deposited into a collection box. Members of the research team matched the waiting times to questionnaires from the hospital database.

### Statistical analysis

Data were entered into SPSS (version 17) for analysis. An overview of respondents was ascertained by calculating descriptive figures for age, gender and practitioner type using percentages and means as appropriate. The continuous variable for age was transformed into categories commonly used in existing data sets for ED in the UK to enable comparison with other data.

Patient satisfaction with quality of care was the primary outcome measure. Initial analysis established the data to be non-parametric (the majority of responses to the satisfaction with quality of care question being 'good' or 'very good') and so Spearman's rank was used to assess correlation between the measure of empathy and perceptions of quality and between waiting time and perceptions of quality.

We transformed responses to the patient satisfaction question into a binary variable. Patients are known to overrate the care they receive, which can result in responses of 'good care' (as opposed to 'very good care'), actually meaning 'substandard care'. Previous studies have found people to be reticent to report negative experiences [33-35]. We managed the validity threat of using patient

satisfaction as an outcome measure by using a high threshold of what constitutes good care. The binary variable was categorised as 'good' and 'not so good' care. 'Good' care was composed of responses from participants who rated quality of care as 'very good' only. 'Not so good' care was composed of all other responses - 'very poor', 'poor', 'fair' and 'good'.

## Results

The sample was generally young and included a roughly equal number of males and females. Of the 81 patients 41 (51%) were male. Twenty-three (28%) were aged 18-29 years (mean 46.5) indicating that the sample was more representative of young people. Almost half of the sample (46%) had their consultation carried out by a doctor and a small proportion (6%) did not know whether they were seen by a doctor or a nurse. Waiting times varied between 11 minutes and 5 hours 17 minutes with a mean of 1 hour 48 minutes, which indicated that waiting times were generally low and mostly within the 4 hour wait target. CARE measure scores varied from 12 to 50 (mean 41.1), indicating wide variation in patients' perceptions of empathy by caregivers, although mostly reporting high levels of perceived empathy. Gender did not appear to make any difference in respondents' ratings of satisfaction with care - comparative analysis indicated no significant differences between male and female responses and overall satisfaction with quality of care (Chi 2 = 1.20,  $p=0.274$ ).

Most patients rated overall satisfaction of care highly, indicating their care to have been either very good (75%) or good (21%) with only a very small proportion (4%) indicating otherwise. Even if we assume those indicating care to be no better than 'good', these figures still indicate that only 25% were at all dissatisfied with care (see Table 3). Women were more likely to rate their care as 'not so good' (30% out of 49.5%); however, this was not statistically significant (Chi Square 1.20,  $p=0.27$ ). A statistically significant difference was found between dissatisfaction rates of younger people (36.3%) compared to older people (20.0%) (Chi Square 4.08,  $p=0.04$ ).

Some patients felt the last 2 questions on the CARE measure were not applicable - Q9: helping you take control (18.5% not applicable) and Q10: Making a plan of action with you (11.1% not applicable). These 2 questions related to patient involvement in their care, which may have been perceived as less relevant in the ED setting. These 2 questions may need refinement to ensure the questionnaire is appropriate in an ED setting.

The Spearman's rho test showed a statistically significant relationship between ratings of patient satisfaction and CARE measure scores with a moderate correlation (Spearman's rho = 0.55,  $p<0.001$ ), whereas no statistically significant correlation was found between satisfaction and waiting time (Spearman's rho = -0.07,  $p=0.56$ ).

**Table 3 Comparison of 'good' and 'not so good' patient satisfaction ratings**

Variable	% Patient Satisfaction Good n=81	% Patient Satisfaction Not So Good n=81
<b>Gender</b>		
Male	80.5	19.5
Female	70.0	30.0
TOTAL	61 (75.3)	20 (24.7)
<b>Age (Years)</b>		
18-29	60.9	39.1
30-39	70.0	30.0
40-49	81.3	18.8
50-59	85.7	14.3
60-69	83.3	16.7
70-79	87.5	12.5
80-89	75.0	25.0
90-99	100.0	0
<b>Seen by</b>		
Doctor	83.8	16.2
Nurse	66.7	33.3
Both	80.0	20.0
Don't know	60.0	40.0

**Discussion**

To our knowledge, this is the first study which aims to explore the measurement of empathy within an ED setting. The results demonstrate a moderate and statistically significant correlation between empathy and satisfaction with quality of care. In other words, those who considered their care to have been of high quality were also more likely to have perceived staff as being more empathetic. This finding suggests relational empathy (CARE measure) to be a valid indicator of healthcare quality from the patient's perspective. This finding has important implications when determining quality indicators for an ED setting, specifically those which encompass the notion of person-centred care.

The study design does not control for other patient and environmental factors which may have influenced patient satisfaction ratings. We do not believe our findings to have been the result of selection bias. Most of our patients were towards the minor end of the spectrum for levels of illness or injury seriousness. Studies of patient satisfaction in the ED have found that patients presenting with urgent conditions were more likely to be satisfied with their care than those who presented with less urgent conditions [36]. Other studies have noted that as the severity of the presenting complaint increases so does the level of communication between patient and practitioner [30]. For ethical reasons, we had to exclude patients if they lacked capacity to give consent, which included those temporarily

incapacitated from opiate analgesics, sedation or altered levels of consciousness. We acknowledge that the study group were predominantly 'walking wounded', which limits the transferability of these findings. Excluded patients were likely to have had more serious conditions or injuries. Individuals who present with major injury or illness may have a longer length of stay within the ED due to the complexity of their condition, which could influence their overall rating of satisfaction; however, again these individuals were likely to have had more serious conditions and so again would have been more likely to have a high level of empathetic care.

We found most care to be of a high quality, with 3-quarters of the respondents indicating care to be 'good' or 'very good'. These findings are consistent with those reported for EDs elsewhere [36]. We found care to be largely empathetic, yet public perceptions of ED may be skewed by the portrayal of negative stories in the media and the under-reporting of positive experiences. However, there is also evidence which suggests that patients do not readily express their dissatisfaction with healthcare quality [37]. Some of the reasons for under-reporting include the patient perception that the issue is outwith the frontline practitioners' control, for example, having to wait a considerable length of time to see a specialist. Also, patients may feel they do not have the expertise to judge the technical aspects of care or indeed automatically assume technical competence of staff [38].

Since the publication of a key paper by Erikson (1987) [38], who questioned the validity of patient satisfaction as a measure of care, many papers have supported the notion that when care meets or exceeds the patients' expectations, then they are more likely to report high levels of satisfaction with care. Leonard (2007) [39] gives a balanced view of the pros and cons of using satisfaction as a measure of quality of care. Despite the debate surrounding the use of patient satisfaction instruments, there is consensus in the literature that the quality of the interaction between patients and practitioners is a strong predictor of quality of care [40].

Empathy and quality of care were significantly and positively correlated, indicating that one is associated with the other, thus valuing the use of patient satisfaction as an outcome measure. This finding may conflict with perceptions of ED being uncaring environments, where the focus is to move the patient through a busy system in a timely manner [41,42]. The brief encounters between patients and clinicians may limit the extent to which a therapeutic relationship can be developed. However, our findings are a reminder that empathy remains central to patients' perceptions of quality of care in an ED.

Consistent with other studies, our sample showed that younger patients were less satisfied with their care; however, no differences were found between other age categories and by gender. The latter finding is consistent with other research in an ED setting [29,36]. Other studies have also found that younger people are more likely to express dissatisfaction than older patients [30,36]. This may be due to the fact that younger patients have higher expectations of the care received in ED or even inappropriate expectations, such as presenting to ED with

**Table 4 Spearman's rho determining the relationship between empathy/waiting time with satisfaction of quality of care ratings**

Variable		How would you rate your overall satisfaction with the quality of care you have received today?	Total empathy score	How many minutes spent in the ED?
How would you rate your overall satisfaction with the quality of care you have received today?	Correlation coefficient	1.000	0.549**	-.066
	Sig. (2-tailed)		0.000	0.559
	N	81	80	81
Total empathy score	Correlation coefficient	0.549**	1.000	-.120
	Sig. (2-tailed)	0.000		0.289
	N	80	80	80
How many minutes spent in the ED?	Correlation coefficient	-.066	-.120	1.000
	Sig. (2-tailed)	0.559	0.289	
	N	81	80	81

\*\* Correlation is significant at the 0.01 level (2-tailed)

minor ailments that could be managed effectively by other healthcare providers. Notably, our study found that empathy remains an important indicator of healthcare quality irrespective of age – young patients who were dissatisfied also reported low levels of empathy by practitioners. Regardless of age, therefore, measuring empathy remains a valid aspect of quality measurement within ED.

Conceptions of quality are likely to change as the discourse of society and healthcare changes. For example, waiting time was likely to be an important indicator within the ED in the UK before substantial improvements were made in this area. Compliance with the 4 hour target in the UK increased from 77% in 2002 to 96% in 2004 [12]. Given that waiting time in ED has improved dramatically over the last few years in the UK, this may no longer be an important predictor of quality of care in the ED. As found in this study, busier time periods do not necessarily result in less satisfied patients [36]. Previous research has also indicated that perceived waiting time is a stronger predictor of patient satisfaction than actual waiting time [36,43-45]. This finding supports the widely recognised disconfirmation paradigm; where perceptions of quality of care are influenced by confirmation or rebuttal of expectations. For example, if a patient expects to be seen within 1 hour of arrival, being seen within 30 minutes would constitute satisfaction; whereas dissatisfaction is likely if the patient has had to wait 1 hour and 15 minutes. Current thinking supports the disconfirmation paradigm, in which dissatisfaction arises when patient expectations are not met [46].

A recent systematic review found no clear evidence to suggest that the 4 hour target actually improved quality of care in the ED [12]. Our study further supports these findings that waiting time is at best weakly linked to quality of care. In April 2011, NHS England introduced a group of measures which aimed to give a more balanced view of performance within the ED. The aim of the 4 hour target was amended to 95% of attendees being seen on time (previously 98%). Other measures, for example, the

number of patients who left before treatments, were also introduced. These changes suggest that the 4 hour wait target alone does not provide an accurate portrayal of quality of care within an ED.

It is difficult to compare average waiting times between studies as there are wide variations in definitions of wait time. For example, wait time within this study was the total time for arrival until departure within the ED (mean 108 minutes). Others have recorded wait time in stages, for example, time from arrival to triage or treatment [43,47] while others have included overnight stays within the ED [48]. However waiting time is defined, its currency as a sole indicator of healthcare quality in the ED appears limited.

Specific indicators of quality are likely to change as practice evolves, for example, as technical and therapeutic interventions advance. Dimensions of quality likely remain constant with some dimensions having more or less prominence over time. For example, time was an important dimension of healthcare quality before significant improvements were made in this area. However, it would be short sighted to banish the dimension of time, as there is a risk that this would eventually result in increased waits for patients. The humanistic and behavioural aspects of caring remain, however, an important and consistent dimension of quality regardless of context or time. Oluwadiya *et al.* [49] studied patient satisfaction with ED care in Nigeria and found that practitioners showing genuine concern, attitude and courtesy were priority areas for improvement. Interpersonal skills such as listening, valuing people as individuals and attempting to understand their concerns, appear to be timeless imperatives of good quality healthcare.

Furthermore, EDs differ from other clinical areas, in that there is a greater degree of uncertainty and anxiety of the patients during their stay, more contact with unknown staff, a constantly changing and busy environment and limited attention to psychosocial issues [29]. Yet, these very factors if anything make the need for empathy even

greater. The measurement of empathy may thus be even more pertinent in EDs than elsewhere.

This study adds to others that have demonstrated that caring behaviours are key predictors of patient satisfaction in an ED setting [15,29,36,44,45,50]. This study, however, suggests one method of measuring a distinct aspect of caring, namely empathy in an ED. The empathy measure reported here, namely the CARE measure, has the potential to be integrated into the daily practice within ED as a quality indicator. Doing so would be especially appropriate with contemporary concerns for ongoing quality improvement.

However, this study found that patients perceived the last 2 questions in the CARE measure to be less relevant than the other questions, 29.6% of patients indicating these questions not to be relevant. This finding is perhaps unsurprising given that these questions centred on the control and planning of care which may be less applicable in an ED setting. These responses may have been attributed to a public perception that ED deal with immediate care, rather than promote self-care and action planning. It is less likely that staff failed to relinquish control to the patient or integrate care planning as patients would have been more likely to rate the response to the question as 'poor' or 'fair', rather than 'does not apply'. As healthcare systems in higher income countries attempt to shift to a more mutual service provision, this finding highlights the cultural shift that will be needed by the public and practitioners to embrace real empowerment and involvement. Mutuality in healthcare requires people actively to influence service provision and enhance their wellbeing in all areas of healthcare including ED.

## Conclusions

Identifying quality indicators is a balance of necessity and sufficiency. There is a risk that dimensions of quality which are difficult to measure will be marginalised in favour of those which are easier to quantify. Despite the continuing changing landscape of healthcare, aspects of caring (of which empathy is an example) remain fundamental to attaining and improving quality of care. This study demonstrates that patients' perceptions of empathy are a measurable indicator of quality of care within the ED. Waiting time, as an isolated indicator, is of little value in determinations of the quality of care in an ED setting.

We suggest a number of areas in which the issues explored in this paper could be further developed. In order to attempt to include the views of patients who were incapacitated, further work would need to be done to ascertain the validity of asking patients' relatives to complete the CARE measure as a proxy. Likewise, parents may also wish to complete the CARE measure after attending the ED with children. The CARE measure offers a potential solution to capture an essential dimension of quality, namely empathy. Capturing the wholeness of quality necessitates inclusions of a person-centered dimension.

## Acknowledgements

This study acknowledges the help and support of ED staff without which the study would not have been possible.

## References

- [1] Institute of Medicine (IOM). (2001). *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington DC: National Academy Press. [http://books.nap.edu/openbook.php?record\\_id=10027&page=R1](http://books.nap.edu/openbook.php?record_id=10027&page=R1) (accessed 5th April 2012).
- [2] Scottish Government. (2010). *The Healthcare Quality Strategy for NHS Scotland*. <http://www.scotland.gov.uk/Resource/Doc/311667/0098354.pdf> (accessed 5th April 2012).
- [3] Quality, Innovation, Productivity and Prevention (QIPP) Right Care. (2011). *The NHS Atlas of Variation in Healthcare: reducing unwarranted variation to increase value and improve quality*. <http://www.rightcare.nhs.uk/index.php/nhs-atlas/> (accessed 11th April 2012).
- [4] Raleigh, V.S. & Foot, C. (2010). *Getting the measure of quality: opportunities and challenges*. London: The Kings Fund.
- [5] Cameron, P., Schull, M.J. & Cooke, M.W. (2011). A framework for measuring quality in the emergency department. *Emergency Medicine Journal* 28, 735-740.
- [6] Department of Health. (2011). *A&E Clinical Quality Indicators: best practice guidance for local publication*. [http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_127776](http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_127776) (accessed 11th April 2012).
- [7] The College of Emergency Medicine. (2011). *Emergency Department Clinical Quality Indicators: A CEM guide to implementation*.
- [8] Beattie, E. & Mackway-Jones, K. (2004). A Delphi study to identify performance indicators for emergency medicine. *Emergency Medicine Journal* 21, 47-50.
- [9] Scottish Government. (2011). *Information Services Division, Accident & Emergency Waiting Times*. <http://isd.scot.nhs.uk/isd/4024.html> (accessed 9th May 2012).
- [10] Department of Health. (2011). *Accident and Emergency: Total Time Spent in A&E*. <http://www.dh.gov.uk/en/Publicationsandstatistics/Statistics/Performanceandstatistics/AccidentandEmergency/index.htm> (accessed 9th May 2012).
- [11] Alberti G. (2004). *Transforming emergency care in England: Department of Health*. 1-48. [http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_4091775](http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4091775) (accessed 9th May 2012).
- [12] Jones, P. & Schimanski, K. (2010). The four hour target to reduce emergency department 'waiting time': a systematic review of clinical outcomes. *Emergency Medicine Australia* 22, 391-298.



- [13] Kelman, S. & Friedman, J.N. (2009). Performance improvement and performance dysfunction: an empirical examination of distortionary impacts of the emergency room wait-time target in the English National Health Service. *Journal of Public Administration Research Theory* 19 (4) 917-946.
- [14] Beattie, M., Shepherd, A. & Howieson, B. (2012). Do the Institute of Medicines' (IOM) dimensions of quality capture the current meaning of quality in health care? – An integrative review. *Journal of Research in Nursing*. 25th April. <http://jrn.sagepub.com/content/early/recent>
- [15] Jolly, E. & Clancy, M. (2009). Waiting times are they that important? A patient survey. *Emergency Medicine Journal* 26, 726.
- [16] Nystrom, M., Dahlberg, K. & Carlsson, G. (2003). Non-caring encounters at an emergency care unit – a life-world hermeneutic analysis of an efficiency-driven organisation. *International Journal of Nursing Studies* 40, 761-769.
- [17] Donabedian, A. (1980). Explorations in Quality Assessment and Monitoring, Vol.1, The Definition of Quality and Approaches to its Assessment. Ann Arbor, MI: Health Administration Press.
- [18] Watson, J. (2009). Measuring caring. In: *Assessing and Measuring Caring in Nursing and Health Sciences*, p10. New York: Springer.
- [19] Reynolds, W. & Scott, P.A. (2000). Do nurses and other professional helpers normally display much empathy? *Journal of Advanced Nursing Studies* 31 (1) 226-234.
- [20] Paley, J. (2001). An archaeology of caring knowledge. *Journal of Advanced Nursing* 36 (2) 190.
- [21] Wolf, Z. R. (1986). The caring concept and nurse identified caring behaviours. *Topics in Clinical Nursing* 8 (2) 84-93.
- [22] Larsen, P. (1984). Important nurse caring behaviours perceived by patients with cancer. *Oncology Nursing Forum* 11 (6) 46-50.
- [23] Mercer, S., Watt, G., Maxwell, M. & Heaney, D. (2004). The development and preliminary validation of the Consultation and Relational Empathy (CARE) Measure: an empathy-based consultation process measure. *Family Practice* 21 (6) 699-705.
- [24] Mercer, S., McConnachie, A., Maxwell, M., Heaney, D. & Watt, G. (2005). Relevance and performance of the Consultation and Relational Empathy (CARE) Measure in general practice. *Family Practice* 22 (3) 328-334.
- [25] Mercer, S. & Murphy, D. (2008). Validity and reliability of the CARE Measure in secondary care. *Clinical Governance: An International Journal* 13, 261-283.
- [26] Mercer, S., Hatch, D., Murray, A., Murphy, D. & Eva, H. (2008). Capturing patients' views on communication with anaesthetists: the CARE Measure. *Clinical Governance: An International Journal* 13, 128-137.
- [27] Mercer, S. & Reynolds, W. (2002). Empathy and quality of care. *British Journal of General Practitioners* 52 (Supplement) S9-12.
- [28] Goldwag, R., Berg, A., Yuval, D. & Benbassat, J. (2002). Predictors of patient satisfaction with emergency care. *Israel Medical Association Journal* 4, 603-606.
- [29] Perez-Carceles, M., Girona, J., Osuna, E., Falcon, M. & Luna, A. (2010). Is the right to information fulfilled in an emergency department? Patients' perceptions of the care provided. *Journal of Evaluation in Clinical Practice* 16, 456-463.
- [30] Raleigh, V., Frosini, F., Sizmur, S. & Graham, C. (2012). Do some trusts deliver a consistently better experience for patients? An analysis of patient experience across acute care surveys in English NHS trusts? *British Medical Journal Quality and Safety*, online first, March. <http://quality.safety.bmj.com/content/early/2012/03/01/bmj.qs-2011-000588.abstract>
- [31] Thiedke, C.C. (2007). What Do We Really Know About Patient Satisfaction? *Family Practice Management* 14 (1) 33-36.
- [32] Ware, J. & Hays, R. (1988). Methods for measuring patient satisfaction with specific medical encounters. *Medical Care* 26 (4) 393-402.
- [33] Kaplan, S. & Ware, J. (1995). The patient's role in health care and quality assessment. In: *Providing Quality of Care: Future Challenges*. 2nd edition, pp. 25-52. (Goldfield N. & Nash D., eds.). Ann Arbor, MI: Health Administration Press.
- [34] Jenkinson, C., Coulter, A., Bruster, S., Richards, N. & Chandola, T. (2002). Patients' experiences and satisfaction with health care: results of a questionnaire study of specific aspects of care. *Quality and Safety in Healthcare* 11, 335-339.
- [35] Nerney, M., Chin, M., Jin, L., Karrison, T., Walter, J., Mulliken, R., Miller, A., Hayley, D. & Friedman, P. (2001). Factors associated with older patients' satisfaction with care in an inner city emergency department. *Annals of Emergency Medicine* 38 (2) 140-145.
- [36] Boudreaux, E., Ary, R., Mandry, C. & McCabe, B. (2000). Determinants of patient satisfaction in a large, municipal ED: the role of demographic variables, visit characteristics, and patient perceptions. *American Journal of Emergency Medicine* 18 (4) 394-400.
- [37] Williams, B., Coyle, J. & Healy, D. (1998). The meaning of patient satisfaction: An explanation of high reported levels. *Social Science and Medicine* 47 (9) 1351-1359.
- [38] Erikson, L. (1986). Patient satisfaction: an indicator of nursing care quality. *Nursing Management* 18 (7) 31-35.
- [39] Leonard, K.L. (2008). Is patient satisfaction sensitive to changes in the quality of care? An exploitation of the Hawthorne effect. *Journal of Health Economics* 27 (2) 444-459.
- [40] Coyle, J. & Williams, B. (1999). Seeing the wood for the trees: defining the forgotten concept of patient dissatisfaction in the light of patient satisfaction research. *Leadership in Health Services* 12 (4) 1-9.
- [41] Wiman, E. & Wikblad, K. (2004). Caring and uncaring encounters in nursing in the emergency department. *Journal of Clinical Nursing* 13, 422-429.
- [42] Coughlan, M. & Corry, M. (2007). The experiences of patients and relatives/significant others of overcrowding in accident and emergency in Ireland: A qualitative descriptive study. *Accident and Emergency Nursing* 15 (4) 201-209.

- [43] Pitrou, I., Lecourt, A., Bailly, L., Brousse, B., Dauchet, L. & Ladner, J. (2009). Waiting time and assessment of patient satisfaction in a large reference emergency department: prospective cohort study, France. *European Journal of Emergency Medicine* 16 (4) 177-182.
- [44] Boudreaux, E. & O'Hea, E. (2004). Patient satisfaction in the emergency department: a review of the literature and implications for practice. *Journal of Emergency Medicine* 26 (1) 13-26.
- [45] Toma, G., Triner, W. & McNutt, L. (2009). Patient satisfaction as a function of emergency department previsit expectations. *Annals of Emergency Medicine* 54 (3) 360-367e.
- [46] Cassidy-Smith, T., Baumann, B. & Boudreaux, E. (2007). The disconfirmation paradigm: throughput times and emergency department patient satisfaction. *Journal of Emergency Medicine* 32 (1) 7-13.
- [47] Booth, A.J., Harrison, C.J. & Gardener, G.J. (1992). Waiting times and patient satisfaction in the accident and emergency department. *Archives of Emergency Medicine* 9 (2) 162-168.
- [48] Ariba, A.J., Thanni, L.O. & Adebayo, E.O. (2007). Patients' perceptions of quality of emergency care in a Nigerian teaching hospital: the influence of patient-provider interactions. *Nigerian Postgraduate Medical Journal* 14 (4) 296-301.
- [49] Ohuwadiya, K., Olatoke, S., Ariba, A., Omotosho, O. & Olakulehin, O. (2010). Patients' satisfaction with emergency care and priorities for change in a university teaching hospital in Nigeria. *International Emergency Nursing* 18, 203-209.
- [50] Brown, A., Sandoval, G., Levinton, C. & Blackstien-Hirsch, P. (2005). Developing an efficient model to select emergency department patient satisfaction improvement strategies. *Annals of Emergency Medicine* 46 (1) 3-10.

### **3.5 Overview of Paper Two**

This section provides additional information and explanation about the cross-sectional survey which were not covered in the publication (Paper Two), due to the word count limit of the target journal.

#### **3.5.1 Methods**

Paper Two described a cross-sectional survey, which usually involves data collection at one point in time from a representative sample of a section of the population of interest (Bland 2001). A cross-sectional survey was chosen as the most appropriate method as the study necessitated the collection of data on the patient perspective of empathy (CARE measure) and overall ratings of satisfaction with quality of care, which are not collected routinely in hospital data. The population of interest was all adult patients who attended the ED during a 10-day period in December, 2011. Given that no funding was obtained to conduct the PhD work, a cross-sectional survey enabled an inexpensive and relatively quick (10 days) means of primary data collection. This method also had the advantage of affecting control over the data collection process; data were collected by myself and the Clinical Educator in the ED. My presence in the ED likely helped remind clinicians to determine the eligibility of patients to participate. Data gathered in cross-sectional surveys are used to establish the relationship between variables (in this case, wait time, empathy and satisfaction with quality of care) which suit the purpose of the research question (Thisted 2006). The research question set out to assess if a measurement of empathy could be effectively used as a measure of quality in an ED setting by assessing: (a) if empathy correlates with a measure of patient satisfaction; and (b) whether this correlation is greater than any found between a measure of waiting time and patient satisfaction.

#### **3.5.2 Team Preparation**

Prior to data collection, the study was discussed with the Clinical Educator from the ED to enable early identification and resolution of any potential challenges. Auditors had recently been in the ED to conduct a time/task in motion study in an attempt to establish efficiency savings by identifying whether specific tasks could be completed by lower-cost staff. Staff reported feeling vulnerable during the audit process. This resulted in efforts to reassure staff that the 'Empathy Study' would not be used for management scrutiny, rather, the data would be used for research purposes. Patient Information Leaflets were distributed at formal staff meetings to reinforce key messages (see Appendix 2); firstly,

that clinical staff would not be required to do any completion of additional paperwork, or even distribute questionnaires. All data would be collected by the researchers (myself and the Clinical Educator). Secondly, the study did not involve identifying empathy scores of individual clinical staff. No identifiable staff data were recorded. Although overall empathy scores for the ED would be known, it was reinforced that the primary purpose of data collection was to establish whether empathy was a more accurate indicator of satisfaction with quality of care in the ED than waiting time, as opposed to how empathetic staff were.

Following ethics approval from the University of Stirling and the National Health Services Research and Ethics Committee (North of Scotland), mutually agreeable dates for data collection were set (see Appendices 3 and 4 for detail of ethics approval). The Clinical Educator identified the clinicians who would be working on the data collection dates and highlighted staff members who had not been present at previous study presentations. Huddles (quick gatherings of staff on shift to discuss pertinent safety issues) were used as a further opportunity to explain the study to remaining staff. Thus, all staff who were to be present in the ED during the data collection period were fully aware of the study.

### **3.5.3 Data Collection**

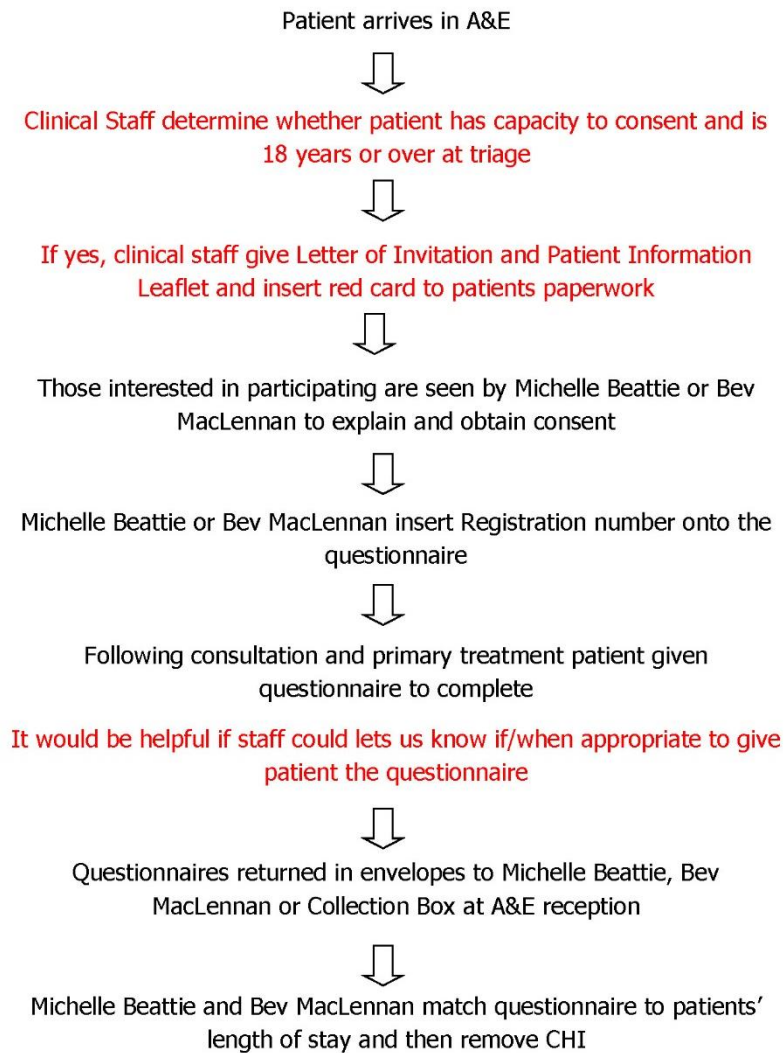
Several patients can arrive at the same time to the ED via different entrances, and they are triaged and treated in different locations and often temporarily leave the department for investigations such as X-rays. Due to this complexity it was important to establish a system to recruit, track and retain eligible patients. Various patient pathways were walked through from admission to discharge for all admission types by myself and the Clinical Educator. For example, those arriving via ambulance did not wait to be triaged and were seen in the area marked 'major' within the ED. Following the walk-through of potential patients' journeys and taking cognisance of the ethical requirement of consent, a Process Flow Diagram for Data Collection was devised, clarified and accepted as feasible by the Clinical Educator (see Figure 4). The text in red highlights the role of clinical staff within the ED.

Patients who agreed to participate were identified by attaching a red card to the front of their clinical notes. For example, a patient arrived on foot to the department, 'checked in' at reception and remained in the waiting room until called for triage. The clinician conducted a brief assessment to determine the severity and urgency of the patient's condition. During this stage the clinician decided whether the patient met the criteria for inclusion in the study ( $\geq 18$  years and having capacity to consent). Eligible patients were

given a Letter of Invitation and Patient Information Leaflet by the clinician. Clinicians advised me or the Clinical Educator of a potential participant. Either of us would explain the study to the patient and allow them time to decide whether they would like to participate. Written consent was obtained by those wishing to participate and a red card was clipped to the patient's notes. As patients are seen and treated, their notes move to correspond to their location, that is, they are moved from the 'waiting' to the 'treated' tray. Similarly, the notes of patients who had been sent to X-ray were stored in the 'X-ray/investigation' tray. Arrival of notes to the 'treated' tray indicated the appropriate time to give the patient the questionnaire to complete.

Participants were given the questionnaire after consultation and treatment, and before discharge, admission or transfer. Completed questionnaires were deposited in the 'Empathy Study Collection Box'. There were often several patients at differing stages of the data collection process at the same time; however, the red card system enabled accurate tracking of multiple participants throughout the ED, which was essential for data collection.

**Figure 4: Process Flow Diagram for Data Collection**



### **3.5.4 Ethical and Legal Considerations**

As highlighted in the discussion section of the Paper, patients who lacked capacity to give informed consent were excluded. Within the ED study, this included those who were incapacitated due to cognitive impairment (i.e. dementia) or those who were suffering from acute mental illness. There were also patients excluded who were temporarily incapacitated from medication (e.g. opiate analgesics), illicit substances (drugs and alcohol) or those with reduced levels of consciousness due to the severity of their illness or injury.

There are ethical principles set out in the Declaration of Helsinki to protect incapacitated patients (World Medical Association 1964). Generally, incapacitated patients are not included in research, but exceptions arise where the subject of interest is a necessary

characteristic of the investigation, for example, unconscious patients in an Intensive Care Unit. Even where exceptions arise, consent must be obtained from a legally authorised representative (World Medical Association 1964). This principle is also enshrined in Scottish Law through the *Adults with Incapacity (Scotland) Act 2000*. Similarly, those incapacitated due to mental health problems are protected under the *Mental Health Act 2007*. The cross-sectional survey did not require the inclusion of incapacitated patients to answer the research question. Patients whose condition changed following the obtaining of their written consent were withdrawn from the study and any associated data were destroyed, that is, those who became incapacitated due to a deterioration in their condition or the effect of medication during their ED visit.

The survey also excluded anyone under the age of 18 years. The legal age of capacity in Scotland is 16 years, whilst in the rest of the UK it is 18 years. There are, however, occasions when those under the age of 18 and 16 can give consent to treatment or research when they are thought to have the capacity to understand the potential implications. These occasions are highly contentious and those under the age of capacity were not required to be included in the ED survey. There is the potential that those aged 16–18 years residing in England, but visiting Scotland, could visit the ED during the study (e.g. during skiing season). To avoid any uncertainty, the upper limit of the legal age of capacity (18 years) was applied to the cross-sectional survey.

### **3.6 Critical Reflection of Paper Two**

On reflection, the ED study provided an extensive learning opportunity; enabling the development of both practical and theoretical research skills. New research skills were obtained from learning techniques to get clinicians on board, completing an ethics application, designing a mechanism to collect the data and learning about statistics. The study also highlighted the limitation of patient satisfaction, therefore alerting me to question the theory and application of measuring the patient perspective, which is discussed in section 3.7.

I felt my nursing background was an advantage to the data collection stage; staff described me as the 'empathy nurse' and I understood when to be sensitive to patient and staff needs. For example, I did not interrupt busy staff to identify patient location or readiness for the questionnaire. Rather, I used non-verbal skills to observe the patient's location and stage in the treatment process. I also knew when it was inappropriate to enter the treatment area, such as when it appeared that the patient and/or relatives were receiving bad news.

There were times when I felt a conflict between my registration as a nurse and my role as researcher within the ED. For example, there were times when the ED was extremely busy. There were patients waiting to be transferred to wards, but clinical staff were unavailable to accompany the patient during transfer, as clinical staff were attending to new emergency arrivals. I could sense the pressure as staff worked hastily in an attempt to move the patients within the four-hour wait target. I wanted to help and felt competent to carry out some activities, for example, transferring patients. Obviously, I was not in the ED as a nurse, therefore I was not protected under NHS employment for vicarious liability (where the employer is accountable for the standard of work only when an employee is performing within their role and job description). I had ethical and managerial approval to be present in the ED to consent and collect patient data, not to deliver patient care. Although I was aware of my legal and professional limitations, I found the conflict of roles to be at odds professionally and morally. I coped by reflecting on such events with the Clinical Educator and my supervisors. The experience reaffirmed the need for me to make a contribution which I felt would have a direct link to patient care in hospitals.

From a methods perspective, the study demonstrated the high ceiling effect of using satisfaction as an outcome measure, which I was previously unaware of. The data collected were mostly towards the high end of the response ratings (high ceiling effect) and therefore required a non-parametric test to assess correlation (Spearman's rank). There is a risk to validity when responses are limited in range; high samples and/or large change is necessary to change categories, therefore limiting responsiveness (Roach 2006). To manage the validity threat, the ratings were regrouped into a binary variable – 'good' and 'not so good' care. 'Good' consisted of only 'very good' responses and 'not so good' was composed of all other responses. This high threshold cut-off point was supported by evidence in the literature that patients overrate satisfaction (Jenkinson et al 2002a, Kaplan and Ware 1995, Nerney et al 2001).

The high ceiling effect did not limit the primary purpose of the ED survey. The survey aimed to establish the relationship between empathy and patient ratings of satisfaction with quality of care, and waiting time, as opposed to how empathetic the ED staff were. Data showed a statistically significant relationship, with moderate correlations, between empathy scores and ratings of satisfaction with quality of care. Conversely, no relationship was found between waiting time and ratings of satisfaction with quality of care. This suggests that even if patients did overrate satisfaction they also treated the CARE measure (empathy) in a similar way.



### **3.7 The Patient Perspective of Hospital Quality of Care: Conceptual Meaning and Measurement**

The ED study (Paper Two) highlighted a limitation of using patient satisfaction to measure the patient perspective of hospital quality of care and the necessity to further explore the conceptual meaning of patient perspective. This section explores quantifying the patient perspective to help describe the learning and redirection of this collection of works from measuring patient satisfaction to measuring patient experience.

The patient perspective became of interest within the UK in the early 1980s with the Griffiths Report recommending the inclusion of patient views to improve healthcare quality and the emergence of consumerism in UK healthcare (Griffiths 1983). The early 1990s witnessed the beginning of competitive market arrangements with the emergence of budget-holding General Practitioners in the UK. Increasingly, more National Health Services are contracted out to other health providers, although less so in Scotland, compared to England. As consumerism rose within UK health services, and society as a whole, so did interest in measuring the patient perspective of healthcare. In fact, the consumerist approach to healthcare influenced the approach to measuring the patient perspective of hospital quality of care (Crow et al 2002). For example, instruments which were developed for retail and banking environments began to be used in healthcare, such as the Service Quality instrument, commonly referred to as SERVQUAL (Chou et al 2005, Parasuraman et al 1988, Shaikh et al 2008). Parasuraman et al (1988) defined 'perceived' quality as the gap between consumers' perceptions and expectations. Perceived quality differs from that of objective quality, in that objective quality is an aspect or a feature of a product or service (for example, the safety kite denotes compliance with UK safety standards), whereas subjective quality is the emotional and behavioural response to the product or service (Holbrook and Corfman 1985). Parasuraman et al (1988) defined expectations as predictions made by consumers about what is likely to happen during a purchase or a transaction. Perceived service quality is then viewed as "the degree and direction of discrepancy between consumers' perceptions and expectations" (Parasuraman et al 1988, p. 17). Application of these instruments to healthcare presumes that patients view healthcare quality in a similar way to commercial and other non-health-related products or services.

Others have described patient satisfaction in a similar way to customer perception (Crow et al 2002): "Satisfaction ... is a relative concept: something that makes one person satisfied (adequately meets their expectations) may make another dissatisfied (falls short of their expectations) (Crow et al 2002, p. 1). If this definition holds, then managing

patient expectations is an important influence on patient satisfaction. For example, as highlighted in the ED study (Paper Two), previous research has indicated that perceived waiting time is a stronger predictor of patient satisfaction than actual waiting time in the ED (Boudreaux and O’Hea 2004, Pitrou et al 2009, Toma et al 2009). This finding can be understood by applying the expectancy-disconfirmation paradigm (Howard and Sheth 1969). Translated to healthcare, the disconfirmation paradigm is where perceptions of quality of care are influenced by confirmation or rebuttal of expectations (Cassidy-Smith et al 2007). For example, if a patient expects to be seen within one hour of arrival to the ED, being seen within 30 minutes would constitute satisfaction; whereas dissatisfaction is likely if the patient had to wait one hour and 15 minutes. The expectancy-disconfirmation paradigm is also related to gap theory, which defines quality as the gap between what a service should provide and the customer perception of what occurred (Boulding et al 1993).

Satisfaction tends to be influenced by patient expectations, and patient expectations are influenced by a variety of factors, other than their healthcare experience (Williams et al 1998). For example, it is known that expectations are likely to shift over time as they are influenced by aging, gender, previous experience and illness severity (Hall et al 1994, Hass et al 2000, Linn et al 1984). Satisfaction is therefore highly individual, which limits its use as a measure of hospital quality. It would be difficult to determine whether changes in satisfaction scores were down to individual factors or change in the quality of hospital care, without using sophisticated statistical modelling (Elliot et al 2010). A study by Salisbury et al (2010) found that when patients were asked a single question on how satisfied they were with their care, only 4.6% of the variance was a result of difference in care; the rest resulted from differences between patients and random error. If an instrument to measure the patient perspective of hospital quality of care is to be used as an ongoing measure of quality improvement at the clinical microsystem it would be unlikely that frontline staff would have the skill or time to use advanced statistical techniques.

Other methodological issues, such as the high ceiling effect (responses clustered at the high end of response options) found in this ED study, have been raised within the literature (Ahmed et al 2014, Coyle and Williams 1999, Fenton et al 2012, Greaves et al 2012, Haggerty 2010, Hendriks et al 2004, Leonard 2008, Mourtzoglou 2000, Salisbury et al 2010, Sofaer et al 2005, Williams et al 1998). For example, results from patient satisfaction surveys suggested that almost all patients are satisfied with their care. Whilst that may seem reassuring, a very narrow response range means such

instruments may lack the ability to distinguish between 'good' and 'not so good' care (Moret et al 2007). For example, if a satisfaction survey was conducted on several hospital wards and most patients responded with 'good' or 'very good' ratings of satisfaction, we would not be able to determine whether there were any differences in quality between the wards. Likewise, if the survey was used as an ongoing improvement measure we would be unable to determine whether the intervention was indeed improving care. The high ceiling effect of patient satisfaction surveys questioned (and questions) their validity; that is, whether such surveys actually measure the construct of interest, the patient perception of hospital quality of care, accurately.

To add to the confusion, the terms 'experience', 'perception' and 'satisfaction' are often used interchangeably; these proxy terms do not overcome the known limitations in using satisfaction as an outcome measure of quality (Parker et al 2003). The literature suggests that patients report high satisfaction, even when their experience has been poor, both for fear of reprisal and due to gratitude bias (Williams et al 1998). With regards to fear of reprisal, patients are vulnerable due to their ill health. The known power imbalances between recipients and providers of healthcare can inhibit an honest response from the patient. Gratitude bias might mean that patients may not be fully open about their quality of care in order to protect frontline staff, especially when poor experience is perceived as being outwith the control of practitioners (Williams 1994). For example, patients may experience a long wait to be seen in the ED, yet still rate satisfaction highly if they felt the wait was not the fault of the nurses and doctors providing treatment. Other research has confirmed that even patients who have suffered an adverse event during hospitalisation rate their satisfaction with hospital quality of care highly (Lopez et al 2009).

Theoretical and methodological issues of defining and measuring patient satisfaction were subject to much debate in the literature, which continues today (Ahmed et al 2014, Coyle and Williams 1999, Fenton et al 2012, Greaves et al 2012, Haggerty 2010, Hendriks et al 2004, Leonard 2008, Moumtzoglou 2000, Rubin et al 1990, Salisbury et al 2010, Sofaer et al 2005, Williams et al 1998). As an alternative to satisfaction, there is evidence to suggest that patient reports of their 'experiences' of healthcare more accurately represent accounts of healthcare quality (Health Foundation 2013, Luxford 2012, Salisbury et al 2010). Measuring patient experience requires questions to be designed around what and/or how often care processes or behaviours occurred, as opposed to patient ratings of care (Dr Foster Limited 2010). For example, a satisfaction survey may ask patients to rate the care process of medicine administration, whereas a

patient experience survey may ask how often they received the right medication at the right time. Rather than asking patients to make a judgement about aspects of their care, patient experience questions are designed to establish factual accounts of whether or how often care processes occurred.

Phenomenologists would argue that hospital quality of care could only be understood through the lived experience of individual patients (Orb 2009), while ethnographers would suggest that patient experience would need to be studied through social interactions, behaviours and group norms (Reeves et al 2008). The term 'patient experience' suggests a qualitative experiential approach to investigating such a complex phenomenon. The literature provides multiple studies of exploring the patient experience of quality of healthcare through qualitative paradigms (Attree 2001, Iedema et al 2011, Sofaer and Firminger 2005). These approaches can support data collection through interviews, observations, focus groups or story telling in order to elicit rich data on patients' experiences of hospital care quality (Creswell 2007, Grassley and Nelms 2009). These methods are particularly useful where an in-depth analysis of the experiential nature of the phenomenon of hospital quality care, from the patient perspective is required. For example, to explore the experiences of quality care in a sub-group of hospitalised patients, that is to say, those with specific conditions or treatments.

The debates between advocates of the use of qualitative and quantitative research have been well rehearsed (Bryman 2006, Buchanan 1992, Pawson and Tilley 1997). However, it is likely that mixed methods are necessary to understand such a complex phenomenon of the patient experience of hospital quality of care (Cornwell and Goodrich 2009, Curry et al 2009, Lagu et al 2013). This thesis, however, is focused on creating a quantitative measure of patient experience of hospital quality of care to be used at the clinical microsystem level, for quality improvement purposes (identified as a current research gap in Chapter 1). There is an assumption inherent within devising a measure of patient experience of hospital quality of care; that the concept can indeed be quantified. There is a degree of realism underpinning this assumption; that the patient experience of hospital quality of care is so complex that absolute truth cannot be confirmed. There is an acknowledgement that all observations are fallible, hence current truth is only an approximation (Onwuegbuzie et al 2009). Attempts to measure the patient experience of hospital quality of care are made to reduce error as much as possible, to be as near to truth as possible. Given the recognised fallibility of observations, findings are based on probabilities as opposed to certainties (Gray 2013).

There is no doubt that quantifying the patient experience within the complexity of hospital care is fraught with difficulties. However, quantifying the patient experience of hospital quality will likely provide an indicator of quality at the clinical microsystem which has the potential to offer a more accurate reflection of quality, from the patient perspective, than current measures.

### **3.8 Study Contribution to the Research Question**

Whilst the limitations of a cross-sectional survey are acknowledged, the study makes three important contributions to the direction of the thesis. Firstly, the study found that empathy is a more accurate indicator of quality, from the patient perspective, than waiting time. Yet, waiting time is the only indicator of quality of care which requires to be collected in the ED. The findings also demonstrate that the domain of caring in Beattie's Model of Healthcare Quality (empathy as an indicator of caring behaviour) can indeed be measured at the clinical microsystem level (ED). Context matters, therefore it cannot be said with certainty that caring (or indicators of it) can be measured in all hospital settings. However, the ability to measure empathy in the ED increases the likelihood that less tangible aspects of healthcare quality can and should be measured at the coalface of clinical practice.

Secondly, the limitation of using patient satisfaction to capture the patient perspective were highlighted; subsequently triggering the exploration of the conceptual meaning of patient perspective which highlighted the benefits of measuring patient experience as opposed to satisfaction. This further exploration redirected efforts to measure patient experience in the development of a measure of hospital quality of care.

Thirdly, the CARE measure was designed for individual practitioner feedback and covers one domain (caring) of what constitutes quality of healthcare. Beattie's and other models of healthcare quality contain more than one domain. Therefore, there remained a need to identify a measure to capture all domains of healthcare quality, from the patient perspective, which is suitable for use at the micro team or unit level of the healthcare system.

Before embarking on the development of a new instrument to measure patient experience of hospital quality of care, it is important to rigorously assess whether an instrument already exists and determine its suitability for use at the clinical microsystem level for quality improvement. The next Chapter includes a systematic review and utility

critique of existing instruments measuring the patient experience of healthcare quality in hospitals.

# Chapter 4

## **What instruments exist to measure the patient experience of hospital quality of care?**

### **4.1 Aim and Linkage to Research Question**

So far, this thesis has established a gap in the measurement of the patient experience of hospital quality of care at the micro level (i.e. hospital ward), which could be contributing to the disparity between reported metrics of national and board-level hospital quality of care and the experiences of individual patients. Chapters 1 and 2 and the associated Paper have identified contemporary domains of healthcare quality. Chapter 3 and its associated Paper identified that a domain of healthcare quality, potentially less amenable to measurement, can indeed be quantified and also confirmed the necessity to measure patient experience as opposed to satisfaction. These findings will be used to help inform the development of a timely and relevant measure of patient experience of hospital quality of care, which could be used for local quality improvement.

The next step, however, was to establish whether an instrument already existed to measure the patient experience of hospital quality of care, which was suitable for use at an operational level for the purpose of team/unit feedback for quality improvement. Developing an instrument is challenging and requires extensive resources, therefore the first step in instrument selection or development is to consider the use of an existing instrument, rather than designing a new one (De Vet et al 2011, Streiner et al 2015). Also, attempts to measure the patient experience of hospital quality of care have been hindered by a proliferation of instruments using various outcome measures (i.e. patient satisfaction, as well as patient experience), with varying degrees of psychometric development and testing (Beattie et al 2014). There has been no previous systematic review to determine the utility of instruments to measure patient experience of healthcare quality in hospitals.

A systematic review was the method selected to determine what instruments existed to measure the patient experience of hospital quality of care and to critique the utility of these instruments. In doing so, this Chapter explores the complexity of psychometrics and the need for a balanced consideration of all aspects of utility in order to select or devise an instrument to measure the patient experience of hospital quality of care for quality improvement purposes. Within this thesis, the utility, or usefulness, of an

instrument is taken to mean the validity, reliability, cost efficiency, acceptability and educational impact of the questionnaire (explained further later).

Whilst devising and conducting the systematic review, this collection of works also makes a contribution to the field of psychometrics and systematic reviews. Usually the quality of papers retained within a systematic review are critiqued using international standards. Whilst international standards exist to critique validity and reliability, there were no standards to critique cost efficiency, acceptability and the educational impact of each instrument. Therefore, standards were devised for these additional, but imperative, aspects of utility and these were applied for the purpose of the systematic review. Also, there were often multiple studies testing different forms of validity and reliability for the same instrument, yet there were no established methods to synthesise the quality and results of studies for the same instrument (Terwee et al 2007). A method of combining (where appropriate) and presenting findings from all five aspects of instrument utility was devised, namely the Beattie and Murphy Instrument Utility Matrix (discussed more fully later in this Chapter).

The first part of this Chapter describes psychometrics and the necessity of instrument utility to be viewed through a wider lens than validity and reliability alone. Van der Vleuten's (1996) aspects of utility are used as a framework to enable a balanced critique of all aspects of utility for instruments measuring the patient experience of hospital quality of care. The second part of the Chapter explains the development of the methods and results of a systematic review and utility critique of instruments measuring the patient experience of hospital quality of care, via the published protocol (Paper Three) and published systematic review (Paper Four). A critical reflection follows each Paper before considering the contribution to the overall thesis. The findings inform the development of an instrument to measure the patient experience of hospital quality of care.

## **4.2 An Introduction to Psychometrics**

The following paragraphs help set the scene for the methods used in the systematic review of instruments measuring the patient experience of hospital quality of care. Psychometrics is the study of the theoretical and statistical methods to quantify abstract or intangible phenomena, an example of which includes the patient experience of hospital quality of care (Polit and Yang 2016). Psychometrics is rooted in psychology and developed from an interest in Darwin's work on differences between animals in adapting to their environment in order to increase survival (Darwin 1872). Two eminent



psychologists, Galton and Cattell, worked on Darwin's theory to determine how individual differences, such as intelligence, could be measured statistically (Cattell 1921, Galton 1874). Modern psychometrics is concerned with the theory and statistical measurement of a wide array of constructs and focuses on the development and testing of measures (mostly questionnaires and tests) of various phenomena. Some psychometricians refer to themselves as clinimetricians as their field of psychometrics focuses on aspects of medicine which cannot be quantified by biophysiological tests or measures, for example, developing a measure of pain assessment (Polit and Yang 2016). Psychometric methods fit with the research aim to develop a valid, reliable, but brief measure of patient experience of hospital quality of care, which can be used at the microsystem level (i.e. the ward). As was observed in Paper Two (ED study), aspects of healthcare quality less amenable to measurement can be quantified when using well designed instruments.

Before an explanation of psychometrics can be given it is important to clarify the terminology used in this collection of works. There is confusion in the literature around definitions of the various types and subdivisions of validity and reliability, compounded by the fact that some terms are used interchangeably (Coaley 2014). This collection of works has used the classifications of validity and reliability as determined by the COSMIN (Consensus-based Standards for the Selection of Health Measurement Instruments) Group for two reasons (see Figure 5 for the COSMIN Domains and Definitions of Measurement Properties). Firstly, as the name suggests, the COSMIN classifications were agreed by an international panel of 43 experts from psychology, epidemiology, statistics and medicine; therefore, a degree of robustness is likely to have been obtained compared to the classification from a single expert (Mokkink et al 2012). Secondly, the standards devised by COSMIN were devised from their classifications and were used to critique the quality of studies found in the systematic review (discussed in part two of this Chapter). Using different classifications from those in COSMIN would have made it difficult, if not impossible, to apply the COSMIN checklists. The following paragraphs provide an overview of the five aspects within the utility framework in order to set the scene for the methods used within the systematic review.

**Figure 5: COSMIN Domains and Definitions of Measurement Properties**

Term			Definition
Domain	Measurement property	Aspect of a measurement property	
Reliability			The degree to which the measurement is free from measurement error
Reliability (extended definition)			The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. using different sets of items from the same health related-patient reported outcomes (HR-PRO) (internal consistency); over time (test-retest); by different persons on the same occasion (inter-rater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater)
	Internal consistency		The degree of the interrelatedness among the items
	Reliability		The proportion of the total variance in the measurements which is due to 'true' <sup>†</sup> differences between patients
	Measurement error		The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured
Validity			The degree to which an HR-PRO instrument measures the construct(s) it purports to measure
	Content validity		The degree to which the content of an HR-PRO instrument is an adequate reflection of the construct to be measured
		Face validity	The degree to which (the items of) an HR-PRO instrument indeed looks as though they are an adequate reflection of the construct to be measured
	Construct validity		The degree to which the scores of an HR-PRO instrument are consistent with hypotheses ( <i>for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups</i> ) based on the assumption that the HR-PRO instrument validly measures the construct to be measured
		Structural validity	The degree to which the scores of an HR-PRO instrument are an adequate reflection of the dimensionality of the construct to be measured
		Hypotheses testing	Idem construct validity
		Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted HR-PRO instrument are an adequate reflection of the performance of the items of the original version of the HR-PRO instrument
	Criterion validity		The degree to which the scores of an HR-PRO instrument are an adequate reflection of a 'gold standard'

<sup>†</sup> The word 'true' must be seen in the context of the CTT, which states that any observation is composed of two components – a true score and error associated with the observation. 'True' is the average score that would be obtained if the scale were given an infinite number of times. It refers only to the consistency of the score, and not to its accuracy.

#### 4.2.1 Why are Psychometrics Important?

If instruments attempting to measure the patient experience of hospital quality are to be used to direct efforts for improvement, there needs to be assurance that the instrument is measuring what it intends to (validity), and that it consistently measures the construct accurately (reliability). Validity and reliability are inextricably linked. An instrument cannot be valid without being reliable, as it needs to measure the construct of interest (whatever that is) consistently. An instrument can, however, be reliable but not valid, as an instrument may reliably measure an invalid construct; it is possible to consistently obtain the wrong answer. Hence, validity is dependent upon reliability, but reliability alone is insufficient (Bannigan and Watson 2009). Reliability increases the upper limit of validity; the higher the reliability of a measure, the higher the possibility of validity, because validity is calculated statistically as the square root of the reliability co-efficient

(Streiner and Norman 2003). Using an instrument which does not capture the patient experience of hospital quality of care will prevent the patient perspective from being fully represented and risks resources being diverted to other aspects of quality of care (as what gets measured, gets prioritised). Similarly, a measure of patient experience of hospital quality of care that is unreliable is untrustworthy in its true measurement of the subject of interest. As a result, an unreliable measure offers no way of determining whether or not patient experience is improving and/or that specific interventions are working, hence the usefulness or futility of improvement efforts will be unknown. It is crucial that an instrument aiming to measure the patient experience of hospital quality of care is both valid and reliable, as well as being usable in practice.

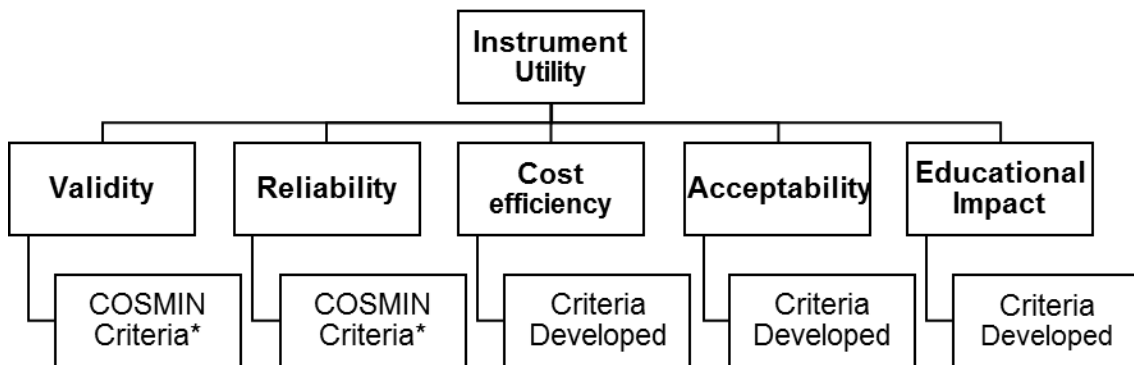
### **4.3 Instrument Utility**

Although validity and reliability are important aspects when selecting and/or designing a measure of patient experience of hospital quality of care, they are not the whole story. As the aim of this collection of works is to devise a measure for team/ward feedback for quality improvement purposes, it needs to be usable in clinical practice. If validity and reliability were the only factors considered in instrument design or selection, there is a risk that an instrument will be chosen that is not fit for purpose. For example, studies may demonstrate that an instrument has high inter-rater reliability (raters agree with a high level of consistency), but the number of raters required may far outweigh the resources available to most, thus rendering the instrument impractical in application. Also, the internal reliability (how well the items are related and accounts for error generated by the items) of an instrument can be improved by increasing the number of items or questions asked, but this needs to be balanced with the burden for the patient completing a lengthy questionnaire (Streiner et al 2015). Given the existing data burden of measures at the micro (hospital ward) level outlined in Chapter 1, there is a need to take a holistic view of aspects of instrument utility to reduce compounding the problem.

Selecting and designing an instrument to measure the patient experience of hospital quality of care for use at the microsystem level of healthcare requires a balanced consideration of all aspects of instrument utility. Van der Vleuten devised a utility framework to critique assessments in education which takes a global view of instrument utility by considering validity, reliability, cost efficiency, acceptability and educational impact (Van der Vleuten 1996). This wide view of utility enables a holistic view of instrument quality, which includes the necessary but not sufficient aspects of validity and reliability. This framework was used to critique the quality of existing instruments in the

systematic review (see Figure 6 for the Framework of Quality Critique of Instrument Utility). The following paragraphs provide an overview of each of the five aspects (validity, reliability, cost efficiency, acceptability and educational impact) to augment the limited explanation within the systematic review Paper.

**Figure 6: Framework of Quality Critique of Instrument Utility**



\*COSMIN provide checklists to critique the quality of different types of validity and reliability i.e. checklist for structural and content validity.

### 4.3.1 Validity

Examining the validity of an instrument necessitates an evaluation of both theoretical and statistical work. Validity is an overall term capturing an array of methods to determine whether an instrument is measuring what it purports to and what conclusions can be drawn from the scores obtained (Streiner et al 2015). Validity is not an ‘all or nothing concept,’ rather, it is a matter of degree. An instrument can never be truly said to be valid, but rather, it can be deemed valid for the population and context in which it was tested (Streiner et al 2015). Validity is cumulative, therefore the more positive results for validity that an instrument has, the more trust users can have that the instrument is measuring what it intends to measure. The following paragraphs describe the classifications of validity as described by COSMIN.

#### 4.3.1.1 Content validity (inclusive of face validity)

Content validity is a judgement of whether or not an instrument adequately reflects the construct of interest, in this case the patient experience of hospital quality of care. The relevance and comprehensiveness of the instrument may have been assessed by exploring the literature to determine how patients define hospital quality of care, or by exploring patient experience of quality of care through focus groups, or interviews, for

example. Face validity is subsumed within this category of validity and is a judgement of whether or not the instrument appears to be representing the construct of interest, in this case, patient experience of hospital quality of care (Streiner et al 2015).

There is criticism in the literature that instrument development is not always theoretically informed, or at least, this is not made explicit in the report of development (Sofaer and Firminger 2005, Wilde et al 1994). Items or domains within an instrument should be derived from theory to make sense of how developers think the instrument will represent the unobservable construct (Edwards and Bagozzie 2000). For example, if the instrument developers' theory suggests that the patient experience of hospital quality care is composed of stages of the patient journey, items (or questions) will be constructed around various stages of that journey, such as admission, ward care and discharge. Once items are constructed to represent the domains of patient experience of hospital quality of care it is necessary to check the sufficiency, or even redundancy of items to capture the construct. For example, Beattie's Model of Healthcare Quality (presented in Chapter 2) was composed of six domains, namely, care that is; safe, effective, timely, caring, enables system navigation and is person-centred. There is usually a set of items or questions for each domain (depending on the type of model). Comprehensiveness of the measure can be tested by getting patients to rate the importance of a list of items, or asking expert panels to add or remove items (De Vet et al 2011). Determining content validity is a subjective assessment, although consistency is likely to be improved by applying standardised criteria, such as those in the COSMIN checklists (Mokkink et al 2012).

Also, a Content Validity Index (CVI) can be used to quantify items devised from previous qualitative work. CVI is a method to reduce the subjectivity when determining whether an instrument has content validity (Lynn 1986). CVI involves establishing an adequate proportion of agreement between experts to determine whether agreement is statistically significant (this issue will be revisited in Chapter 5). As explained, theoretical work and qualitative approaches are essential steps in instrument development in order to accurately represent the construct of interest, such as that of the patient experience of hospital quality of care.

#### **4.3.1.2 Criterion validity**

Once the instrument has been constructed and content validity has been established, it is necessary to use statistical methods to check, verify and potentially amend the instrument to pursue validation (Coaley 2014). For example, if a new instrument was

designed to measure patient experience of hospital quality of care, validity could be tested by asking patients to complete both the new and an existing 'gold' standard instrument and compare the results. This procedure is known as criterion validity (McDowell and Newell 1996). The new instrument would be deemed positive for validity if the relationship (measured as correlations) between the new and existing instrument move in the expected direction and are sufficient. What is deemed sufficient is subject to much debate, but total correlations of 0.7 or above tend to be used in health instruments to indicate validity (Terwee et al 2007). However, this method is difficult to apply when measuring the patient experience of hospital quality of care as there is no established 'gold' standard instrument in which to use for comparison (Beattie et al 2005).

#### **4.3.1.3 Construct validity (inclusive of hypothesis testing, structural and cross-cultural validity)**

##### **4.3.1.3a Hypothesis testing**

Where no established 'gold' standard exists, hypothesis testing can be used (Streiner et al 2015). Multiple hypotheses can be generated a priori based on known attributes of the population or using other measures generated from empirical findings, for example, it is known that patients with poorer health often report a poorer hospital quality care experience (Hewitson et al 2014). Therefore, it can be hypothesised that a measure of health would correlate positively (0.7 or above) with a measure of patient experience of hospital quality care (Terwee et al 2007). A new measure should reflect the known differences between patient groups.

##### **4.3.1.3b Structural validity**

Structural validity tests the degree to which the structure of the instrument reflects the construct (Mokkink et al 2010). Structural validity is examined statistically by conducting factor analysis to explore how many 'factors' are within an instrument or questionnaire. To be structurally credible, factors should explain at least 50% of the variance within an instrument (Terwee et al 2007). That is to say that at least 50% of the items within the questionnaire should be measuring the construct intended, such as patient experience of hospital quality of care.

Identifying such factors will inform developers how many domain areas there are within a questionnaire. For example, the theoretical model for a patient experience instrument might suggest four domains, such as safe, effective, timely and caring. However, factor

analysis might identify five factors (domains) which would suggest that the theory needs to be reconsidered; either there is an additional domain which needs to be named and interpreted, or one of the existing domains may constitute two factors (domains). For example, the domain of 'caring' may need to be split into two sub-domains, such as 'technical' and 'interpersonal'. Given that factors are the collective variance of items (questions) within an instrument, they should explain more variance than any single item.

Factors are explained as percentage variance of the instrument. For example, suppose we had ten items within an instrument and the factor analysis identified two factors where the first factor scored four. This would represent  $4/10 \times 100 = 40\%$  variance. The second factor may score two, which would represent 20% variance. If there were no other factors, this would mean that the model has two factors (domains) which collectively captured 60% of the variance of the instrument. This would be deemed to be a positive result for structural validity. The factor analysis may find other factors, but only factors with an eigenvalue of  $\geq 1$  are retained as factors. Eigenvalues are the statistics used to measure the amount of variation in the total sample accounted for by each factor (Larsen and Warne 2010).

#### **4.3.1.3c Cross-cultural validity**

All positive results of validity tests confirm that the instrument is valid for the population and context in which it has been tested. Instruments to be used in different contexts, such as another country, should be subjected to cross-cultural validity testing (Wong et al 2013). As the name suggests, the instrument is usually adapted using the opinions of those who will use the instrument, such as inpatients in a culturally different healthcare setting. The instrument is then subjected to backward forward translation and tested with results compared to the original instrument version (Leplege and Verdier 1995). Instruments measuring patient experience of quality of hospital care would be likely to differ between Western and low-income countries, for example.

#### **4.3.2 Reliability**

The reliability of an instrument is concerned with the repeatability and consistency of how a construct is measured. The premise with reliability is that whenever an elusive concept is measured there is a degree of error. The less error, the more reliable the instrument, or, in other words, these are inversely related. There is a general acceptance that a degree of error is apparent in all measures, for example, most bathroom scales have an error rate of plus or minus 2lbs, hence error is a small fraction of the true range

of scores (Streiner et al 2015). However, suggesting that the error rate in a patient experience instrument is plus or minus 2 is of little value as there is no common understanding of whether this is an acceptable deviation from the true score. The measurement error of instruments is, therefore, calculated using the ratio of variation between individual scores and the variation between all scores, which is known as variance (Streiner et al 2015).

Potential sources of error differ depending on the type of instrument and the way in which it is to be used. For example, internal consistency reliability is the most common form of reliability testing as it only requires a single administration of the instrument (De Vet et al 2011, Streiner et al 2015). It tests the consistency in which items within the instrument are answered. The internal consistency reliability of the instrument can be improved by removing, adding or refining existing items. When a positive result for internal consistency reliability (Cronbach's alpha > 0.70) is found, it indicates that the items are consistently measuring the construct of interest (Terwee et al 2007). Internal consistency discriminates the subject of interest by accounting for the measurement error generated by the questions asked. It does not, however, account for other potential sources of error, such as those generated by the raters (patients) or between themselves (inter-rater reliability), nor does it account for the stability of the measure at different times of administration (test-retest reliability). It is therefore important to consider whether all the potential sources of error have been accounted for when choosing the form of reliability of interest. These sources of error are relevant when measuring the patient experience of hospital quality of care. For example, patient experience of hospital quality of care may be influenced by recall bias (Black and Jenkinson 2009). Therefore, the ability of an instrument to measure patient experience of quality of hospital care over time is important if the tool is to be used for quality improvement purposes. Also, an instrument aiming to measure patient experience of quality of hospital care for national comparison would need to examine potential sources of error between patients in different health boards, such as levels of illness severity or differences in specialities and age, as these are known to influence patient experience (Bleich et al 2009). Understanding the potential sources of error of an instrument allows the determination of the best balance of items, times of administration and number of patients needed to provide stable feedback for measurement, monitoring and improvement.

Finding a patient experience instrument with positive internal consistency does not, in itself, guarantee that the instrument is reliable for the purpose for which the data will be used. There is a risk that rating the reliability of an instrument without determining



whether the right test has been done could give a false sense of security. Indeed, on occasion there can be a trade-off between validity and internal consistency. For example, a measure to assess the patient experience of hospital quality of care may be interested in different aspects of the patient journey, that is, from admission to discharge. The likely low correlation (a poor patient experience of quality of care on admission may not necessarily mean the patient will have a poor experience on discharge) between these domains of interest would depress the instruments' internal consistency, but be key to the validity of its content. Similar to validity, reliability is not an 'all or nothing' concept, but rather a matter of degree (Streiner et al 2015). Therefore, each study testing an aspect of reliability adds to the evidence that an instrument can measure the patient experience of hospital quality consistently, with an acceptable level of error.

#### **4.3.3 Cost efficiency**

In the current financial climate, cost has become a key consideration when selecting and devising an instrument for patient and/or healthcare practitioner use (McCull 2001). Cost is considered in this collection of works to represent the resources necessary to utilise the instrument for its primary purpose. Obtaining a large, standardised sample will be expensive, for example, questionnaires requiring administration by nurses, or other clinical staff, are likely to be more expensive in comparison to self-completion questionnaires. In the general population the effect of the response burden for lengthy questionnaires remains debateable (Rolstad et al 2011). However, patients are often required to convalesce at home, therefore they are often still unwell at the point of hospital discharge, and thus the length of time it takes for them to complete a questionnaire is an important consideration in this context. Again, cost is often a trade-off for high reliability and validity; obtaining sample sizes necessary for a stable measure may be expensive, and lengthy instruments increase the likelihood of validity, especially when covering complex concepts such as the patient experience of hospital quality of care.

#### **4.3.4 Acceptability**

The term 'acceptability' in this work considers the suitability of the instrument from the user's perspective. This differs from validity as it considers the tolerability of the instrument. For example, studies may demonstrate validity from a statistical test, but users (patients, clinicians and managers) may feel the instrument does not ask the correct questions, or that the results will not be used appropriately. Patients may think a questionnaire has an unacceptably high number of questions, despite internal

consistency (reliability) being improved by increasing the number of items. If the measure is not accepted by those expected to use the instrument there is an increased risk of 'gaming' and the measure becoming an end in itself. Measures need to be credible to clinicians and other users if they are to be used appropriately (Davies 2006). Users' perceptions of the instrument are important to ensure that the measure captures what they think is important and relevant. Validity might ensure that the patient experience of hospital quality of care is being captured, but poor acceptability will likely limit its use in practice. Again, a balanced consideration is necessary as some instruments may demonstrate content validity but have only been tested in a simulated environment or have a high number of questions, subsequently reducing the acceptability of the instrument by users.

#### **4.3.5 Educational Impact**

Educational impact considers evidence around the instruments' ease of use for learning or decision-making. Using a validated and reliable instrument is futile if not followed by action, learning or impact. This category determines how easy it is to make use of the instrument results as intended. Again, this is largely dependent upon the primary purpose of data use. For example, if the data are to be used for ranking hospital performance, these data will likely need to be subjected to complex statistical processes, but may also need to be available in a mode that is easily interpreted for general public use. If, however, the data are to be used for local improvement, they would need to be easily interpreted without the necessity for complex statistical analysis to enable timely interpretation for the frontline team.

#### **4.4 Summary**

Importantly, instruments measuring the patient experience of hospital quality of care need to be of use in the real world (Bannigan and Watson 2009). There is little point in having a valid and reliable instrument that cannot be used in practice. Van der Vleuten emphasises the importance of weighing all of these aspects to select the right instrument, for the right purpose. For example, if results are to be used for high stakes (the outcome has important consequences for an individual or organisation), there is a necessity for high reliability, whilst tolerating high cost. Whilst data used for team improvement may tolerate lower levels of reliability, at the same time, they must contribute to educational impact and acceptability. Critiquing instruments to measure the patient experience of hospital quality of care using the five aspects of utility will aid a balanced consideration for instrument selection and development.

#### **4.5 Paper Three: Instruments to measure patient experience of healthcare quality in hospitals: a systematic review protocol**

Identifying an instrument to measure the patient experience of hospital quality care is complex. A systematic review was conducted to robustly critique the utility of published instruments to measure patient experience of hospital quality of care, therefore enabling instrument selection and identifying whether an instrument existed which could be used to measure local quality improvement at an operational level. A systematic review was selected as it was known there were various instruments measuring patient experience of hospital quality of care, but their robustness and use for quality improvement at a ward level was not known. The systematic review process enabled a rigorous review of all published instruments, as well as examining their validity, reliability, cost, acceptability and educational impact. The review brought together, or synthesised, findings from separate studies of individual instruments to provide an overview of the instrument's utility. Doing so will aide decision-making for those identifying the right patient experience instrument for the right purpose. A systematic review aims to identify, evaluate and summarise findings from relevant studies to make evidence more accessible to decision-makers (Centre for Reviews and Dissemination 2009). The methods were published in a protocol (Paper Three) and registered with PROSPERO (Prospectively Registered Systematic Reviews) CRD42013006754.

##### **Objective Three:**

To identify and critique the utility of instruments which measure the adult inpatient experience of hospital quality of care.

##### **Associated Publications**

###### **Publication Three**

**Beattie M**, Lauder W, Atherton I, Murphy D. (2014) Instruments to measure patient experience of healthcare quality in hospitals: a systematic review protocol *Systematic Reviews Journal*, 3; 4. **Highly accessed.**

<http://www.systematicreviewsjournal.com/content/3/1/4>

#### Publication Four

**Beattie M**, Murphy D, Atherton I, Lauder W. (2015) Instruments to measure patient experience of healthcare quality in hospitals: a systematic review. *Systematic Reviews Journal*, 4; 97. **Highly accessed.**

<http://www.systematicreviewsjournal.com/content/4/1/97>

PROTOCOL

Open Access

# Instruments to measure patient experience of health care quality in hospitals: a systematic review protocol

Michelle Beattie<sup>1\*</sup>, William Lauder<sup>2</sup>, Iain Atherton<sup>1</sup> and Douglas J Murphy<sup>3</sup>

## Abstract

**Background:** Improving and sustaining the quality of care in hospitals is an intractable and persistent challenge. The patients' experience of the quality of hospital care can provide insightful feedback to enable clinical teams to direct quality improvement efforts in areas where they are most needed. Yet, patient experience is often marginalised in favour of aspects of care that are easier to quantify (for example, waiting time). Attempts to measure patient experience have been hindered by a proliferation of instruments using various outcome measures with varying degrees of psychometric development and testing.

**Methods/Design:** We will conduct a systematic review and utility critique of instruments used to measure patient experience of health care quality in hospitals. The databases Medical Literature Analysis and Retrieval System Online (MEDLINE), Cumulative Index to Nursing and Allied Health Literature (CINAHL), Psychological Information (Psych Info) and Web of Knowledge will be searched from inception until end November 2013. Search strategies will include the key words; patient, adult, hospital, secondary care, questionnaires, instruments, health care surveys, experience, satisfaction and patient opinion in various combinations. We will contact experts in the field of measuring patient experience and scrutinise all secondary references. A reviewer will apply an inclusion criteria scale to all titles and abstracts. A second reviewer will apply the inclusion criteria scale to a random 10% selection. Two reviewers will independently evaluate the methodological rigour of the testing of the instruments using the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist. Disagreements will be resolved through consensus. Instruments will be critiqued and grouped using van der Vleuten's utility index. We will present a narrative synthesis on the utility of all instruments and make recommendations for instrument selection in practice.

**Discussion:** This systematic review of the utility of instruments to measure patient experience of hospital quality care will aid clinicians, managers and policy makers to select an instrument fit for purpose. Importantly, appropriate instrument selection will provide a mechanism for patients' voices to be heard on the quality of care they receive in hospitals.

PROSPERO registration CRD42013006754.

## Background

Improving and sustaining the quality of hospital care experienced by patients continues to be a challenge worldwide [1-4]. Current quality improvement thinking advocates the use of measurement to determine whether change initiatives are indeed improving care [2,3]. Measurement, however, is difficult and no single measure can capture the

multitude of facets and outcomes of modern, complex health care systems. The net result has been a proliferation of instruments to measure quality of care.

It is important to establish what constitutes quality of care from the perspective of patients, as well as having the views of clinicians and health care managers, as views differ [5]. Patients, through their unique experience, can offer insights into hospital quality that would be unseen from other perspectives, such as the way a treatment, process or interaction has made them feel and, subsequently, behave. Yet, the majority of measurement plans only include

\* Correspondence: michelle.beattie@stir.ac.uk

<sup>1</sup>School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Centre for Health Science, Old Perth Rd, Inverness M2 3JH, UK  
Full list of author information is available at the end of the article

aspects of quality defined from the perspectives of clinicians and managers. Despite efforts to improve hospital care, the challenge of assuring and improving health care in hospitals remains. There is the potential that measuring and acting on issues of quality raised by patients can be a solution to this intractable problem. There is also increasing evidence that patients who have positive health care experiences have improved outcomes [6] resulting in a more efficient health care system [7]. The necessity to hear the patients' perspective is not new. However, recent aspirations for 'person-centred' care and 'mutual' health care services [3,8] have reaffirmed the imperative for clinicians and health care managers to listen to patients' experiences and act on them to implement improvements.

However, attempts to assess the quality of hospital care by measuring patient experience are challenging. Firstly, there is confusion over the terms 'experience', 'perception' and 'satisfaction'; [5,7] secondly, what constitutes quality within existing instruments is not always defined from the patient's perspective (validity); [9] thirdly, instruments need to produce consistent and reproducible results (reliability) and, essentially, instruments need to be usable in real world practice [10].

First, confusion over the terms 'experience', 'perception' and 'satisfaction' often result in these being used interchangeably, despite known limitations of using satisfaction as a measure of quality [11-14]. Satisfaction has been defined as the gap between a patient's expectations and the actual care he or she received [15]. Yet, many factors influence patients' expectations and these are not static, which threatens the validity of using satisfaction as an outcome measure. Patients do not readily express dissatisfaction with the actual care received for fear of reprisal or because of feeling empathy for those providing frontline care [16,17]. It is thought that a more accurate account of quality of care can be captured if questionnaires are designed around what patients have actually experienced, as opposed to their opinions of the experience [7,18,19]. We need to distinguish between instruments measuring patient experience and those measuring satisfaction/perceptions.

Secondly, instruments attempting to measure a patient's experience of hospital quality care need to do just that. There needs to be sound theoretical and empirical evidence that instruments have been constructed that are representative of patients' views of quality of care (content validity). There are multiple definitions of what constitutes quality of care and views differ between those providing and receiving health services [20-22]. There is a risk that people, with good intent, have developed instruments from supposition about important aspects of quality to patients. We need to determine the validity of existing instruments purporting to measure patient experience of hospital care.

Thirdly, instruments measuring patient experience of hospital quality care need to produce consistent and

reproducible results if they are to be trusted in practice (reliability). Data arising from such an instrument may be used to direct limited resources therefore; the results need to be credible. A recent literature scan highlighted that many studies utilising instruments to measure patient experience provided limited information on their reliability and validity [5]. It is also unlikely that patient feedback instruments developed in-house would have undergone any reliability testing. There is an element of futility in employing an unreliable instrument to help deliver quality hospital care more reliably.

Importantly, instruments need to be usable in real world practice otherwise their sustainability, and therefore their purpose, will be jeopardised [10]. Instruments measuring the patients experience must be acceptable and interpretable to both patients and clinicians. The length and coherence of the instrument needs to be considered to ensure maximum returns and an adequate sample size. The skills required to score and interpret the results of the instrument are another consideration, to ensure timely feedback and use of the findings. Also important is the financial cost of instrument administration, interpretation and feedback mechanisms. These practicalities need to be balanced with other aspects of utility. For example, we know that the more items or questions an instrument contains, the more likely we are to be measuring the construct under enquiry (construct validity). Yet, instruments with multiple questions will be less easy to use in clinical practice due to the length of time it takes for patients to complete them and for staff to analyse and interpret them. There are balances and trade-offs to be made to identify an instrument fit for purpose.

The utility index developed by van der Vleuten [23] provides a useful framework to enable selection of the right instrument for the right purpose. The index consists of five components, namely; validity, reliability, educational impact, cost efficiency and acceptability. The importance of each component is largely dependent upon the purpose of the instrument. For example, an instrument measuring patient experience of hospital quality care to determine the performance rating of a hospital would likely weight more importance on reliability and validity; whereas an instrument used to provide team feedback for improvement would require an emphasis on educational impact, cost efficiency and acceptability. Where the outcome is associated with high stakes, evidence of validity and reliability are required, potentially to the detriment of other aspects of utility. To make a judgement on an instrument measuring patient experience of hospital quality, it is essential, therefore, to establish its intended purpose.

Measuring and acting on patient experience could offer a solution to the complex problem of improving the quality of hospital care. There is a necessity to balance these empirical and theoretical issues to be able to select the right

instrument for the right purpose in the real world. There is a need to identify the range of instruments available to measure patient experience of health care quality, to establish the instruments intended use and assess all aspects of utility. To our knowledge there has been no previous systematic review to determine the utility of instruments to measure patient experience of health care quality in hospitals. There is, therefore, a dear gap in the existing literature, necessitating the proposed review.

#### Study aim and objectives

The aim of this study is to systematically review and critique the utility of instruments available to measure patient experience of health care quality in hospitals. Study objectives are to:

1. Identify the range of instruments available to measure patient experience of hospital care.
2. Determine the intended use of the results of the instrument.
3. Examine the theoretical basis for each instrument.
4. Determine the reliability and validity of each instrument to measure patient experience of hospital care.
5. Categorise instruments according to purpose and outcome of utility critique.
6. Make recommendations on the use of existing patient experience instruments for policy, practice and research.

#### Methods/Design

##### Study method

A systematic review will allow relevant instruments to be identified, evaluated and summarised. This will enable efficient and accessible decision-making of instrument selection to measure patient experience of the quality of hospital care. The review will follow the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) flow diagram and guidance set out by the Centre for Reviews and Dissemination [24,25].

##### Search strategy

We are aiming to identify published instruments measuring patient experience of general hospital care. Therefore, combinations of key words (with appropriate truncation) will be devised in relation to the population (that is, adult patient), context (that is, hospital, secondary care, care setting), measure (that is, questionnaires, health care surveys, instrumentation) and outcome of interest (that is, patient experience/perspective or opinion). The following databases will be searched: Medical Literature Analysis and Retrieval System (MEDLINE), Cumulative Index to Nursing and Allied

Health Literature (CINAHL), Psychological Information (Psych Info) and Web of Knowledge from their inception until July 2013. As per Centre for Review and Dissemination (CRD) Guidance a sample search strategy from MEDLINE is presented below (see Table 1). Experts in the field of measuring patient experience will also be contacted or their websites searched to identify any relevant studies. Duplicate studies will be removed using RefWorks and double checked by one researcher.

#### Inclusion criteria

A reviewer will apply an inclusion criteria scale to all titles and abstracts. A second reviewer will apply the inclusion criteria scale to a random 10% selection. Disagreements will

**Table 1 Search strategy Ovid MEDLINE(R)**

	Advanced search
1	Patient-centred care/
2	Exp *quality indicators, health care/
3	is,fs.
4	**Process assessment (health care)/
5	**Health care surveys*/is (instrumentation)
6	Patient-reported.mp.
7	**Questionnaires*/st (standards)
8	Quality of care.mp.
9	Health care surveys/ or questionnaires/
10	Patient experience.mp.
11	**Outcome assessment (health care)/
12	**Inpatients/
13	is,fs. or measure*.mp. or validation.mp.
14	Inpatients/
15	Secondary care/
16	Hospita*.mp.
17	(Acute adj (service* or care or setting*))/mp.
18	(Patient* adj3 experience*)mp.
19	(Quality* adj3 (care or healthcare))/mp.
20	1 or 10 or 18
21	14 or 15 or 16 or 17
22	5 or 13
23	20 and 21 and 22
24	2 or 8 or 19
25	23 and 24
26	(Patient* adj2 (perspective* or opinion* or experience*))/mp.
27	25 and 26

Footnote for Table 1: An asterisk (\*) represents the most significant concept in Medical Subject Headings within MEDLINE. The slash (/) is used to describe more completely an aspect of a subject. A major topic asterisk before a subheading (/) indicates that that major concept and subheading are associated.

be resolved through consensus. We will ascertain the level of inter-reviewer agreement by calculating Cohen's kappa statistic. As the result of instrument selection from the review could be used for high stakes purposes (that is, ranking in hospital ratings league tables) we would aim for a high level of agreement ( $k > 0.8$ ) [26]. If agreement of the 10% falls below a high standard ( $k < 0.8$ ), a second reviewer will screen the remaining 90%. If this high threshold is not met with two reviewers, we will consider the feasibility of increasing the number of reviewers, or make the level of agreement explicit whilst acknowledging the limitations of increased error. Where decisions are unable to be made from title and abstract alone, we will retrieve the full paper. An Inclusion Selection Form has been devised to ensure standardisation of this procedure (see questions below). This form has been designed on a criteria scale basis; therefore, if the reviewer answers 'no' to the first question, the paper is rejected. This approach will enable progression to further inclusion questions only as necessary, thus enabling a speedy, yet thorough and transparent process. All exclusion decisions will be documented in a tabulated form. Secondary references will be scrutinised for additional instruments not identified in the literature search.

#### Inclusion selection questions

1. Does the study test the psychometrics, theoretical development, or use of an instrument?  
Yes  Go to question 2 No  Reject
2. Is the context of the study a hospital?  
Yes  Go to question 3 No  Reject
3. Is the population adult in-patients in general surgery or medicine?  
Yes  Go to question 4 No  Reject
4. Is the tool measuring the patients' perspective, as opposed to staff or others?  
Yes  Go to question 5 No  Reject
5. Is the tool in relation to hospital care as opposed to being condition specific i.e. quality of osteoporosis care?  
Yes  Go to question 6 No  Reject

6. Is the tool measuring general experience as opposed to satisfaction with a specific profession, i.e. nursing?  
Yes  Go to question 7 No  Reject

7. Is the tool measuring the patients' experience, as opposed to satisfaction?  
Yes  Retain paper No  Reject

Studies that meet the following inclusion criteria will be retained:

- **Date:** We will search retrospectively to the database inception to ensure we examine all catalogued papers available in this field.
- **Language:** Studies in the English language. Studies reported in a language other than English will be excluded due to translation costs.
- **Study Type:** Studies that examine the theoretical or conceptual background or psychometric properties of an instrument measuring patient experience of health care quality in hospitals.
- **Setting:** Instruments that have been tested in a hospital setting, including general surgery or medical ward/facility. Thus, instruments developed and tested in primary care, out-patient centres and other day care clinics will be excluded. Also, we will exclude areas specific to psychiatric or learning disabilities as they would be likely to need instruments developed specific to their needs. We will also eliminate instruments designed specifically for specialist areas such as intensive care, obstetrics and palliative care, as patients in highly specialised areas would be likely to have different determinants of what constitutes quality of care.
- **Participants:** Only adult inpatients will be included. We will, therefore, exclude instruments devised for the paediatric or neonatal population.
- **Global experience of hospital care:** Instruments that aim to measure patient experience of their general hospital care. Thus, condition- or procedure-specific instruments will be excluded (for example, those used to measure aspects of osteoporosis or surgical care). Whilst instruments such as Patient Reported Outcome Measures (PROMS) [27] and Patient Reported Experience Measures (PREMS) are important to determine whether patients have received optimum specialist care and treatment, they will not provide a global measure of patient hospital experience.
- **Patient experience:** We are keen to identify instruments that measure quality from patient experience of direct care. There are a multitude of questionnaires to measure patient satisfaction; however, we intend to exclude these due to the



- methodological limitations identified earlier in this paper.
- **Defining quality:** We will include all definitions or conceptions of quality if they have been devised from the patients' perspective. Exploring how instruments have derived a definition of quality will be an important critique in terms of instrument validity. Ensuring the patient is the subject of interest will remove studies that utilise practitioners', families' and carers', or even managers' definitions of health care quality.

#### Data extraction

A Data Extraction Form will standardise the information recorded and aid analyses. The Data Extraction Form includes study characteristics and the five aspects of van der Vleuten's utility index. Two researchers will independently extract the data for all included studies and agree, through consensus, the accuracy and completeness of the data. Where consensus is difficult to achieve we will use a third researcher to reach agreement (Table 2).

#### Assessment of study quality

We will apply the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist to evaluate the methodological rigour and results of the instruments [28-30]. The checklist has been designed by international experts in the field of health status measurement, but is equally applicable to measuring elusive concepts, such as experiences of hospital care quality. One of the main purposes of the checklist is to evaluate the methodological rigour of instruments for a systematic review [31]. The checklist is made up in modular fashion that enables specific criteria to be applied to certain tests. It is highly likely that one instrument may have several associated studies. The flexibility of various checklists ensures that the same level of scrutiny is applied to judge various studies of instruments, even if they have conducted different validity and reliability tests. See the section 'Judging reliability and validity' for further explanation on implementation of the COSMIN checklist.

Using the information from the Data Extraction Form and results of the application of the COSMIN checklist we will determine the relative importance of each utility item by categorising them as essential, desirable or supplementary (see detail of Utility Index Matrix below). This will enable instruments to be grouped according to purpose and comparisons made with similar instruments. This judgement will be determined by two reviewers through consensus. An independent, third person will be used to arbitrate where necessary. As this will require individual

**Table 2 Data extraction form**

<b>General information</b>	Author
	Year
	Country of origin
	Papers
<b>Instrument detail</b>	Outcome measure
	Purpose/use instrument
	Number and type of categories
	Number of items
	Scale design
<b>Utility characteristics</b>	Type of patients
	Type of environment
	Validity
	Theoretical/conceptual framework
	Types of validity tests conducted and results
	Reliability
	Type of tests conducted and results
	Response rate
	Sample size
	Educational impact
	Ease and usefulness of interpretation
	Feedback mechanism
	Cost efficiency
	Number of rates required to detect difference
Level of expertise required for scoring and analysis	
Acceptability	
Content validity outcomes-appropriateness of language	
Time required to complete the instrument	
Timing of administration	
Mode of administration (that is, self-completion)	
Acceptability by clinical teams and managers	

judgement we will ensure our decision-making is explicit in an accompanying narrative.

#### Application of van der Vleuten's Utility Index Matrix

Each Instrument would be judged (dependent upon extent of testing and purpose) with the following criteria and rated as essential, desirable or supplementary

- Purpose
- Validity
- Reliability
- Educational Impact
- Cost Efficiency
- Acceptability

### Instrument detail

We will need to know how the instrument was administered and used in order to assess the risk and type of measurement error to determine whether psychometric testing was sufficient. For example, we know that the timing of a questionnaire is likely to affect the patient's recall of his/her hospital experience; hence this is a potential source of measurement error. Therefore, if an instrument is measuring patient experience of hospital quality care at three months post-discharge we would expect some testing to determine the stability of the instrument over time (for example, test-retest reliability).

### Examining instrument theoretical development

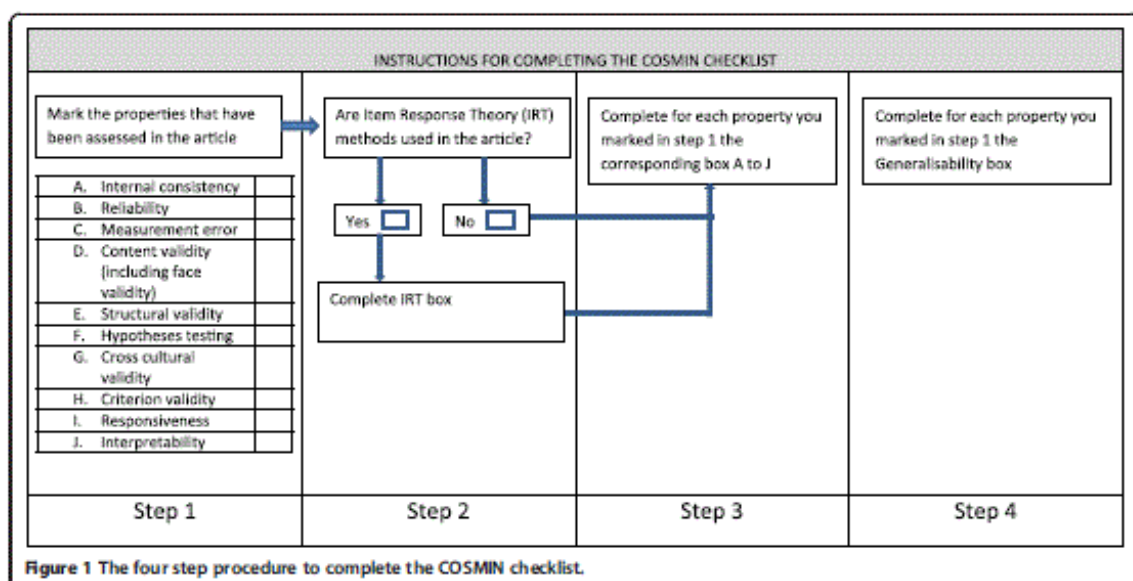
The theory of psychological measurement begins with identification and examination of the theoretical/conceptual development of an instrument, known as content validity. Where the theory underpinning the construction of an instrument is not presented we will search reference lists in an attempt to locate relevant/associated papers. Where evidence of theoretical or conceptual development is not evident we will report this finding. We will critique whether the development of the instrument was informed from the patients' perspective of quality and comment on whether the process of content validity included a theoretical construction and quantification as identified by Lynn (1986) [32].

### Judging reliability and validity

Determining what constitutes sufficient psychometric testing is complex as validity and reliability are matters

of degree, as opposed to 'all or nothing.' However, whilst accepting that psychometric results are dependent upon the purpose, theory and number of items within an instrument, it is also important to establish the rigour of the studies conducted. We will examine the extent of the validity and reliability testing using the COSMIN checklist (see Figure 1). The checklist is applied in a four step process. Firstly, the properties that are being assessed in the study are selected, for example, internal consistency. Secondly, statistical methods used in the study are assessed to distinguish between Classical Test Theory (CTT) and Item Response Theory (IRT). For those using IRT this checklist should be completed. Thirdly, the appropriate checklist is applied depending on type of assessment determined in step one. The checklists contain relevant questions to rate the standards for methodological quality. The final step is to complete the generalizability checklist for each property identified in step one. Using the quality criteria set out by the COSMIN expert group [33] we will classify individual studies of instruments as rating positive, indeterminate or negative. The COSMIN checklist does not quantify an overall quality score as this would wrongly assume that all quality criteria have equal importance [33].

Again, the checklist will be applied by two reviewers independently before they meet to discuss and agree collectively. We will not be excluding studies on the basis of this evaluation. Rather, we will report on all the instruments we have critiqued, as the purpose of the review is to identify and assess the utility of all instruments measuring patient experience of hospital quality of care.



### Data analysis

Where applicable, we will use the general framework and specific tools outlined in the ESRC Guidance on the Conduct of Narrative Synthesis in Systematic Reviews [34]. Numerical counts will be presented to describe general information and instrument detail. We will present individual results of the COSMIN checklist application and the individual study results. We will then collectively compare and contrast instruments with similar purposes for their quality rigour and results. It would be inappropriate to conduct a meta-analysis of results of different instruments due to the variations in the way they are utilised and other heterogeneous conditions. There is currently no empirical method to pool together results of measurement properties; therefore synthesis is recommended [33]. We will categorise instruments with similar purposes and explore the individual and collective findings of application of the utility index. Given that the balance of utility is complex and specific to the function of each instrument, the analysis will be presented as a narrative synthesis. A narrative synthesis of instrument purpose, rigour and findings will enable recommendations to be made on the selection of patient experience measures for policy, practice and future research.

### Discussion

Improving and sustaining health care within hospitals continues to challenge practitioners and policy makers. Patients have unique insights into the quality of care in hospitals, but as yet are an underutilised resource in terms of measurement of quality health care. This systematic review of the utility of instruments to measure patient experience of hospital quality care will enable clinicians, managers and policy makers to select a tool fit for purpose. Ensuring this difficult, yet essential perspective of quality is included could divert resources to improve aspects of care that are important to patients. Harnessing their experience could offer the leverage needed for improvements in the quality of hospital care. We believe that this systematic review is timely and will make a valuable contribution to fill an existing research gap.

### Abbreviations

CRD: Centre for review and dissemination; PRISMA: Preferred reporting items for systematic reviews and meta-analysis.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

MB conceived and designed the study, devised search strategies, drafted the inclusion selection form and drafted the manuscript. WL participated in study design, statistical advice, piloting of inclusion selection form and revision of manuscript. IA participated in study design, piloting of inclusion selection form and revision of the manuscript. DM provided direction for the study idea and design, provided statistical advice and helped revise the manuscript. All authors have read and approved the final manuscript.

### Acknowledgements

I would like to acknowledge the contribution of Kathleen Irvine, Subject Librarian, who shared her understanding of health databases freely and helped to shape the search strategies. Kathleen died suddenly and tragically before this protocol was devised.

### Funding

No specific funding has been allocated for the study; however, the University of Stirling is funding the post of the first author.

### Author details

<sup>1</sup>School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Centre for Health Science, Old Perth Rd, Inverness IV2 3JH, UK. <sup>2</sup>School of Nursing, Midwifery and Health, University of Stirling, Stirling FK9 4LA, UK. <sup>3</sup>Quality, Safety and Informatics Research Group, University of Dundee, Mackenzie Building, Kirsty Semple Way, Dundee DD2 4BF, UK.

Received: 29 October 2013 Accepted: 13 December 2013  
Published: 4 January 2014

### References

1. Institute of Medicine (IOM): *Crossing the quality chasm: a new health system for 21st century*. Washington DC: National Academy Press; 2001.
2. Department of Health: *High quality care for all. Gateway Ref 2008*. 10106.
3. Scottish Government: *The healthcare quality strategy for NHS Scotland*. [http://www.scotland.gov.uk/Resource/Doc/311667/0098354.pdf].
4. Quality Improvement and Prevention (QIP): *The NICE atlas of variation in healthcare: reducing unwanted variation to increase value and improve quality*. United Kingdom: Right Care; 2011.
5. Foundation H: *Measuring patient experience: No. 18, evidence scan*. England: Health Foundation; 2013.
6. Doyle C, Lennox L, Bell D: *A systematic review of evidence on the links between patient experience and clinical safety and effectiveness*. *Br Med J Open* 2013, **3**:e001570.
7. Sofaer S, Fimminger K: *Patient perceptions of the quality of health services*. *Annu Rev Public Health* 2005, **26**:513–559.
8. Department of Health: *A promise to learn – a commitment to act, improving the safety of patients in England*. England: National Advisory Group on the Safety of Patients in England; William Lea publishers; 2013.
9. Lynn M, McMillen B, Sidani S: *Understanding and measuring patients' assessment of the quality of nursing care*. *Nurs Res* 2007, **56**:59–166.
10. Barrigan K, Watson R: *Reliability and validity in a nutshell*. *J Clin Nurs* 2009, **18**:3237–3243.
11. Williams B, Coyle J, Healy D: *The meaning of patient satisfaction: an explanation of high reported levels*. *Soc Sci Med* 1998, **47**:1351–1359.
12. Parker J, Nester CJ, Long AF, Barrie J: *The problem with measuring patient perceptions of outcome with existing outcome measures in foot and ankle surgery*. *Foot Ankle Int* 2003, **24**:56–60.
13. Coulter A: *Can patients assess the quality of health care?* *Br Med J* 2006, **333**:1–2.
14. Byrne K, Sims-Gould J, Frazee K, Martin-Mathews A: *'I am satisfied...but': clients' and families' contingent responses about home care*. *Home Health Care Sev Q* 2011, **30**:161–177.
15. Crow R, Gage H, Hampson S, Hart J, Kimber A, Storey L, Thomas H: *The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature*. *Health Technol Assess* 2002, **6**:1–6.
16. Erikson L: *Patient satisfaction: an indicator of nursing care quality*. *Nurs Manage* 1986, **1**:831–835.
17. Williams B: *Patient satisfaction: a valid concept?* *Soc Sci Med* 1994, **38**:509–516.
18. Sikma H, Kessens J, Campen C, Peters L: *Quality of care from the patients' perspective: from theoretical concept to new measuring instrument*. *Health Expect* 1998, **1**:82–95.
19. Coulter A, Fitzpatrick R, Cornwell J: *The point of care measures of patients' experience in hospital: purpose, methods and uses*. *The Kings Fund* 2009:1–32.
20. Donabedian A: *Explorations in quality assessment and monitoring, Vol. 1, The definition of quality and approaches to its assessment*. Ann Arbor, MI: Health Admin Press; 1980.
21. Maxwell RJ: *Quality assessment in health*. *Br Med J* 1984, **288**:1470–1472.

22. Beattie M, Shepherd A, Howieson B: Do the Institute of Medicine's (IOM) dimensions of quality capture the current meaning of quality in health care? – An integrative review. *J Res Nurs* 2012, **18**:288–304.
23. van der Vleuten C: The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ* 1996, **1**:41–67.
24. Centre for Reviews and Dissemination: *Systematic reviews: CRD's guidance for undertaking reviews in health care: Centre for Review and Dissemination*. York: University of York; 2009.
25. Moher D, Liberati A, Tetzlaff J, Altman DG: Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *J Clin Epidemiol* 2009, **62**:1006–1012.
26. Edwards P, Clarke M, D'Giseppi C, Pratap S, Roberts I, Wentz R: Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening tools. *Stat Med* 2002, **21**:1635–1640.
27. Chow A, Mayer E, Dassi A, Athanasiou T: Patient-reported outcome measures: the importance of patient satisfaction in surgery. *Surgery* 2009, **146**:435–443.
28. Mokkink LB, Terwee CB, Knol DL, Alonso J, Patrick DL, Bouter LM, de Vet HCW: Protocol for the COSMIN study: Consensus-based Standards for the selection of health Measurement Instruments. *BMC Med Res Methodol* 2006, **6**:2.
29. Mokkink L, Terwee C, Stratford P, Alonso J, Patrick D, Riphagen I, Knol D, Bouter L, de Vet H: Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Qual Life Res* 2009, **18**:313–333.
30. Mokkink L, Terwee C, Patrick D, Alonso J, Stratford P, Knol D, Bouter L, de Vet H: The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010, **19**:539–549.
31. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCW: *COSMIN checklist manual*. Amsterdam: Institute for Health and Care Research; 2012.
32. Lynn M: Determination and quantification of content validity. *Nurs Res* 1986, **35**:382–386.
33. Terwee C, Bet S, Boer M, van der Windt D, Knol D, Dekker J, Bouter L, de Vet H: Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007, **60**:34–42.
34. Popay J, Roberts H, Sowden A, Peckicrew M, Arai L, Rodgers M, Britten N: *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews*. ESRC methods programme. England: Lancaster University; 2006.

doi:10.1186/2046-4053-3-4

Cite this article as: Beattie et al: Instruments to measure patient experience of health care quality in hospitals: a systematic review protocol. *Systematic Reviews* 2014 **3**:4.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)



## **4.6 Overview of Paper Three**

The following paragraphs provide additional details of the research process undertaken for the protocol development not covered in the associated publication, due to the limitations of the journal's word count. This is followed by a critical reflection of the work not included in the original article.

### **4.6.1 Clarifying the Concept**

Devising the protocol enabled the development of a clear plan to tackle the theoretical and methodological challenges inherent in conducting a systematic review to critique the utility of instruments to measure patient experience of hospital quality of care. For example, Chapter 1 explained the conceptual and methodological challenges of using satisfaction as an outcome measure and the cross-sectional survey in the ED (detailed in Chapter 3) confirmed this finding. However, developing the search strategy highlighted the necessity of including the term 'satisfaction', even though this was not the outcome of interest. Many relevant studies had been filed under 'satisfaction' within Medical Index Subject Heading (MeSH) hierarchies, even when their outcome of interest was patient 'experience.' To ensure retention of the most appropriate studies, the inclusion criteria also stipulated that only studies attempting to measure patient experience should be retained. The difference between patient 'experience' and 'satisfaction' also required explanation to enable other reviewers to apply the inclusion criteria consistently.

### **4.6.2 Application of Inclusion Criteria**

There is a necessity to ensure a robust procedure when determining which studies to include to ensure all appropriate studies are included. The evidence suggested that having an independent dual review of papers can improve the robustness of the application of study inclusion and reduce bias (McDonagh et al 2013). Given that there was no funding for the systematic review, it was impossible to obtain the necessary resource for a second reviewer to apply the inclusion criteria to all records (1,000), however, a 10% random sample of the records was feasible. It was proposed that the Cohen's kappa statistic would be calculated to determine the level of reliability between reviewers' decisions and aim for a high level of agreement ( $k > 0.8$ ). This seemed sensible as Cohen's kappa statistic calculates inter-rater agreement for categorical items, whilst taking account of the error by chance agreement (Streiner et al 2015).

#### **4.7 Critical Reflection of Paper Three**

On reflection, development of the protocol helped to provide a guide on how to conduct the systematic review. However, there were areas of the protocol that highlighted my novice level of knowledge of psychometrics and statistics. For example, stipulating that a low Cohen's kappa statistic for inter-rater agreement would result in a duplicate application of the inclusion criteria for all papers was a risky strategy. If reliability fell below 0.8, there was no resource to have the remaining 900 records checked by a second reviewer. Future solutions would be to ensure there is adequate resource for duplicate review steps, such as including adequate costing in funding applications. Alternatives could have been to acknowledge the increased error as a limitation of the study and determine sufficient numbers needed to achieve an acceptable level of agreement.

Determining the most appropriate method to critique the utility of instruments was complex for the following reasons:

- there was a need to critique the quality of methods, as well as the results of retained studies,
- reliability and validity are not 'all or nothing' concepts, making definitive decisions difficult,
- one instrument could have multiple studies, using different psychometric methods of testing, hence requiring multiple quality appraisal criteria,
- there are no established methods for critiquing wider aspects of utility, that is to say, educational impact, cost and acceptability, hence criteria had to be devised, tested and applied,
- there are no established methods for synthesising and presenting results for multiple studies of the same instrument.

The methods used to critique the quality of instruments in systematic reviews were the COSMIN criteria, which include various checklists and scoring systems to critique the quality of the methods of studies examining the validity and/or reliability of instruments (Mokkink et al 2006, Mokkink et al 2009, Mokkink et al 2010). Criteria also needed to be applied to assess the quality of the results of studies examining instrument validity and/or reliability. A member of the COSMIN group, Terwee et al (2007), has led the development of standards to determine 'cut off' points for positive and negative results of studies examining the validity and/or reliability of instruments. These standards were applied to studies retained within the systematic review. There were, however, no

standards to critique the wider aspects of utility, namely; educational impact, cost and acceptability, requiring further development following the protocol development. The systematic review would have likely benefited from detailing how to critique these additional aspects of utility at the protocol development stage.

At the protocol stage, the intention was to determine the relative strength of each of the five aspects of utility, depending on the instrument's primary purpose. However, following retrieval of studies, it became apparent that some instruments had specified two purposes, hence making decisions about the five aspects of utility impossible. Also, it was difficult to obtain agreement between other researchers in relation to decision-making, as 'relative importance' was, and is, a slippery concept, thus highlighting an unreliable process. Instead, the primary purpose of the instrument and details of each aspect of utility per instrument were presented in a utility matrix to enable users to have an overview of each instrument's properties to inform instrument selection. Although this was a deviation from the published protocol, the change was made explicit in the systematic review publication.

#### **4.8 Study Contribution to the Research Question**

Whilst, with hindsight, there were aspects of the protocol which could have been further developed, the protocol enabled the development of a plan to navigate most of the complexity of critiquing the utility of instruments measuring the patient experience of hospital quality of care. Feedback from journal reviewers encouraged clarity in devising the inclusion criteria and application of the methods to critique the quality of the studies. Developing the protocol helped to devise a robust approach to the systematic review. Also, learning about the complexity of psychometrics informed the development of a valid, reliable, yet brief instrument to measure the patient experience of hospital quality of care (discussed more fully in Chapter 5).

Exploring and understanding that instruments to measure the patient experience of hospital quality of care have different purposes provided further understanding about the potential research gap, highlighted in Chapter 1; that there is a disconnect between macro level (National) reports of hospital quality of care and the experiences of individual patients (Local or ward level). This disconnect may be due to different types of data being used for different purposes. For example, a national level survey of patient experience used for comparative league tables would need credible evidence of validity and be highly reliable, whilst tolerating high cost. However, an instrument being used for local quality improvement could potentially tolerate less reliability in favour of lower

cost and user acceptability. Each instrument is capturing different data at different levels of the healthcare system, which may account for the difference in reports. The systematic review then becomes not only about whether a valid and reliable instrument exists to measure the patient experience of hospital quality of care, but whether an instrument exists to measure the patient experience of hospital quality of care that is appropriate for the primary purpose of the research gap highlighted in this collection of works – quality improvement at a ward level for team feedback.

The remaining half of this Chapter embeds the published systematic review (Paper Four), followed by a critical reflection, before summarising the contribution to the research question.



## 4.9 Paper Four: Instruments to measure patient experience of healthcare quality in hospitals: a systematic review

Beattie *et al. Systematic Reviews* (2015) 4:97  
DOI 10.1186/s13643-015-0089-0



RESEARCH

Open Access

# Instruments to measure patient experience of healthcare quality in hospitals: a systematic review



Michelle Beattie<sup>1\*</sup>, Douglas J. Murphy<sup>2</sup>, Iain Atherton<sup>3</sup> and William Lauder<sup>4</sup>

### Abstract

**Background:** Improving and sustaining the quality of hospital care is an international challenge. Patient experience data can be used to target improvement and research. However, the use of patient experience data has been hindered by confusion over multiple instruments (questionnaires) with unknown psychometric testing and utility.

**Methods:** We conducted a systematic review and utility critique of questionnaires to measure patient experience of healthcare quality in hospitals. Databases (Medical Literature Analysis and Retrieval System (MEDLINE), Cumulative Index to Nursing and Allied Health Literature (CINAHL), Psychological Information (PsychINFO) and Web of Knowledge until end of November 2013) and grey literature were scrutinised. Inclusion criteria were applied to all records with a 10 % sample independently checked. Critique included (1) application of COSMIN checklists to assess the quality of each psychometric study, (2) critique of psychometric results of each study using Terwee *et al.* criteria and (3) development and critique of additional aspects of utility for each instrument. Two independent reviewers completed each critique. Synthesis included combining findings in a utility matrix.

**Results:** We obtained 1157 records. Of these, 26 papers measuring patient experience of hospital quality of care were identified examining 11 international instruments. We found evidence of extensive theoretical/development work. The quality of methods and results was variable but mostly of a high standard. Additional aspects of utility found that (1) cost efficiency was mostly poor, due to the resource necessary to obtain reliable samples; (2) acceptability of most instruments was good and (3) educational impact was variable, with evidence on the ease of use, for approximately half of the questionnaires.

**Conclusions:** Selecting the right patient experience instrument depends on a balanced consideration of aspects of utility, aided by the matrix. Data required for high stakes purposes requires a high degree of reliability and validity, while those used for quality improvement may tolerate lower levels of reliability in favour of other aspects of utility (educational impact, cost and acceptability).

**Systematic review registration:** PROSPERO CRD42013006754

**Keywords:** Systematic review, Patient, Experience, Satisfaction, Quality, Hospital, Acute care, Instruments, Questionnaires, Surveys, Utility

\* Correspondence: michelle.beattie@stir.ac.uk

<sup>1</sup>School of Health Sciences, Centre for Health Science, University of Stirling, Highland Campus, Old Perth Rd, Inverness IV2 3JH, UK

Full list of author information is available at the end of the article



© 2015 Beattie *et al.* This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

## Background

Despite an array of improvement initiatives in hospitals, the quality of care delivered remains open to question [12, 14, 18]. Patients who have experienced hospitalisation can offer unique insights into quality of care, which can be used for improvement. Yet, patients' views of quality of care are not always included in hospital measurement plans [4]. However, if patient experience data is to be used to improve quality of care in hospitals, it needs to be reliable and valid yet usable in practice [11, 54].

Measurement is fundamental to improving the quality of hospital care [43]. We will only know whether interventions or changes are working if processes and outcomes are measured. Measuring the patient experience in a robust way enables facts to be established from the complex phenomena of quality of care [32]. Patient experience data can be used to benchmark hospital performance, monitor effectiveness of interventions, establish hospital rankings and secure funding for research and innovation. Quantitative data can be combined with patient stories to create compelling evidence to evoke reflection and improvements within clinical teams [30]. Measuring the patient experience can highlight potential solutions, opportunities to improve hospital care.

Although a combination of tools is required to capture the complexity of hospital care, surveys are likely to remain the core method for measuring patient experience [11]. Surveys or questionnaires can be used to capture large samples of standardised data, which is essential if the patient perspective is to be equally represented alongside other aspects of care easier to quantify, such as waiting times.

There are, however, challenges to measuring the patient perspective of hospital care using questionnaires. Firstly, quality of care is difficult to quantify and define [5]. There is no widely accepted definition of quality of care; rather, there is an understanding that it is multi-dimensional, with varying interpretations dependent on who is being asked [16]. The widely accepted STEEEP acronym (Safety, Timeliness, Effectiveness, Efficiency, Equity and Person Centeredness) is most commonly used to describe the dimensions of quality of care [23]. There is consensus that quality of care consists of technical (knowledge and expertise) and interpersonal divisions (i.e. empathetic behaviour) [5, 16]. For example, the explanation of treatment options (technical) is improved if they are explained in an empathic and person-centred way (interpersonal).

Secondly, the terms 'satisfaction' and 'experience' are often used interchangeably despite their different meanings. Satisfaction is the gap between patient expectations and experience. Patients tend to overrate satisfaction, due to gratitude bias and other factors. Therefore, the validity and usefulness of satisfaction

data is limited; thus, there are calls for the patients' perspective of quality of care to focus on measuring experience, as opposed to satisfaction [31, 57, 58]. Patient experience is defined as things that happen to people and the extent that people's needs are met [17]. Questions are, therefore, designed around what actually occurred during hospitalisation. For example, a question might be asked as to whether or not patients received the right medication, at the right time as opposed to asking patients to rate their satisfaction with medicine administration. The emphasis is on asking patients whether or not, or how often, they have experienced certain care processes, rather than on rating aspects of care or treatment.

Thirdly, instruments need to be valid and reliable. That is, they accurately represent the patient experience of hospital care (validity), and this is measured consistently (reliability). An example of validity would be ensuring the patient experience is being measured, rather than the clinicians' perspective, as these are known to differ [16]. An unreliable tool would not be able to monitor improvement over time, consistently and without error.

Finally, instruments need to have high utility if they are to be used in real-world practice [3]. Van der Vleuten considered instrument utility from five aspects, namely validity, reliability, cost efficiency, acceptability and educational impact [52]. Each of these aspects is important to users of patient experience instruments. In the current financial climate, cost had become a key consideration when selecting an instrument. For example, obtaining a large, standardised sample will be expensive. Acceptability considers the suitability of the instrument from the users' perspective. This includes not only measuring a valid construct but also the tolerability of the instrument. For example, users (patients, clinicians and managers) may think a questionnaire has an unacceptably high number of questions, despite internal consistency (reliability) being improved by increasing the number of items [10]. Educational impact is also a factor to consider. How easy is it for an organisation, or individual within it, to drill down and make use of the data? Van der Vleuten emphasises the importance of weighing all of these aspects to select the right instrument, for the right purpose. For example, if survey results are to be used for high stakes (the outcome has important consequences for an individual or organisation), there is a necessity for high reliability, while tolerating high cost. Data used for team improvement may tolerate lower levels of reliability but require educational impact and acceptability.

This systematic review critiques the utility of published questionnaires aiming to measure the adult inpatient experience of hospital quality of care. The findings will aid appropriate instrument selection,

which will ultimately increase the likelihood of the patient's voice improving hospital quality of care.

### Study objectives

1. Identify questionnaires available to measure the adult inpatient experience of general (medical/surgical) hospital quality care.
2. Identify studies conducted to examine the measurement properties (validity and reliability) of questionnaires quantifying the adult inpatient experience of quality care.
3. Identify papers exploring the cost efficiency, acceptability and educational impact of questionnaires measuring the adult inpatient experience of hospital quality care.
4. Critique the quality of the methods and results of the measurement properties using recognised criteria for each instrument.
5. Determine the utility of each questionnaire by integrating results on the quality of validity, reliability, cost efficiency, acceptability and educational impact.

### Methods

Our methods were published in a protocol [4] prior to conducting the review, and this study was registered with PROSPERO (registration number CRD42013006754). A PRISMA (2009) Checklist aided the study design (see Additional file 1).

### Search strategy

Search strategies were devised, and the following databases were searched from inception until end of November 2013 as follows: Medical Literature Analysis and Retrieval System (MEDLINE), Cumulative Index to Nursing and Allied Health Literature (CINAHL) and Psychological Information (PsychINFO). No restrictions were applied to language, publication type or year. The word 'satisfaction' was included in our strategies, as some papers pertaining to 'experience' were filed under satisfaction within Medical Index Subject Headings (MeSH) within databases. Other literature was identified by contacting experts in the field and searching specialist websites (see Additional file 2 for MEDLINE search strategy and resources searched). Some e-mails were not responded to; we set a definitive deadline for response for July 2014. All records were exported into Ref Works for removal of duplicates and reference management. Duplicate removal was second checked within Ref Works and amended by hand by MB.

### Selection criteria

An inclusion selection form was applied to all titles and abstracts, enabling a transparent and focused selection of papers of interest: [4]

*Study type:* examining any measurement properties, theoretical development or utility of a questionnaire.

*Population:* adult in-patients, thus excluding clinicians, family members and paediatric perspectives.

*Setting:* surgical or medical care, thus excluding specialist areas, such as palliative and psychiatric care as patients in specialist areas have different determinants of what constitutes quality of care [38, 44].

*Global perspective:* patients' overall experience of hospital quality of care. Therefore, we eliminated condition-specific instruments and those measuring quality of specific professional groups.

*Construct of interest:* quality of care. We included all definitions or conceptualisations of quality, so long as they were defined from the patients' perspective. Studies measuring patient satisfaction were eliminated due to the theoretical and methodological limitations identified earlier.

Where decisions could not be made on title or abstract alone, full papers were retrieved. A second reviewer independently applied the inclusion criteria to a random 10 % of the records, retrieving full papers where necessary.

### Data extraction/instrument overview

We used a data extraction form to standardise the information recorded and aid analyses [31]. Some instruments have been considered by multiple studies; therefore, papers were grouped according to the instrument type to reduce duplication of data extraction. Data was extracted from the most recent version of the instrument only. All data extracted were checked for accuracy by a second, independent researcher.

### Assessment of study quality

The Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist was used to evaluate the methodological rigour of the studies [34, 51], and Quality Criteria for Measurement Properties [50] was used to critique the results of the studies. Studies were not rejected on the basis of this quality critique; rather, the results were synthesised to enable appropriate instrument selection.

The COSMIN checklists have been designed and validated for use in evaluating the rigour of psychometric studies of healthcare instruments [34]. The COSMIN checklist provides separate checklists (referred to as boxes) for each type of measurement property, for example, box A is for internal consistency, B for reliability and so forth. Boxes A–H are for different types of

psychometric studies and have their own associated quality questions. See Mokkink et al [34] for a full explanation of the COSMIN checklist. The checklists for interpretability and generalisability were not used as these are recommended for data extraction use only and are not scored for quality. All quality grading of studies were scored independently by two researchers (MB, DM) before reaching consensus.

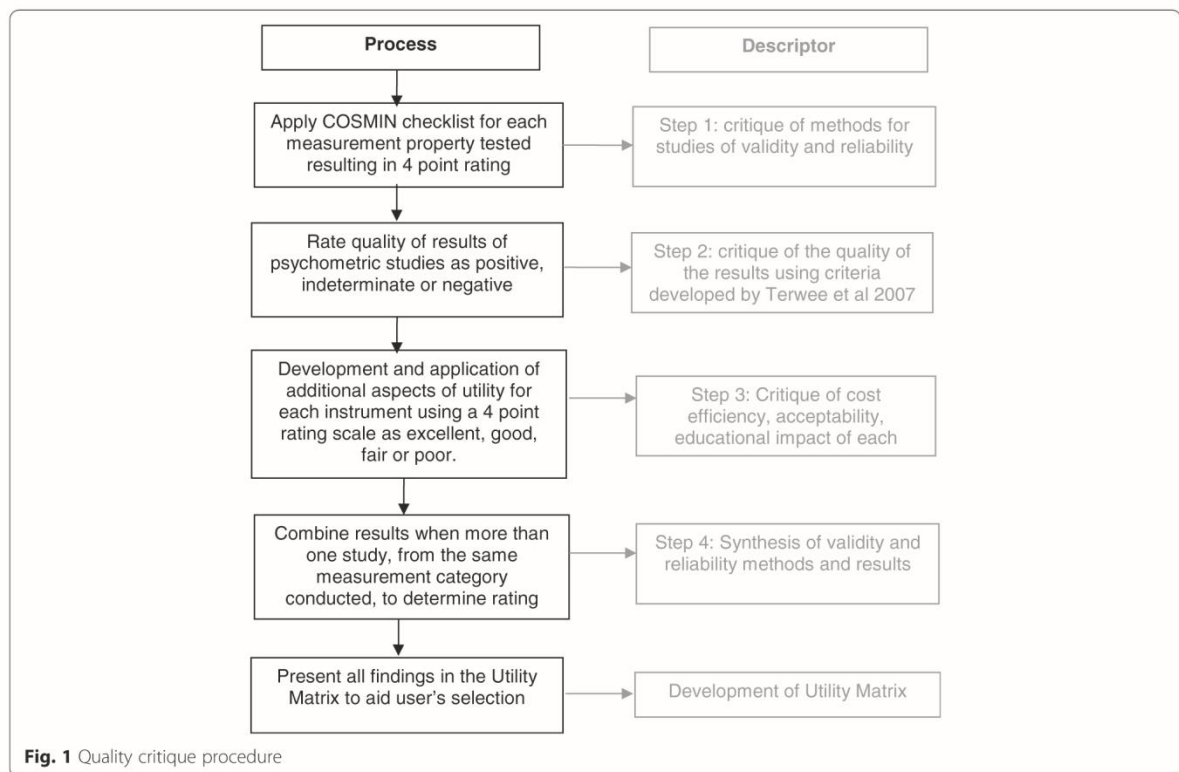
There were several steps in the quality critique of retained studies and instruments (Fig. 1 of quality critique procedure). Firstly, we applied the appropriate A–H checklist to critique the methodological quality of how each measurement property was being tested within each study. Responses within individual checklists were given a methodological score by applying the COSMIN four-point checklist scoring system. The scoring system is designed to ensure that items are scored as ‘excellent’ when there is evidence of adequate methodological quality, ‘good’ when relevant information is not fully reported but adequate quality can be assumed, ‘fair’ if the methodological quality is in doubt and ‘poor’ when there is evidence that the methodological quality is not adequate. Where answers to checklist questions were of variable ratings (i.e. some excellent, some poor), the overall score was determined

by taking the lowest rating of any item. In other words, the worst score counted [51].

Secondly, we rated the quality of the results of the psychometric studies by using the Quality Criteria for Measurement Properties devised by Terwee et al. (see Table 1) [50]. Results were rated as positive (+), indeterminate (?) or negative (-) according to the quality criteria for each measurement property. For example, positive ratings for internal consistency are given, using Terwee et al. criteria, if Cronbach’s alpha is  $\geq 0.70$ . Studies with Cronbach’s alpha results of  $< 0.70$  would be categorised as negative, or where Cronbach’s alpha was not determined, the result would be categorised as indeterminate. A full explanation, with justification for all COSMIN criteria results, is available from Terwee et al. [50].

#### Development of quality matrix

The COSMIN checklists only enable a critique of the validity and reliability aspects of utility; as a third step in devising a quality matrix, we developed additional questions to rate the cost efficiency, acceptability and educational impact of instruments (Table 2). Each question response has a four-point rating criteria of excellent, good, fair or poor.



**Table 1** Quality Criteria for Measurement Properties (Terwee et al. 2007) [50]

Property	Rating	Quality criteria
Reliability		
Internal consistency	(+)	(Sub)scale unidimensional AND Cronbach's alpha(s) $\geq 0.70$
	?	Dimensionality not known OR Cronbach's alpha not determined
	(-)	(Sub)scale not unidimensional OR Cronbach's alpha(s) $< 0.70$
Measurement error	(+)	MIC > SDC OR MIC outside the LOA
	?	MIC not defined
	(-)	MIC $\leq$ SDC OR MIC equals or inside LOA
Reliability	(+)	ICC/weighted Kappa $\geq 0.70$ OR Pearson's $r \geq 0.80$
	?	Neither ICC/weighted Kappa, nor Pearson's $r$ determined
	(-)	ICC/weighted Kappa $< 0.70$ OR Pearson's $r < 0.80$
Validity		
Content validity	(+)	The target population considers all items in the questionnaire to be relevant AND considers the questionnaire to be complete
	?	No target population involvement
	(-)	The target population considers all items in the questionnaire to be irrelevant OR considers the questionnaire to be incomplete
Construct validity		
Structural validity	(+)	Factors should explain at least 50 % of the variance
	?	Explained variance not mentioned
	(-)	Factors explain $< 50$ % of the variance
Hypothesis testing	(+)	Correlation with an instrument measuring the same construct $\geq 50$ % OR at least 75 % of the results is in accordance with the hypotheses) AND correlation with related constructs is higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	(-)	Correlation with an instrument measuring the same construct $< 50$ % OR $< 75$ % of the results is in accordance with the hypotheses OR correlation with related constructs is lower than with unrelated constructs

+ positive, - negative, ? indeterminate, AUC area under the curve, MIC minimal important change, ICC intraclass correlation, SDC smallest detectable change, LOA limits of agreement

Cost efficiency was rated in terms of the resources necessary to utilise the instrument for its primary purpose. The higher the resource/cost required, the lower the rating. Sample sizes detailed in instrument papers were used to answer the first question 'What are the number of observations (patients, raters, times) needed to reach the required level of reliability for the purpose of the instrument?' The number of observations needed to achieve the desired level of reliability is important to establish in terms of feasibility [35]. An instrument may be highly reliable but require extensive resource to obtain a reliable sample. Therefore, we are determining the resources necessary to achieve the level of reliability necessary for the instrument's primary purpose. For example, the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) instrument requires a minimum of 300 questionnaires per hospital to achieve a minimum of 0.8 reliability for all reported measures [20]. Also, if an instrument requirement was use on two or more

occasions to obtain reliability (i.e. test re-test reliability) where time affected the instrument performance, there would be a need to multiply the number of assessments by the given number of administrations.

Another question estimated the resource required to administer the questionnaire, for example, assessments requiring to be conducted by experts are more expensive in comparison to self-completion questionnaires. Completion time was also included; where developers had not published information on completion times, estimates were calculated by comparing with similar instruments. Question 4 brought together the preceding three questions on cost efficiency to estimate the cost of obtaining a reliable sample: minimal, moderate, considerable or extensive. These categories transformed into an inverse rating scale from poor to excellent, 'extensive', for example, becoming a rating of 'poor' for cost efficiency.

For the utility dimension of acceptability, questions were designed around evidence of the subjects'

**Table 2** Additional aspects of utility scoring criteria

	Excellent (****)	Good (***)	Fair (**)	Poor (*)
Questions for cost efficiency				
1. What are the number of observations (patients, raters, times) needed to reach the required level of reliability for the purpose of the instrument?	Only a small sample needed (<30)	A moderate sample size (30–49)	Not explicit but can be assumed or (50–99 assessments needed)	No details given or (≥100 assessments needed)
2. How long does an assessment take to complete	≤15 min	≤ 30 min	30–60 min	>60 min
3. What are the administrative costs of completing the assessment?	Easily embedded within existing resource. Little additional support required	Some administrative resource but no specialist resource required	Large amount of resource required to assess and administer	Significant specialist expertise and administrative time required to assess and administer
4. What is the cost to complete a reliable sample?	Minimal	Moderate	Considerable	Extensive
Questions for acceptability				
1. Is there evidence of subjects understanding of the instrument/assessment?	Investigations of subjects understanding (i.e. cognitive testing of instruments)	Estimated evidence of subjects understanding (i.e. high number of questions missed)	Subject understanding not explicitly stated but some can be assumed (i.e. student guide to OSCE)	No evidence of subject understanding
2. How many assessments are not completed?	There are low numbers of missing items (<10 %) and adequate response rates (>40 %)	There are a high number of missing items (≥10 %) and an adequate response rates (>40 %)	There are low numbers of missing items or poor (<10 %) and an inadequate response rate (<40 %)	There are high numbers of missing items (≥10 %) and poor response rates (<40 %)
3. Has the instrument/assessment been tested in an appropriate context?	Evidence of successful administration/use within an appropriate setting	Tested in vivo and changes recommended would be achievable	Testing in vivo and changes recommended would be difficult or only partially tested in vivo	Testing has only been conducted in vitro/simulation
Questions for educational impact				
1. There is evidence of the instruments intended purpose being achieved (i.e. if aim is to enable hospital ranking for patient selection, is there evidence that the results are actually influencing patient choice?)	Clear evidence of intended purpose being fulfilled	Explanatory or theoretical link between intended and actual use but no clear evidence	Evidence of theoretical work but relationship between intended and actual purpose poorly or not described	No evidence of intended purpose becoming actual
2. The scoring system is easily translated or available in an easy to use format?	Explicitly stated and easy to calculate	Explicitly stated but not easy to calculate	Scoring only calculated by resource with statistical knowledge	Scoring not explained well enough to calculate
3. The feedback from the results can be readily used for action where necessary?	Feedback is readily available in a format that enables necessary action	Feedback is readily available but not drilled down enough to enable targeted action	Minimal feedback available or delay results in limited impact	No explanation to determine adequacy of feedback. No direct feedback could be readily used without additional expertise

perception of the instrument, where less acceptance would result in a lower rating. There is an overlap between this category and content validity. However, the COSMIN checklist for content validity does not cover all aspects of user acceptability, e.g. cognitive testing. Also, some instruments may demonstrate content validity but have only been tested in a simulated environment or have an unacceptably high number of

questions. Grading was determined on a four-point rating scale of excellent, good, fair and poor. The overall rating of acceptability was determined by the worst score.

Questions for educational impact required evidence around an instrument's ease of use for learning or decision-making. Using a validated and reliable instrument is futile if not followed by action, learning or impact. This category

determines how easy it is to make use of the instrument results as intended. Again, question responses were graded using four rating responses, with the final rating determined by the worst score.

Where responses within individual categories of utility dimensions differed, the overall score was determined by the worst score counts, except for cost efficiency, where scoring was based on a balance of responses. Questions and categorised responses were refined following the testing of application to one instrument. Two researchers independently scored all papers and resolved disagreements through consensus.

#### **Beattie and Murphy instrument utility matrix**

All results were integrated into a utility matrix to aid instrument selection for users. The matrix enabled a synthesis of the quality of the methods used in the studies and results of all measurement properties from each study of each instrument, from the application of COSMIN and Terwee et al. criteria [50]. To simplify, the results from validity studies were merged into three headings: content, construct and criterion validity. Content validity included any study on the theoretical development of the instrument construction. Studies empirically testing any other type of validity, except criterion, were grouped together as construct validity. Construct validity is an overarching term for validity as opposed to a distinct form [10]. However, criterion validity was retained as a separate category as this is viewed as the 'gold standard', indicating the ability of an instrument to predict future outcomes, which would be of interest to those selecting an instrument.

Reliability was presented in the matrix in two categories: internal consistency and other forms of reliability. Internal consistency is the relationship between items and accounts for error generated by the questions or items asked by the instrument [49]. Measurement of internal consistency is only relevant when instruments have been designed from a reflective model. To determine whether instruments derived from a reflective model, we asked the question 'Do we expect all items to change when the construct changes?' If changes to the patient experience of quality of care did not result in changes in all domains, we classified the questionnaire as derived from a formative model. Also, measures of internal consistency are based on a single administration of the instrument and essentially represent the average of correlations among all the items in the instrument [49]. However, this does not account for the potential error between different observers or from one time interval to another. Generalizability G-theory and its associated decision D-studies can be used to further explore the reliability of an instrument and research the most effective blend of relevant resources

(times of administration, number of observers or raters) needed to explain error and attain reliability [20, 49]. To address the potential for misinterpreting an instrument as reliable when demonstrating high internal consistency but where other sources of error had not been examined, we added a question to the matrix to indicate whether or not all relevant sources of errors were investigated.

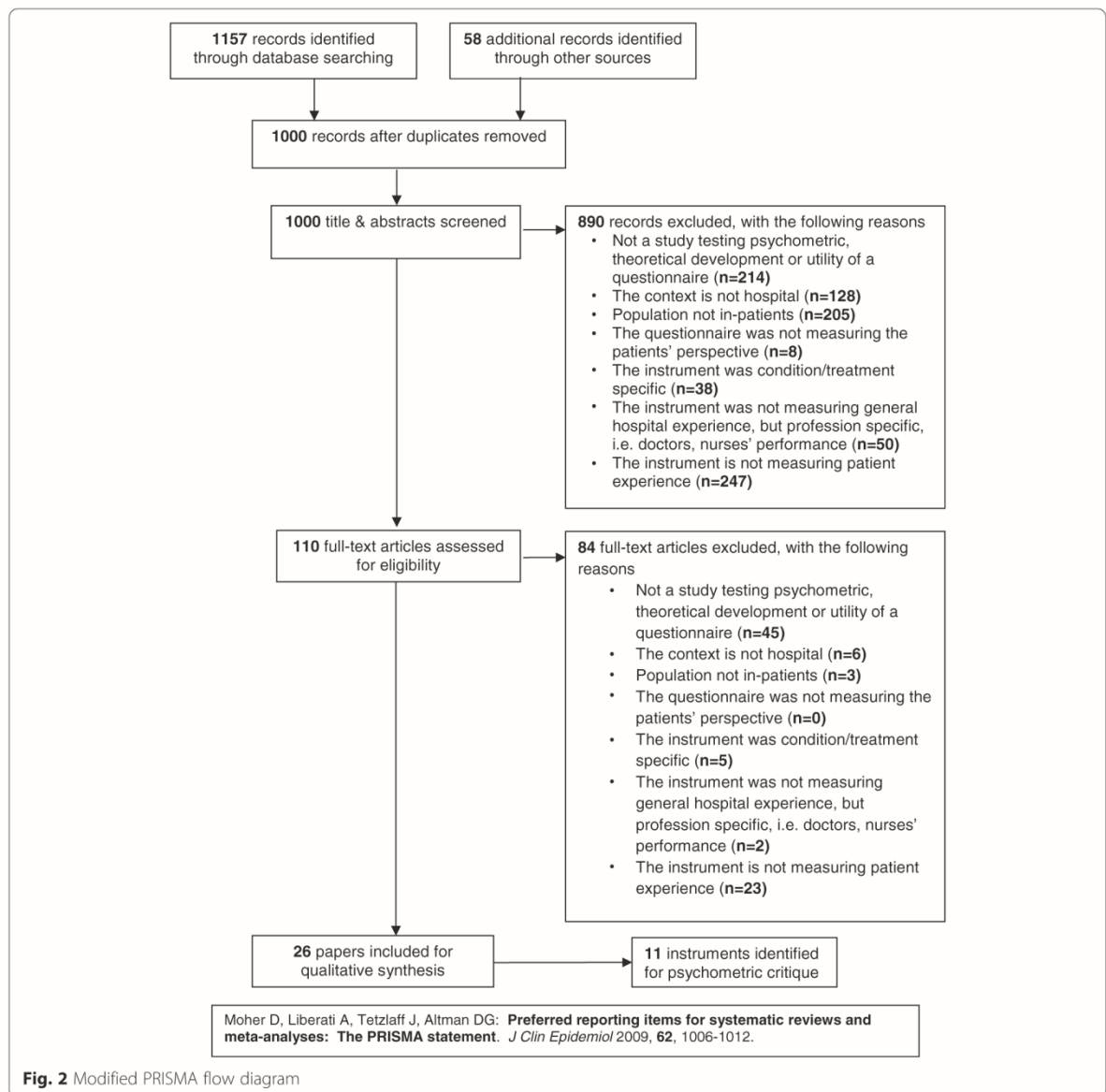
We presented ratings of study quality in star ratings: excellent (\*\*\*\*), good (\*\*\*) and fair (\*\*) and poor \* and the quality of results as positive (+), (?). Where more than one study from the same measurement category had been conducted, we determined the average point to rate the quality of the study methods. We provide two examples of combining validity and reliability scores to further explain. Example 1: if structural validity scored 'excellent' and cross-cultural validity scored 'fair', our overall rating would be 'good'. If, however, structural validity scored 'excellent' and cross-cultural validity scored 'good', we would rate validity overall as good to excellent (represented as \*\*\*/\*\*\*\*). Example 2: if the same instrument had two studies on reliability with study quality for one scoring 'excellent' and the other scoring 'good', we would rate reliability overall as good to excellent (represented as \*\*\*/\*\*\*\*). Where the quality of study results varied, within the same measurement property, we presented these as mixed. For example, if structural validity results scored positive and cross-cultural validity scored negative, we presented these as mixed (+/-).

## **Results**

Results of the search strategy were documented within the PRISMA flow diagram (see Fig. 2) [33]. We obtained 1157 records from our searches. Following removal of duplicates, 1000 records were screened for inclusion criteria. Application of the inclusion criteria to titles and abstracts resulted in the exclusion of 890 records. We retrieved 110 full-text articles where we were unable to make decisions from the title and abstract. Following application of inclusion criteria to full-text articles, we rejected 84 and retained 26 papers.

### **Screening results**

A second reviewer applied the inclusion criteria to a random 10 % of the 1000 papers ( $n = 100$ ). Where the second reviewer was unable to make a decision on title and abstract alone, full-text papers were retrieved ( $n = 17$ ). We rejected numerous papers where the outcome of interest, or theoretical model, was patient satisfaction, as opposed to patient experience (see Fig. 2 for specific exclusion results). The percentage of agreement between both reviewers was 90 %, therefore demonstrating a highly reliable process. Reviewers reached consensus following discussion on the remaining ten



papers. The process resulted in 26 papers being retained in relation to 11 instruments measuring the patient experience of hospital quality of care.

#### Characteristics of included instruments

The range of instruments and associated papers can be found in Table 3. Instruments were available across the World: Ethiopia (1), Hong Kong (1), India (1), Scandinavia (4), UK (3) and USA (1). Most instruments had generated multiple versions as they developed over time; therefore, we critiqued the most recent instrument version

and associated psychometric studies published in November 2013. For example, we used the Scottish Inpatient Patient Experience Survey (SIPES) measure version 2012 [46] as there is approximately a 1-year time lag between the instrument's use and results. Some instruments had extensive developmental histories, for example, the National Health Service Inpatient (NHSIP) Survey has been operating annually since 2002 [40], but its theoretical development work can be traced back to as early as 1991, to the original Picker Adult In-Patient survey [2, 9, 19]. We included the most recent works only. The Hospital



**Table 3** Instrument overview

Instrument/abbreviation	Associated papers	Country of origin	Domains covered	Conceptual framework	No. of items	Mode of administration	Timing of administration
Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS)	Sofaer et al. [48]	USA	Nurse communication	Reflective	24	Mail	48 h—6 weeks of discharge
	Keller et al. [26]		Doctor communication			Telephone	
	O'Malley [36]		Physical comfort			Mail with telephone follow-up	
	Levine et al. [29]		Pain control				
	Giordano et al. [20]		Medicine communication				
	Agency for Healthcare Research and Quality [1]		Discharge information			Interactive Voice Recognition (IVR)	
Quality from the Patients' Perspective (QPP)	Wilde et al. [56]	Sweden	Medical-technical competence	Reflective	68	Self-completion questionnaire	At discharge
	Wilde et al. [55]		Physical technical conditions				
	Larsson et al. [28]		Personal necessities				
			Characteristics				
Quality from the Patients' Perspective Shortened (QPPS)	Larsson et al. [27]	Sweden	Identity-orientated approach	Reflective	24	Self-completion questionnaire	At discharge
			Situation				
			Participation				
			Commitment				
			Socio-cultural atmosphere				
Positive treatment of significant others							
Picker Patient Experience Questionnaire (PPE-15)	Jenkinson et al. [25]	England	Information and education	Reflective	15	Self-completion postal questionnaire	Within 1 month of discharge
	Jenkinson et al. [24]		Coordination of care				
	Reeves et al. [42]		Physical comfort				
			Emotional support				
			Respect for patient preferences				
			Involvement of family and friends				
Continuity and transition							

**Table 3** Instrument overview (*Continued*)

NHS Inpatient Survey (NHSIP)	Boyd [6] Sizmur and Redding [47] Picker Institute Europe [40] Decourcy et al. [13]	England (originated in the USA)	Overall impression Admission to hospital The hospital and ward Doctors Nurses Your care and treatment Operations and procedures Leaving hospital	Formative	70	Postal survey	Between 4 and 5 months of discharge
Scottish Inpatient Patient Experience Survey (SIPES)	Scottish Government [46] Scottish Government [45]	Scotland	Admission to hospital The hospital and ward Care and treatment Hospital staff Arrangements for leaving hospital Care and support services after leaving hospital	Formative	30	Postal survey, questionnaire (also available online, by telephone and via text phone)	Between 4 and 5 months of discharge
Hong Kong Inpatient Experience Questionnaire (HKIEQ)	Hospital Authority [22] Wong et al. [59]	Hong Kong	Prompt access Information provision Care and involvement in decision-making Physical and emotional needs Coordination of care Respect and privacy Environment and facilities Handling patient feedback Overall care of health professionals and quality of care	Reflective	62	Mixed 92 % interviewed by telephone 8 % face-to-face home interviews	48 h—1 month after discharge
Patient Experience Questionnaire (PEQ)	Pettersen et al. [39]	Norway	Information future complaints Nursing services Communication Information examinations Contact with next-of-kin Doctor services Hospital and equipment Information medication Organisation General satisfaction	Reflective	35	Postal self-completion questionnaire	6 weeks after discharge

**Table 3** Instrument overview (Continued)

Norwegian Patient Experience Questionnaire (NORPEQ)	Oltedal [37]	Norway	Whether doctors were understandable Doctors professional skills Nurses professional skills Nursing care Whether doctors and nurses were interested in the patients problems Information on tests	Reflective	8	Self-completion	Within 7 weeks of discharge
Patient Experiences with Inpatient Care (I-PAHC)	Webster et al. [53]	Ethiopia	Nurse communication Doctor communication Physical environment Pain management Medication and symptom communication	Reflective	16	Interviewer-assisted completion	After first day of admission
Patient Perceptions of Quality (PPQ)	Rao et al. [41]	India	Medicine availability Medical information Staff behaviour Doctor behaviour Hospital infrastructure	Reflective	16	Interviewer-assisted completion	Not specified

Consumer Assessment of Healthcare Providers and Systems (HCAHPS) originated in 2002 [1], but we used version 2012 [2].

Instruments covered similar domains to capture the patient experience of their hospital care. Some focused on stages of the patient's journey, from admission to hospital discharge [6, 46]. Others were structured around dimensions of hospital quality, i.e. communication and coordination of care, such as HCAHPS [2] and Quality from the Patients' Perspective Shortened (QPPS) [56]. All instruments covered aspects of technical and interpersonal components of quality of care. There were some cultural differences in content. For example, the Patient Perceptions of Quality (PPQ) [41] included questions around medicine availability, reflective of the low-income context in which the instrument was tested. Importantly, all instruments were measuring the patient experience, as opposed to satisfaction.

Most instruments were devised from a reflective model (see Table 3). That is to say, collectively, factors within the questionnaire reflect the construct of interest, patient experience of hospital quality. For example, changes made to improve the quality of hospital care (construct) would likely cause variation in all indicators, i.e. safety and person centeredness

within these instruments. The NHSIP and SIPES instruments were exceptions, based on a formative model. Domains within their questionnaire were designed around the patient journey, i.e. from admission to discharge home. A poor experience during admission to hospital (indicator) would decrease the patient's score of quality of care, but not necessarily influence other indicators, i.e. the patient's experience of hospital discharge.

The number of items within the instruments varied from 8 to 70, excluding demographic questions. All instruments were self-completed instruments, except Patient Experiences with Inpatient Care (I-PAHC) and PPQ which required interviewer assistance due to the prevalence of illiteracy in the countries in which they were tested [41, 53]. Most instruments were mailed, although some offered telephone assistance (HCAHPS, SIPES, NHSIP) and HCAHPS was available in several formats (mail only, telephone only, mail followed by telephone and interactive voice response) [8].

All instruments were administered following discharge from hospital, except I-PAHC which was completed any time during the admission, but after the first day of hospitalisation [53]. Timings varied, from instruments being distributed on discharge to several months following hospitalisation.

### Instrument quality and results

The type and quality of the methods and results of the psychometric studies was variable but mostly of a high standard (see Table 4). Every instrument had evidence of examining at least one aspect of validity and reliability.

#### Validity

Content validity was tested for all instruments by exploring which aspects of hospital quality care mattered most to patients. Scores for content validity were rated as good or excellent, except for HCAHPS [48]. HCAHPS was rated as poor as no information was provided to determine whether aspects of quality suggested by patients had been integrated within their instrument, as well as patients having concurred with pre-determined items. While the quality of the methodology and results was limited for HCAHPS, in all other instruments, the questionnaire items were relevant and sufficient, therefore rating positive for content validity.

All instruments had examined other types of validity, except NHSIP and SIPES. Comments in NHSIP documentation referred to previous structural validity, but the detail required to judge criteria was unavailable [47]. Criterion validity is considered when an instrument is compared with a gold standard. While no gold standard exists for measures of patient experience, the COSMIN criteria include comparisons of shortened with original longer versions as criterion validity. Three studies comparing shortened versions with their original longer versions (QPP [55], QPPS [27], PPE-15 [24, 25]), rated fair, excellent and good, respectively, with positive results. Some developers had tested the validity of their instrument extensively, namely QPP, HKIEQ and NORPEQ which had conducted three or more validity studies. The methodological quality of all construct validity studies was mostly good or excellent (HCAHPS), except HKIEQ. [22] HKIEQ was rated as fair as no description was given on how the authors handled missing items within their study. Most results of construct validity were categorised as positive, as factor analysis explained at least 50 % of the variance or met other Quality Criteria for Measurement Properties identified by Terwee et al. (see Table 1) [50]. Several studies were rated as indeterminate as they did not meet the Quality Criteria for Measurement Properties' results. For example, structural validity was thoroughly examined for the HCAHPS instrument but was categorised as indeterminate as structural equation modelling does not report factor loadings [26]. This result needs to be interpreted with caution as the HCAHPS study demonstrated an excellent fit for structural validity. The methodological quality of criterion validity for the QPP instrument was rated as poor as there were flaws identified in the study

design [55]. The validity of one QPP study [55] was in doubt as student nurses were given scenarios to act as simulated patients to answer questionnaire items in the instrument.

#### Reliability

All instruments studied internal consistency to determine the interrelatedness among items. All instruments achieved positive internal consistency results, except NHSIP [47] which was indiscriminate as Cronbach's alpha was not determined. Importantly, two instruments [45, 47]. were derived from formative models and did not have unidimensional subscales, which is reflected in their indiscriminate results and lower quality findings [25, 47]. However, the quality of the study methods for five instruments (NHSIP [47], SIPES [45], HKIEQ [22, 59], PEQ [39] and NORPEQ [37]) did not clarify how missing items were handled. Four instruments examined types of reliability in addition to internal consistency (HCAHPS [26], HKIEQ [22], PEQ [39] and NORPEQ [37]). All had positive results, but one HCAHPS study was indeterminate as the minimal important change was not determined as per the Quality Criteria for Measurement Properties (as detailed in Table 1).

#### Results of instrument utility

The cost efficiency was rated as good for QPPS [27], NORPEQ [37] and I-PAHC [53]. All other instruments were rated as poor or fair, highlighting that considerable or extensive resource would be required to obtain an adequate sample (see Table 5). All instruments, except QPP, were rated excellent or good for the dimension of acceptability, as there was evidence of user acceptability in an appropriate context. QPP was rated as fair due to the evidence of testing in a simulated setting only [56].

Educational impact was good for five of the instruments (HCAHPS [26, 29, 48], SIPES [45, 46], NORPEQ [37], I-PAHC [37], PPQ [53]) as there was evidence of the instruments being easily used for their intended purpose, i.e. hospital ranking or quality improvement. Five instruments (QPP [55], QPPS [27], PPE-15 [25], NHSIP [13, 40], HKIEQ [22]) were rated as fair as there was some evidence of educational impact, and PEQ was rated as poor as there was no enough information to determine educational impact.

#### Utility matrix results

All results (critique of methods, results and additional aspects of utility) were embedded in our utility matrix to enable an easy overview and aid instrument selection (see Table 6). We found two main purposes of patient experience instrument use to compare performance across hospitals and local quality improvement. Overall,

**Table 4** Quality of methods and results of psychometric studies

Instrument/abbreviation	Associated papers	Measurement property	Result	Quality rating of results	Quality rating of methods
Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS)	Sofaer et al. [48]	Content validity	Patients considered other aspects of hospital care which appear to have not been included	Negative	Poor
	Keller et al. [26]	Internal consistency	Cronbach's alpha 0.70	Positive	Excellent
	Keller et al. [26]	Reliability	ICC 0.70	Positive	Excellent
	Keller et al. [26]	Structural validity	7 categories for 16 items. Factor loadings 0.57–91. Uniqueness of error reported	Indeterminate	Excellent
	O'Malley [36]	Measurement error	Correlation between same composites different services Surgery 0.76 Obstetrics 0.73 Medical 0.85	Indeterminate	Good
Quality from the Patients' Perspective (QPP)	Wilde et al. [56]	Content validity	35 patient interviews—development of relevant questionnaire	Positive	Excellent
	Wilde et al. [55]	Internal consistency	Cronbach's alpha 0.80	Positive	Excellent
	Wilde et al. [55]	Content validity	High patient ratings of item clarity and comprehensiveness	Positive	Excellent
	Wilde et al. [55]	Structural validity	Factor solutions Medical/technical competence 50.4 % Physical/technical conditions 44.8 % Identity-orientated approach 66.9 % Socio-cultural atmosphere 65.8 %	Positive	Good
	Wilde et al. [55]	Criterion validity	Correlation between long and short version in their entirety was 0.90	Positive	Poor
	Larsson et al. [28]	Structural validity	RMSEA of 0.050 was obtained indicating the model was an acceptable fit	Indeterminate	Good
	Quality from the Patients' Perspective Shortened (QPPS)	Larsson et al. [27]	Internal consistency	Cronbach's alpha 0.74 for overall scale	Positive
Larsson et al. [27]		Criterion validity	Pearson correlation coefficients all results statistically significant 0.0025 when Bonferroni corrections made	Positive	Excellent
Picker Patient Experience Questionnaire (PPE-15)	Jenkinson et al. [25]	Internal consistency	Cronbach's alpha 0.8	Positive	Good
	Jenkinson et al. [24]	Internal consistency	0.89 for 4 pages 0.87 for 12 pages	Positive	Excellent
	Reeves et al. [42]	Content validity	Focus groups, cognitive testing, amendments—research did not identify any missing items from patients' perspective	Positive	Excellent
	Jenkinson et al. [25]	Criterion validity	Correlations between short and long version between 0.93 ( $P < 0.001$ ) and 0.95 ( $P < 0.001$ )	Positive	Good
	Jenkinson et al. [24]	Hypothesis testing	Item correlations were above recommended levels for all PPE items in both survey versions (0.37–0.61)	Positive	Excellent
NHS Inpatient Survey (NHSIP)	Boyd [6]	Content validity	Tested and modified with group of inpatients	Positive	Excellent
	Sizmur and Redding [47]	Internal consistency	Item correlations given but Cronbach's alpha not reported	Indeterminate Positive	Fair Excellent

**Table 4** Quality of methods and results of psychometric studies (Continued)

Scottish Inpatient Patient Experience Survey (SIPES)	Scottish Government [45]	Content validity	Extensive work with patient groups: survey, focus groups, stakeholder consultations, cognitive testing. Findings, the patient found the items relevant and comprehensive		
	Scottish Government [45]	Internal consistency	Cronbach's alpha over 0.70 for each survey section	Positive	Poor
Hong Kong Inpatient Experience Questionnaire (HKIEQ)	Hospital Authority [22]	Internal consistency	Cronbach's alpha 0.75 for overall scale	Positive	Fair
	Hospital Authority [22]	Reliability	Intraclass correlation 0.42–0.96 and test re-test 0.78	Positive	Fair
	Hospital Authority [22]	Content validity	Participants found the questionnaire to be clear, understandable, and appropriate	Positive	Excellent
	Hospital Authority [22]	Structural validity	17 factors explained 74 % of the variance	Positive	Fair
	Wong et al. [59]	Internal consistency	Cronbach's alpha 0.75 for overall scale	Positive	Fair
	Wong et al. [59]	Structural validity	18 factors explained 75.5 % of the variance	Positive	Fair
	Hospital Authority [22]	Cross-cultural validity	Translated but not cross-culturally validated	Indeterminate	Fair
	Patient Experience Questionnaire (PEQ)	Pettersen et al. [39]	Internal consistency	Cronbach's alpha greater than 0.70 for overall scale	Positive
Pettersen et al. [39]		Reliability	Test re-test 0.62–0.85 with ICC exceeding 0.7	Positive	Fair
Pettersen et al. [39]		Content validity	Grouped more than 600 m written comments and held focus groups with previous inpatients to ensure relevant and sufficient items were covered	Positive	Good
Pettersen et al. [39]		Structural validity	20 items, 6 factors accounted for 67 % total variance	Positive	Excellent
Pettersen et al. [39]		Hypothesis testing	Associations between rating scale and external measures, i.e. gender, age, fulfilment of expectations. Only mean differences computed	Indeterminate	Poor
Norwegian Patient Experience Questionnaire (NORPEQ)	Olteidal [37]	Internal consistency	Item correlation 0.59–0.71 and Cronbach's alpha 0.85	Positive	Fair
	Olteidal [37]	Reliability	Intraclass correlation 0.45–0.79 and test re-test 0.88	Positive	Good
	Olteidal [37]	Content validity	Patient interviews found questions and scaling easy to understand and all relevant questions covered	Positive	Good
	Olteidal [37]	Structural validity	6 items explained 57.7 % variance	Positive	Good
	Olteidal [37]	Construct validity	Hypothesised scales scores would correlate 0.6–0.8 with satisfaction (correlation significant, range from high to low)  Scale scores would correlate 0.4–0.6 perceptions of incorrect treatment (moderate result)  Scores would correlate 0.1–0.3 with patient health and physical health. (Result 0.19–0.27)	Positive	Good
Patient Experiences with Inpatient Care (I-PAHC)	Webster et al. [53]	Internal consistency	Cronbach's alpha >0.78	Positive	Excellent
	Webster et al. [53]	Content validity	Focus groups, revisions by stakeholders, translated, cognitively tested and patient groups reported clear questions covering all aspects important to them	Positive	Excellent

**Table 4** Quality of methods and results of psychometric studies (*Continued*)

Patient Perceptions of Quality (PPQ)	Webster et al. [53]	Structural validity	Kept if item loadings greater than 0.40. Variance not reported	Indeterminate	Excellent
	Webster et al. [53]	Construct validity	5 factors with loadings 0.48–0.86. Results in accordance with priori hypothesis	Positive	Excellent
	Webster et al. [53]	Cross-cultural validity	Translation done but not empirically tested	Indeterminate	Fair
	Rao et al. [41]	Internal consistency	Cronbach's alpha >0.70	Positive	Excellent
	Rao et al. [41]	Content validity	Questionnaire devised from qualitative interviews with patients	Positive	Excellent
	Rao et al. [41]	Structural validity	5 dimensions explained 73 % variance	Positive	Excellent

HCAHPS, NORPEQ, PPE-15 and I-PAHC demonstrated the most evidence that their instruments were valid and reliable. NHSIP and SIPES demonstrated the least evidence of validity and reliability. All other instruments were found to have a degree of psychometric evidence. The most cost-effective instruments were QPPS, NORPEQ and I-PAHC. All instruments demonstrated good or excellent acceptability, except QPP. Several instruments (HCAHPS, SIPES, NORPEQ, I-PAHC and PPQ) were rated as good for educational impact.

### Discussion

To our knowledge, this is the first systematic review to identify and critique the utility of instruments aiming to measure patient experience of hospital quality. We found 11 international instruments measuring the patient experience of hospital care, while we dismissed numerous measuring patient satisfactions. We critiqued utility from a wide perspective, using international standards where they were available and devising additional criteria where needed.

Reassuringly, all instruments reported some psychometric testing and published information on other aspects of utility. Similar literature reviews have found that studies do not report sufficient psychometric information to enable critique, although this has improved over the last 10 years [7, 21]. We found enough reported psychometric information to critique the retained instruments, although some missing data may have resulted in studies being apportioned lower scores for study quality.

Of course, validity and reliability are not 'all or nothing' concepts; rather, they are a matter of degree. Evidence of validity tends to be cumulative, as each new study provides further confirmation of the ability of an instrument to measure patient experience of hospital quality care. As validation develops over time, it is important not to dismiss newer instruments with only some validation. The reliability of an instrument is also strengthened over time as developers refine the tool and identify ways in which to reduce the margin of error,

such as the establishment of a training manual and, of course, developments in psychometrics.

While the longevity of instruments is an identified strength, there should also be a note of caution. Well-established instruments may rely on historical data to establish theories and concepts of quality of hospital care. What constitutes Quality from the Patients' Perspective is likely to shift over time [4]; therefore, we suggest that elements of hospital care which are important to patients are re-explored at least every few years, to re-ensure continued instrument validity. We also found evidence of items being added to instruments to fit the current healthcare policy context [6, 46]. While this seems reasonable, there is a risk that an instrument becomes a measure of healthcare policy implementation as opposed to measuring the patient experience of the quality of hospital care. Conducting interviews or surveys to assess the impact of additional items addressing policy aims should also ensure that such changes do not alter the overall validity of questionnaire content from the patient's perspective. We found extensive work in terms of theoretical and conceptual development of instruments in this area, which is necessary for an elusive and evolving concept of quality of health care.

We found no studies assessing the ability of an instrument to detect change over time in the construct to be measured, otherwise known as responsiveness [15]. This was surprising given that one of the main uses of patient experience instruments is to measure hospital care quality for evaluation of local improvement work. This review highlights both the need for and the current gap in studies assessing responsiveness of these instruments.

This systematic review highlights that there is no 'one-size-fits-all' approach in selecting an instrument to measure the patient experience of hospital quality of care. Rather, there are a range of instruments available with varying strengths and limitations of instrument utility. Instrument choice will, therefore, be dependent upon a number of factors, specifically the purpose for

**Table 5** Results of additional aspects of utility

	HCAHPS		QPP		QPPS		PPE-15		NHSIP	
	F	R	F	R	F	R	F	R	F	R
<b>Cost efficiency</b>										
1. What are the number of observations (patients, raters, times) needed to reach the required level of reliability for the purpose of the instrument?	≥300 [20]	Poor	Not reported	Poor	Not reported	Poor	330 per group [24]	Poor	Not Reported	Poor
2. How long does an assessment take to complete?	8 min [8]	Good	30 min [28]	Good	≤15 min	Excellent	12 min [42]	Excellent	20 min (estimate)	Good
3. What are the administrative costs of completing the assessment?	V large numbers and expertise [8]	Poor	Considerable [28]	Fair	Brief and easy scoring [27]	Excellent	Large no. and standardised data	Fair	Large no. and standardised	Fair
4. What is the cost to complete a reliable sample?	Extensive	Poor	Considerable	Fair	Minimal	Good	Considerable	Fair	Extensive	Poor
Overall Rating	POOR		FAIR		GOOD		FAIR		POOR	
<b>Acceptability</b>										
1. Is there evidence of subjects understanding of the instrument/assessment?	Yes [29,48]	Excellent	Yes [55]	Excellent	Yes [8]	Excellent	Yes [42]	Excellent	Yes [47]	Excellent
2. How many assessments are not completed?	25 % miss RR 47 %	Good	13 % miss RR [55]	Good	25 % miss RR [55]	Good	29 % miss RR [42]	Good	No info [47]	Good
3. Has the instrument/assessment been tested in an appropriate context?	Yes [26]	Excellent	Tested in simulation [55]	Fair	Yes [55]	Good	Yes [42]	Excellent	Yes [47]	Excellent
Overall Rating	Good		Fair		Good		Good		Good	
<b>Educational impact</b>										
1. Is there evidence of the instrument being used for its intended purpose? (i.e. if aim is to provide hospital ranking for patient selection, is there evidence that the results are influencing patient choice?)	Evidence of purpose [20]	Excellent	Discussion of purpose but no evidence [55]	Fair	Discussion of purpose but no evidence [27]	Fair	Explanatory use for national comparison	Good	Clear evidence of purpose [47]	Excellent
2. Is the scoring system easily translated or available in an easy to use format?	Easy scoring	Excellent	Easy scoring	Excellent	Easy scoring	Excellent	Easily scored	Excellent	Statistical knowledge	Fair
3. Can the results be readily used for action where necessary?	Available but not at unit/team level	Good	Results actionable at local level	Excellent	Results actionable at local level	Excellent	Adjustments needed (Jenkinson comparison)	Fair	Expertise required to enable local action	Fair
Overall Rating	Good		Fair		Fair		Fair		Fair	



**Table 5** Results of additional aspects of utility (Continued)

	SIPE		HKIEQ		PEQ		NORPEQ		I-PAHC		PPQ	
	F	R	F	R	F	R	F	R	F	R	F	R
<b>Cost efficiency</b>												
1. What are the number of observations (patients, raters, times) needed to reach the required level of reliability for the purpose of the instrument?	Variable but >100	Poor	300–500 [45]	Poor	Not specified	Poor	Not specified	Poor	≥230 [53]	Poor	Not specified	Poor
2. How long does an assessment take to complete?	20 min [46]	Good	25 min [59]	Good	<30 min (estimate)	Good	>15 min (estimate)	Excellent	15 min [53]	Excellent	<30 min (estimate)	Good
3. What are the administrative costs of completing the assessment?	V large numbers and expertise	Poor	V large numbers and expertise	Poor	Considerable	Fair	Brief and simple scoring	Good	Interviewers required	Fair	Interviewer required [41]	Fair
4. What is the cost to complete a reliable sample?	Extensive	Poor	Extensive	Poor	Considerable	Fair	Moderate	Good	Moderate	Good	Considerable	Fair
Overall Rating	POOR		POOR		FAIR		GOOD		GOOD		FAIR	
<b>Acceptability</b>												
1. Is there evidence of subjects understanding of the instrument/assessment?	Yes [45]	Excellent	Yes [22]	Excellent	Yes [39]	Excellent	Yes [37]	Excellent	Yes [53]	Excellent	Yes [41]	Excellent
2. How many assessments are not completed?	No info RR 50 % [13]	Good	21 % miss RR 49 % [22]	Good	>10 % miss RR 53 % [39]	Excellent	42.5 % miss RR 48 % [37]	Excellent	High No RR 95 % [53]	Good	0 % miss RR 85 % [41]	Excellent
3. Has the instrument/assessment been tested in an appropriate context?	Yes [45]	Excellent	Yes [22]	Excellent	Yes [39]	Excellent	Yes	Excellent	Yes	Excellent	Yes	Excellent
Overall Rating	Good		Good		Excellent		Excellent		Good		Excellent	
<b>Educational impact</b>												
1. Is there evidence of the instrument being used for its intended purpose? (i.e. if aim is to provide hospital ranking for patient selection, is there evidence that the results are influencing patient choice?)	Explanatory use for national comparison [45]	Good	Explanatory use for national benchmarking [22]	Good	Clear evidence of purpose [39]	Excellent	Explanatory use described [37]	Good	Explanatory use described [53]	Good	Explanatory use described [41]	Good
2. Is the scoring system easily translated or available in an easy to use format?	Easy colour coding	Excellent	Statistical expertise	Fair	Not explained	Poor	Easy scoring	Excellent	Easy scoring	Excellent	Easy scoring	Excellent
3. Can the results be readily used for action where necessary?	Results at hospital level	Good	Results at hospital level	Good	No information	Poor	Readily available	Excellent	Readily available	Excellent	Readily available	Excellent
Overall Rating	Good		Fair		Poor		Good		Good		Good	
<b>F findings, R ratings</b>												

**Table 6** Results of Beattie and Murphy instrument utility matrix

Instrument	Primary purpose	Validity			Reliability			Cost efficiency Rating	Acceptability Rating	Educational impact Rating
		Content/ theoretical development	Construct (structural, cross- cultural)	Criterion validity	Internal consistency	Other reliability	Was the correct error source investigated?			
HCAHPS	National comparisons	*(-)	****(?)		****(+)	***/***(?)	Y	*	***	***
QPP	Quality improvement	****(+)	***(?)	*(+)	****(+)		Y	**	**	**
QPPS	Quality improvement			****(+)	****(+)		P	***	***	**
PPE-15	National performance indicators	****(+)	****(+)	***(+)	***/***(+)		P	**	***	**
NHSIP	National performance indicators	****(+)			**(?)		N	*	***	**
SIPES	National comparisons	****(+)			*(+)		N	*	****	***
HKIEQ	National comparisons	****(+)	**(?)		**(+)	**(+)	Y	*	***	**
PEQ	Quality improvement and national surveillance	***(+)	**/***(?)		**(+)	**(+)	Y	**	****	*
NORPEQ	Cross-national comparisons in Nordic countries	***(+)	***(+)		**(+)	***(+)	Y	***	****	***
I-PAHC	Quality improvement in low-income settings	****(+)	***/***(?)		****(+)		P	***	***	***
PPQ	Local quality improvement	****(+)	****(+)		****(+)		P	**	****	***

Ratings of study quality: \*poor, \*\* fair, \*\*\*good, \*\*\*\*excellent. Ratings of measurement results: (+) positive rating, (-) negative rating, (?) indeterminate rating, (?) mixed. Correct source of error: Y yes, N no, P partial

which the data will be used, available resource and local context. For example, where an instrument is to be used for high stakes purposes (perhaps attached to a financial incentive, public league tables or an outcome measure in a research study), an instrument with high reliability should be selected, such as HCAHPS. However, high costs in terms of resource would need to be accepted as HCAHPS requires compliance with standardised sampling, data collection and statistical expertise to analyse the data. Alternatively, if an instrument is to be used to measure the effectiveness of local quality improvement work, then QPPS may be the instrument of choice, as it rated good for user acceptability and cost efficiency. Similarly, but in a low-income setting, I-PAHC could be a useful instrument as it has scored 'good' and 'excellent' in all dimensions of instrument utility. Also, brief instruments, such as QPPS or PPE-15, may be used as

screening instruments to determine a sample for more detailed exploration.

Context is also important, particularly in relation to theoretical development and content validity. For example, if work has been carried out to determine what quality of hospital care means to a local population, as with SIPES in Scotland, then this would be the instrument of choice in Scotland in terms of its content validity. Where instruments are utilised in other countries, studies of cross-cultural validity should be conducted before instrument use.

As with all literature reviews, our findings are dependent upon the quality of detail available in the published literature. There are risks that unpublished instruments have been missed. While our literature search did not include the EMBASE database for pragmatic reasons, we did conduct a thorough search of MEDLINE,

CINHAL and PsychINFO, as well as specialist databases in the field of patient experience. We also acknowledge that only 10 % of the inclusion criteria was independently checked by two reviewers. Despite checking secondary references, we found no other instruments meeting our inclusion criteria.

Also, there is a possibility that included instruments have been harshly critiqued. We used the COSMIN criteria which reduces scores for methodological quality when insufficient information is available and applies the 'lowest score counts' for an overall score [3]. Some psychometric studies may have only been rated as poor or fair on one item response, subsequently giving a low overall rating. However, a design strength of the COSMIN four-point rating scale was to ensure that only fatal flaws are categorised as poor. Therefore, some item responses cannot be categorised as poor. For example, some checklists determine whether or not the percentage of missing items was reported. Responses are either 'yes' or 'no'. A response of 'no' could still achieve a 'good' quality rating as this question did not offer a 'poor' response option. While having missing items is not regarded as good practice, COSMIN developers determine that the overall quality of the study could still be good or excellent [51]. We limited bias by making reasonable attempts to contact instrument developers for further information and complete scoring independently before arriving at definitive results.

Using the criteria from Terwee et al [50] for results of measurement properties offered a rigorous, equitable and transparent critique of study results. Some instruments may have just fallen below the criteria set and therefore been rated as a negative. That is not to say the instrument cannot be used; rather, some caution should be applied when considering instrument selection. Depending on the purpose of the instrument, lower levels of reliability may have been acceptable; however, the cut-off point needed to be set somewhere.

There were also some psychometric results which did not fit the Quality Criteria for Measurement Properties' results [50], such as studies which used structural equation modelling, which were subsequently categorised as indeterminate. Applying the quality criteria was extremely time-consuming; for example, some studies took several hours. Some criteria required to be more explicit; for example, the criteria for structural validity required factors to explain more than 50 % of variance. It was unclear whether 50 % was required for each factor or total factors. We used total factors and reached decisions on anomalies through consensus discussion.

We do not suggest that the additional dimensions of utility are definitive; rather, this paper offers a starting point of a method to critique these additional, but fundamental, aspects of instrument use. Although offering

a degree of face validity, further work is required to determine application to instruments measuring other constructs. A working manual would also provide explanatory guidance for other users. As well as instrument selection, the matrix can also be used to identify research gaps for existing instruments, for example, further validity testing for the SIPES instrument or reliability studies for NHSIP. Instrument development should start with a sound theoretical development of what constitutes Quality from the Patients' Perspective. New instruments may be necessary if there are revised theoretical and conceptual developments of what constitutes quality of hospital care. Advances in how to quantify patient experience may also necessitate the development of new instruments.

## Conclusions

Patient experience data could be used to drive improvements in hospital care at national, local and healthcare team levels. To date, there are a range of instruments available to measure the patient experience of hospital quality care. Clinicians, managers, policy makers and researchers need to select patient experience instruments which are fit for purpose. This study aims to aid this choice by providing a framework to allow consideration of a wide perspective of the utility of instruments. Users can weigh the importance of each dimension, depending on the purpose of data collection, thus aiding instrument choice. Selecting the right patient experience instrument for the right purpose can aid improvements in hospital quality of care.

## Additional files

**Additional file 1: PRISMA (2009) Checklist.** This file contains a checklist of the criteria necessary for reporting a systematic review, detailing where specific criteria are covered in the paper.

**Additional file 2: Search results.** This file contains the search strategy conducted in MEDLINE and results of all database and grey literature searching.

## Abbreviations

CINAHL: Cumulative Index to Nursing and Allied Health Literature; COSMIN: Consensus-based Standards for the Selection of Health Measurement Instruments; HCAHPS: Hospital Consumer Assessment of Healthcare Providers and Systems; HKIEQ: Hong Kong Inpatient Experience Questionnaire; I-PAHC: Patient Experiences with Inpatient Care; MEDLINE: Medical Literature Analysis and Retrieval System Online; NHSIP: National Health Service Inpatient Survey; NORPEQ: Norwegian Patient Experience Questionnaire; PEQ: Patient Experience Questionnaire; PPE-15: Picker Patient Experience Questionnaire; PPQ: Patient Perceptions of Quality; PsychINFO: Psychological Information; QPP: Quality from the Patients' Perspective; QPPS: Quality from the Patients' Perspective Shortened; SIPES: Scottish Inpatient Patient Experience Survey.

## Competing interests

The authors declare that they have no competing interests.

**Authors' contributions**

MB conceived and designed the study, devised search strategies, applied inclusion criteria, applied quality scoring, developed the matrix and drafted the manuscript. DM provided direction for the study idea and design, provided statistical advice, applied quality scoring and helped devise matrix and the manuscript. IA participated in the study design, piloting of inclusion selection form and revision of the manuscript. WL participated in the study design, provided statistical advice, applied inclusion criteria, applied quality scoring and revision of the manuscript. All authors have read and approved the final manuscript.

**Acknowledgements**

Many thanks to Rob Polson, Subject Librarian, who assisted with database searching and Ref Works support. Thanks to instrument developers and study authors who responded and provided information via e-mail queries and telephone discussions.

**Funding**

No specific funding has been granted for the study; however, the University of Stirling is funding fees for a PhD by publication by the first author. It is proposed that this paper will constitute part of the PhD by publication submission.

**Author details**

<sup>1</sup>School of Health Sciences, Centre for Health Science, University of Stirling, Highland Campus, Old Perth Rd, Inverness IV2 3JH, UK. <sup>2</sup>Quality, Safety and Informatics Research Group, University of Dundee, Dundee, UK. <sup>3</sup>School of Nursing, Midwifery and Social Care, Edinburgh Napier University, Edinburgh, UK. <sup>4</sup>School of Health Sciences, University of Stirling, Stirling, UK.

Received: 13 March 2015 Accepted: 8 July 2015

Published online: 23 July 2015

**References**

- Agency for Healthcare Research and Quality. HCAHPS Fact Sheet CAHPS Hospital Survey. USA; 2012.
- Agency for Healthcare Research and Quality. HCAHPS Survey 2012. [http://www.hcahpsonline.org/Files/HCAHPS%20V7%200%20Appendix%20A1%20-%20HCAHPS%20Expanded%20Mail%20Survey%20Materials%20(English)%20July%202012.pdf].
- Bannigan K, Watson R. Reliability and validity in a nutshell. *Jn Clin Nurs*. 2009;18:3237–43.
- Beattie M, Lauder W, Atherton I, Murphy D. Instruments to measure patient experience of health care quality in hospitals: a systematic review protocol. *Syst Rev*. 2014;3:4.
- Beattie M, Shepherd A, Howieson B. Do the Institute of Medicines' (IOM) dimensions of quality capture the current meaning of quality in health care?—an integrative review. *J Res Nurs*. 2012;18:288–304.
- Boyd J. The 2006 inpatients importance study. Oxford, Picker Institute Europe: The Acute Co-ordination Centre for the NHS Acute Patient Survey Programme; 2007.
- Castle NG, Brown J, Hepner K, Hays R. Review of the literature on survey instruments used to collect data on hospital patients' perceptions of care. *Health Serv Res*. 2005;40:1996–2017.
- Centers for Medicare and Medicaid (CMS). CAHPS Hospital Survey (HCAHPS) quality assurance guidelines. Baltimore: CMS; 2013.
- Cleary P, Edgman-Levitan S, Robert M. Patient evaluate their hospital care: a national survey. *Health Aff*. 1991;10:254–67.
- Coaley K. An introduction to psychological assessment and psychometrics. London: Sage publications; 2014.
- Coulter A, Fitzpatrick R, Cornwell J. The point of care measures of patients' experience in hospital: purpose, methods and uses. *The Kings fund*. 2009. p. 1–32.
- Council of Medical Speciality Societies. The measurement of health care performance: a primer from the CMSS. USA: United Health Services Inc; 2007.
- DeCourcy A, West E, Barron D. The National Adult Inpatient Survey conducted in the English National Health Service from 2002 to 2009: how have the data been used and what do we know as a result? *BMC Health Serv Res*. 2012;12:71.
- Department of Health. Patients first and foremost: the initial government response to the report of the Mid Staffordshire NHS Foundation Trust Public Inquiry. London: Stationary Office; 2013.
- DeVet H, Terwee C, Mokkink L, Knol D. Measurement in medicine. New York: Cambridge University Press; 2011.
- Donabedian A. Explorations in quality assessment and monitoring, Vol. 1, The definition of quality and approaches to its assessment. Ann Arbor: Health Admin Press; 1980.
- Foundation H. Measuring patient experience: No. 18, evidence scan. Health Foundation: England; 2013.
- Francis R. Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry. London: Stationary Office; 2013.
- Gerteis M, Edgman-Levitan S, Delbanco T. Through the patient's eyes. Theoretical development of what defines quality from the patients' perspective. San Francisco: Jossey-Bass; 1993.
- Giordano L, Ellito M, Goldstein E, Lehrman E, Spencer P. Development, implementation and public reporting of the HCAHPS Survey. *Med Care Res Rev*. 2009;67:27–37.
- Groene O, Botje D, Sunol R, Andree-Lopez M, Wagner C. A systematic review of instruments that assess the implementation of hospital quality management systems. *Int J Qual Health Care*. 2013;25:525–41.
- Hospital Authority. Patient satisfaction survey in Hong Kong Hospital Authority; Final Report. Hong Kong; 2010
- Institute of Medicine (IOM). Crossing the quality chasm: a new health system for 21st century. Washington DC: National Academy Press; 2001.
- Jenkinson C, Coulter A, Reeves R, Bruster S, Richards N. Properties of the Picker Patient Experience questionnaire in a randomized controlled trial of long versus short form survey instruments. *J Public Health Med*. 2003;25:197–201.
- Jenkinson C, Coulter A, Bruster S. The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *Int J Qual Health C*. 2002;14:353–8.
- Keller S, O'Malley A, Hays RD, Mathew RA, Zaslavsky AM, Hepner AM, et al. Methods used to streamline the CAHPS Hospital Survey. *Health Serv Res*. 2005;40:2057–77.
- Larsson BW, Larsson G. Development of a short form of the Quality from the Patient's Perspective (QPP) questionnaire. *J Clin Nurs*. 2002;11:681–7.
- Larsson G, Larsson BW, Munck IME. Refinement of the questionnaire 'quality of care from the patient's perspective' using structural equation modelling. *Scand J Caring Sci*. 1998;12:111–8.
- Levine RE, Fowler FJ, Brown JA. Role of cognitive testing in the development of the CAHPS Hospital Survey. *Health Serv Res*. 2005;40:2037–56.
- Locock L, Robert G, Boaz A, Vougioukalou S, Shuldham C, Fielden J, et al. Using a national archive of patient experience narratives to promote local patient-based co-design. *J Health Serv Res Policy*. 2014;4:200–7.
- Luxford K. What does the patient know about quality? *Int Jn Qual Health Care*. 2012;24:439–40.
- Manary MP, Boulding W, Staelin R, Glickman SW. The patient experience and health outcomes. *N Engl J Med*. 2013;368:201–3.
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol*. 2009;62:1006–12.
- Mokkink L, Terwee C, Patrick D, Alonso J, Stratford P, Knol D, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19:539–49.
- Murphy DJ, Bruce DA, Mercer SW, Eva KW. The reliability of workplace-based assessment in postgraduate medical education and training: a national evaluation in general practice in the United Kingdom. *Adv Health Sci Edu*. 2009;14:219–32.
- O'Malley AJ, Zaslavsky AM, Hays RD, Hepner KA, Keller S, Cleary PD. Exploratory factor analyses of the CAHPS Hospital Pilot Survey responses across and within medical, surgical and obstetric services. *Health Serv Res*. 2005;40:2078–94.
- Oldedal S, Garrat A, Bjertnaes O, Bjornsdottir M, Freil M, Sachs M. The NORPEQ patient experience questionnaire: data quality, internal consistency and validity following a Norwegian inpatient survey. *Scand J Caring Sci*. 2007;35:540–7.
- Osse B, Vernooij-Dassen M, Schade E, Grol R. A practical instrument to explore patients' needs in palliative care: the problems and needs in palliative care. *Palliative Med*. 2007;21:391–9.
- Pettersen KL, Veenstra M, Guldvog B, Kolstad A. The Patient Experiences Questionnaire: development. *Int J Qual Health C*. 2004;16:453–63.

40. Picker Institute Europe. Guidance manual for the NHS Adult Inpatient Survey 2012. Oxford: Picker Institute Europe; 2012.
41. Rao KD, Peters DH, Bandeen-Roche K. Towards patient-centered health services in India—a scale to measure patient perceptions of quality. *Int J Qual Health C.* 2006;18:414–21.
42. Reeves R, Coulter A, Jenkinson C, Cartwright J, Bruster S, Richard N. Development and pilot testing of instruments for use in the acute NHS Trust Inpatient Survey Programme. Oxford: Picker Institute; 2002.
43. Scales Jr C, Schulman K. Triggering management for quality improvement. *Health Serv Res.* 2014;49(5):1401–6.
44. Schroder A, Larsson B, Ahlstrom G. Psychometric properties of the instrument quality in psychiatric care and descriptions of quality of care among in-patients. *IJHCQA.* 2010;23:554–70.
45. Scottish Government. Scottish Inpatient Patient Experience Survey 2010, Volume 2: Technical Report. (<http://www.scotland.gov.uk/Publications/2010/09/30111425/0>).
46. Scottish Government. Scottish Inpatient Patient Experience Survey 2012, Volume 2: Technical Report. (<http://www.scotland.gov.uk/Publications/2012/08/5616/0>).
47. Szymur S, Redding D. Core domains for measuring inpatients' experience of care. Oxford: Picker Institute Europe; 2012.
48. Sofaer S, Crofton C, Goldstein E, Hoy E, Crabb J. What do consumers want to know about the quality of care in hospitals? *Health Serv Res.* 2005;40:2018–36.
49. Streiner DL, Norman GR. Health measurement scales (3rd ed.). Oxford: Oxford Medical Publications; 2003.
50. Terwee CB, Bot SD, de Boer M, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status instruments. *J Clin Epidemiol.* 2007;60:34–42.
51. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res.* 2012;21:651–7.
52. Van der Vleuten C. The assessment of professional competence: developments, research and practical implications. *Adv in Health Sci Educ.* 1996;1:41–67.
53. Webster TR, Mantopoulos J, Jackson E, Cole-Lewis H, Kidane L, Kebede A, et al. A brief questionnaire for assessing patient healthcare experiences in low-income settings. *Int J Qual Health C.* 2011;23:258–68.
54. Wensing M, Elwyn G. Research in patients' views in the evaluation and improvement of quality of care. *Qual Saf Health Care.* 2002;11:153–7.
55. Wilde B, Larsson G, Larsson M, Starrin B. Quality of care: development of a patient-centred questionnaire based on a grounded theory model. *Scand J Caring Sci.* 1994;8:39–48.
56. Wilde B, Starrin B, Larsson G, Larsson M. Quality of care from a patient perspective: a grounded theory study. *Scand J Caring Sci.* 1993;7:113–20.
57. Williams B. Patient satisfaction: a valid concept? *Soc Sci Med.* 1994;38:509–16.
58. Williams B, Coyle J, Healy D. The meaning of patient satisfaction: an explanation of high reported levels. *Soc Sci Med.* 1998;47:1351–9.
59. Wong EL, Coulter A, Cheung AW, Yam CH, Yeoh EK, Griffiths SM. Validation of inpatient experience questionnaire. *Int J Qual Health C.* 2013;25:443–51.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)



## **4.10 Overview of Paper Four**

The following paragraphs provide additional details of the research process undertaken for the systematic review not covered in the associated publication, due to the limitations of journals word count. This is followed by a critical reflection of the work not included in the original article.

### **4.10.1 The complexity of Instrument Origin**

Once papers were retained through the application of inclusion criteria there was then a necessity to group papers into their respective instruments. This was a complex process, as some instruments had extensive histories and some were derivatives of other included instruments. For example, the Picker Patient Experience Questionnaire (PPE-15) had an extensive development history relating to the original Adult Picker Questionnaire from the early 1990s (Cleary et al 1991, Cleary et al 1992, Cleary et al 1993, Gerteis et al 1993). This often results in multiple versions of the same instrument. Also, the NHS Inpatient Survey (NHSIP) and the Hong Kong Inpatient Experience Questionnaire (HKIEQ) were built on some of the original theoretical work from the original Adult Picker Questionnaire (Boyd 2007, Hospital Authority 2010). To manage the complexity of the systematic review, clear decisions were made about inclusions. For example, only the most recent version of each instrument was included. The research question was: what instruments exist to measure the patient experience of hospital quality of care?, therefore it made pragmatic sense to include current, as opposed to old, versions of instruments. This included the most recent theoretical development of each instrument.

### **4.10.2 The Use of COSMIN Criteria**

The methods of critiquing the quality of the psychometric testing of each instrument were also challenging. For example, critiquing one aspect of quality, such as internal consistency, would necessitate the following process: stage one would involve application of the COSMIN checklist for internal consistency (11 questions, such as “was there a description of how missing items were handled?”) with dichotomous ‘yes/no’ answers. These then had to be translated into rating response of excellent, good, fair or poor by applying the COSMIN checklist with a 4-point scale (see the example in Figure 7).

**Figure 7: Example of Question and Response Using the COMIN 4-Point Rating Scale to Assess Methodological Quality**

Question	Excellent	Good	Fair	Poor
3. Was there a description of how missing items were handled?	Described how missing items were handled	Not described but can be deduced how missing items were handled	Not clear how missing items were handled	

Stage two would involve rating the quality of the results as positive, negative or indeterminate using Terwee et al's (2007) criteria. For example, a positive result for internal consistency is given if Cronbach's alpha is  $\geq 0.70$ , negative if less than 0.70 and indeterminate if Cronbach's alpha was not reported. These steps were replicated for every study for each instrument. For example, the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) instrument had studies reporting content validity, internal consistency, reliability, structural validity and measurement error, therefore requiring the application of five different checklists and associated scoring plus the application of the criteria for results (CMS 2013). To aid robustness of the procedure, a second researcher (Dr Douglas Murphy) then independently repeated the whole scoring process before meeting with me to compare results and reach consensus.

#### **4.10.3 The Development of Additional Aspects of Instrument Utility**

Given that there was no existing criteria to critique the additional, yet essential, aspects of utility, it was necessary to devise these. Criteria were developed to critique instrument cost efficiency, acceptability and educational impact (Van der Vleuten 1996). Some aspects of the COSMIN checklists influenced the design of the additional criteria, such as the use of the same rating scale (poor – excellent) and applying the 'lowest score counts' rule to determine an overall grade. There was also a desire to keep the criteria simple and brief to avoid time-consuming critique. The criteria were tested on two instruments independently before both reviewers met to clarify and refine the criteria before application to all instruments. Both reviewers agreed on the final ratings of all instruments for the additional aspects of utility, suggesting there was consistency in interpretation and application. The criteria and rating responses are described in Table 2, Paper Four.

#### **4.10.4 Beattie and Murphy Instrument Utility Matrix**

The primary purpose of the systematic review method was to pull together, or synthesise, the results of various studies to answer a key question. Whilst results of different types of psychometric studies cannot be synthesised for each instrument (i.e. it would be impossible to combine results from validity with those from reliability testing due to the diversity, or heterogeneity, of statistical tests and purpose), it would aid users to select the right instrument for the right purpose if results within types of validity and reliability were combined and presented in an easy-to-view format. No method to collectively synthesise the *quality* of the studies and *results* for individual instruments exists. Current recommendations are to conduct a narrative synthesis (Popay et al 2006, Terwee et al 2007), but whilst some narrative synthesis is useful, it was felt that there was strength in combining results, where possible, to present a simple overview, hence the development of the utility matrix. This involved combining ratings of study quality where possible. For example, if structural validity scored 'excellent' and cross-cultural validity scored 'fair', the overall rating for construct validity would be 'good'. If, however, structural validity scored 'excellent' and cross-cultural validity scored 'good', the overall rating would be 'good to excellent'. Where this was not possible, such as combining and thus misinterpreting reliability and internal consistency, these were kept separate. To establish whether other sources of error had been examined we added a question to the matrix to indicate whether all relevant sources of errors were investigated.

#### **4.11 Critical Reflection of Paper Four**

Conducting the systematic review and psychometric technique provided further opportunity for learning, as well as highlighting that there was no 'one-size-fits-all' approach to selecting an instrument to measure the patient experience of hospital quality of care. Submitting the systematic review and receiving feedback from the journal reviewers enabled further learning about psychometrics. The following paragraphs provide a critical reflection of the learning and methods used in the systematic review.

##### **4.11.1 Reflective Versus Formative Models**

The systematic review did not, initially, distinguish between reflective and formative models, as the implications for psychometric testing were not fully understood. The Paper was subject to open access peer review and this important point was highlighted by two international experts in the field of healthcare psychometrics. This important information was then applied to the systematic review. However, it would have been



beneficial to have known this at the outset and to have detailed this in the protocol. Instruments are either derived from a reflective or formative model. Distinguishing the difference is important as measuring internal consistency reliability using Cronbach's alpha is only relevant when an instrument has been derived from a reflective model (De Vet et al 2011). A reflective model assumes that all parts of the construct (i.e. patient experience of hospital quality of care) are composed of domains which are interrelated and reflect the construct (De Vet et al 2011). Changes to responses in one domain are likely to reflect changes in the other domains. For example, if quality of hospital care had a domain for *safety* and a domain for *effectiveness* there would be an assumption that changes in patient response to safety items would also result in changes in responses to items on effectiveness, as these aspects are interrelated. Theoretical models constructed in this way are known as reflective models.

As domains of reflective models are expected to be interrelated, the reliability of an instrument (internal consistency) could be determined using Cronbach's alpha, whereas, domains derived from a formative model are not expected to necessarily be interrelated. For example, a patient experience of hospital quality of care instrument might be designed around stages in the patient journey, such as admission, ward care, and discharge, among others. If the patient experienced poor care during hospital admission this would not necessarily mean that they would have the same experience during their ward stay. If items and domains are not expected to be interrelated there is little point in measuring the reliability of the interrelatedness of the items. It is necessary therefore to determine from which kind of model an instrument is derived to determine whether it was appropriate to test for internal consistency using Cronbach's alpha. It is important therefore to distinguish between formative and reflective models, otherwise, formative instruments which did not test for internal consistency could be perceived as having evidence of lesser reliability than those which did, as opposed to the test not being applicable. Learning the difference between these two types of models and the implications for internal consistency testing enabled a greater appreciation of psychometrics; highlighting another aspect of complexity which required more than a dichotomous 'yes or no' response.

#### **4.11.2 Application of COSMIN Checklists**

Application of the COSMIN checklists required extensive resources. For example, an application of the COSMIN checklists to one particular paper took eight hours (although there were 26 papers retained, some of these had conducted more than one psychometric test which therefore required application of more than one COSMIN

checklist). The time necessary to apply the COSMIN checklists was, in part, due to the ongoing learning required to understand and apply the criteria correctly. All applications of the COSMIN checklists were double-checked by Dr Douglas Murphy, Senior Clinical Research Fellow, who has extensive psychometric experience. He also reported that the use of COSMIN is extremely time-consuming. However, the checklists did provide a robust and consistent approach to critiquing the instruments' psychometric properties.

There were other challenges encountered when applying the COSMIN criteria. For example, initially both reviewers obtained different results from applying the COSMIN checklists to a few papers. Meeting to discuss the differences revealed that, on occasion, researchers had applied different checklists. Further investigation revealed disparity in the use of checklists for *hypothesis testing* and *criterion validity*. This was due to the fact that one reviewer knew that no gold standard instrument existed to measure patient experience of hospital of care, therefore assuming that there would be no studies testing criterion validity. Criterion validity is the method used to compare the validity of a new instrument to the gold standard (McDowell and Newell 1996). However, the COSMIN manual confirmed that criterion validity checklists should be applied to studies which involved the testing of a brief instrument from an original validated instrument (Mokkink et al 2012). Both reviewers then agreed to apply the criterion validity checklist to these studies.

Similarly, both reviewers were scoring the criteria for handling missing data differently within the COSMIN checklists. For example, one reviewer had rated two studies as 'poor' as it was not clear how missing data had been handled, whilst the other reviewer had rated the study as 'fair' as they felt that the missing data could be deduced from information within the paper. This second example highlighted the fact that having prior psychometric knowledge helped to apply the COSMIN checklists more accurately. Once agreement was sought on application of categories and handling missing data, there was complete consistency in both reviewers' ratings of the studies.

#### **4.11.3 Application of Quality Criteria for Results of Measurement Properties**

Quality criteria were devised by Terwee et al (2007) to determine whether the *results* of psychometric studies were positive, negative or indiscriminate. Applying the criteria in the systematic review provided useful 'cut off' points to aid decision-making as to the quality of the results of included studies. However, there were some limitations found when applying the criteria. For example, the results of a study testing the structural validity of the HCAHPS (Hospital Consumer Assessment of Healthcare Providers and

Systems) instrument (Keller et al 2005) could not have the Terwee et al (2007) criteria applied. The study used structural equation modelling, which is a form of confirmatory factor analyses. The result criteria set by Terwee et al (2007) are that structural validity is positive if factors explain at least 50% of the variance, indeterminate if variance is not mentioned and negative if factors explain < 50% of the variance. Variance was not reported in the structural equation modelling results, therefore using Terwee et al (2007) criteria, the study was rated as indeterminate. Yet, the results of the structural equation modelling in Keller et al (2005) reported a Comparative Fit Index (CFI) of 0.97. A CFI of 0.95 or higher is a positive result (Hu and Bentler 1999). There is a risk that the structural validity evidence of the HCAHPS instrument is not recognised, despite the fact that the findings were positive, and some would say superior, to other methods of factors analysis. Terwee et al (2007) recognise that the criteria for results are in ongoing development. There is likely a balance to be had between necessity and sufficiency to enable the criteria to remain interpretable and user-friendly.

#### **4.12 Study Contribution to the Research Question**

As well as developing some tools and techniques to navigate the complexity of selecting/devising a patient experience instrument which is fit for purpose, the review also established that there was no instrument to measure the patient experience of hospital quality of care which could be used for team feedback for local improvement. The systematic review and psychometric critique found that the quality of the instruments was variable, but mostly of a high standard. However, those which were brief (<20 questions) were unsuitable for other reasons. For example, the Picker Patient Experience Questionnaire (PPE-15) has 15 questions, but is not intended to be used as a standalone instrument (Jenkinson et al 2002b). The Norwegian Patient Experience Questionnaire (NORPEQ) was derived from inpatients' experience in Nordic countries (Oltedal et al 2007), which is reflective of a high-cost healthcare system. By contrast, the Patient Experience with Inpatient Care (I-PAHC) and Patient Perceptions of Quality (PPQ) were developed in non-Western, low-income healthcare settings (Rao et al 2006, Webster et al 2011). For NORPEQ, PPQ and I-PAHC, the context of their development limits their transferability. There was no instrument measuring patient experience of hospital quality of care as represented in Beattie's Model of Healthcare Quality, namely, care which is; safe, effective, timely, caring, enables system navigation and is person-centred.

The problem identified in Chapter 1 was the apparent disconnect between reported metrics of national and board-level hospital quality of care and the experiences of individual patients. It was thought that the different instruments measuring patient experience of hospital quality of care, at different levels of the healthcare system, could be contributing to this disparity. The systematic review clarified what instruments were available to measure the construct of interest and what their primary purpose and utility were. The systematic review findings also suggest that different instruments should be used for different purposes, including measurement at the micro and meso levels of the healthcare system and for quality improvement and research purposes.

There remains a need to devise an instrument to measure the patient experience of hospital quality of care which can be used at a ward level as an ongoing measure of quality improvement. The next Chapter describes the development and testing of an instrument to measure the patient experience of hospital quality of care for use at the clinical microsystem (i.e. the hospital ward).

# Chapter 5

## Developing and testing a measure of patient experience of hospital quality care

### 5.1 Aim and Linkage to Research Question

Results from the systematic review presented in Paper Four and Chapter 4 confirmed that no current measure of patient experience of hospital quality of care is suitable for quality improvement purposes at a ward/unit level, nor is there an instrument which measures the patient experience using the domains of healthcare quality, as identified in Beattie's Model of Healthcare Quality (outlined in Chapter 2). Hence, there is a need to devise a new instrument, despite the effort and resource required to do so (De Vet et al 2011, Streiner et al 2015). This research gap can be filled by devising a timely and relevant measure of hospital quality of care, from the patient experience of healthcare quality. Doing so may help reduce the chasm between metrics reporting national and board-level hospital quality of care and the experiences of individual patients. A valid, reliable and brief measure of patient experience is likely to more accurately reflect the quality of hospital care experienced locally by patients, as opposed to pooled data lacking discrimination of location. Data from such a brief instrument would be more amenable to providing measurement of ongoing improvement efforts in comparison to lengthy national surveys which do little to reveal local trends over time. Hence, this study's aim was the development of a new instrument, one that will be informed and critiqued by the standards devised for all aspects of instrument utility (validity, reliability, cost efficiency, acceptability and educational impact) identified in Chapter 4 (Van der Vleuten 1996).

Instrument development is an ongoing and complex process (Coaley 2014, Streiner et al 2015). The first part of this Chapter embeds Paper Five, which describes the preliminary, yet essential, stages of the development and psychometric testing of a brief measure of patient experience of hospital quality of care, namely; the Care Experience Feedback Improvement Tool (CEFIT). To increase the likelihood of creating an instrument which is fit for purpose and provide transparency of the development process, all aspects of instrument utility will be considered.

The second part of the Chapter includes an overview of the Paper, which provides further explanation and justification for decisions made in Paper Five. The section also provides

links to other Chapters of the thesis, demonstrating how the journey has informed the development and testing of CEFIT. This section is followed by a critical reflection of the Paper to examine the limitations of CEFIT and demonstrate associated learning. Finally, the contribution of the Chapter and its associated publication (Paper Five) in answering the research question is considered.

**Objective Four:**

To develop a brief measure of patient experience of hospital quality of care which is structurally valid and reliable.

**Publication Five**

**Beattie M**, Shepherd A, Lauder W, Atherton I, Cowie J, Murphy D (2016) Development and preliminary psychometric properties of the Care Experience Feedback Improvement Tool (CEFIT). *BMJ Open*;6e010101.doi:10.1136/bmjopen-2015-010101.

## 5.2 Paper Five: Development and Preliminary Psychometric Properties of the Care Experience Feedback Improvement Tool (CEFIT)

Downloaded from <http://bmjopen.bmj.com/> on June 16, 2016 - Published by group.bmj.com

Open Access

Research

# BMJ Open Development and preliminary psychometric properties of the Care Experience Feedback Improvement Tool (CEFIT)

Michelle Beattie,<sup>1</sup> Ashley Shepherd,<sup>2</sup> William Lauder,<sup>3</sup> Iain Atherton,<sup>4</sup> Julie Cowie,<sup>2</sup> Douglas J Murphy<sup>5</sup>

**To cite:** Beattie M, Shepherd A, Lauder W, *et al*. Development and preliminary psychometric properties of the Care Experience Feedback Improvement Tool (CEFIT). *BMJ Open* 2016;6:e010101. doi:10.1136/bmjopen-2015-010101

► Prepublication history for this paper is available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2015-010101>).

Received 24 September 2015  
Revised 7 April 2016  
Accepted 26 April 2016



CrossMark

For numbered affiliations see end of article.

**Correspondence to**  
Dr Michelle Beattie;  
[michelle.beattie@stir.ac.uk](mailto:michelle.beattie@stir.ac.uk)

### ABSTRACT

**Objective:** To develop a structurally valid and reliable, yet brief measure of patient experience of hospital quality of care, the Care Experience Feedback Improvement Tool (CEFIT). Also, to examine aspects of utility of CEFIT.

**Background:** Measuring quality improvement at the clinical interface has become a necessary component of healthcare measurement and improvement plans, but the effectiveness of measuring such complexity is dependent on the purpose and utility of the instrument used.

**Methods:** CEFIT was designed from a theoretical model, derived from the literature and a content validity index (CVI) procedure. A telephone population surveyed 802 eligible participants (healthcare experience within the previous 12 months) to complete CEFIT. Internal consistency reliability was tested using Cronbach's  $\alpha$ . Principal component analysis was conducted to examine the factor structure and determine structural validity. Quality criteria were applied to judge aspects of utility.

**Results:** CVI found a statistically significant proportion of agreement between patient and practitioner experts for CEFIT construction. 802 eligible participants answered the CEFIT questions. Cronbach's  $\alpha$  coefficient for internal consistency indicated high reliability (0.78). Interitem (question) total correlations (0.28–0.73) were used to establish the final instrument. Principal component analysis identified one factor accounting for 57.3% variance. Quality critique rated CEFIT as fair for content validity, excellent for structural validity, good for cost, poor for acceptability and good for educational impact.

**Conclusions:** CEFIT offers a brief yet structurally sound measure of patient experience of quality of care. The brevity of the 5-item instrument arguably offers high utility in practice. Further studies are needed to explore the utility of CEFIT to provide a robust basis for feedback to local clinical teams and drive quality improvement in the provision of care experience for patients. Further development of aspects of utility is also required.

### Strengths and limitations of this study

- The psychometric findings demonstrate the structural validity and internal consistency of Care Experience Feedback Improvement Tool (CEFIT) to quantify the patient experience of quality of healthcare.
- While the large sample (n=802) enabled exploration of the CEFIT structure, the findings are limited to an Australian community population, with a healthcare experience, as opposed to inpatients.
- Validity and reliability are not all or nothing approaches. Rather, each study with positive results adds to the psychometric strength of the instrument. Further testing of CEFIT is required to establish the utility of CEFIT to measure patient experience of hospital quality of care for quality improvement at a ward/unit level.

### INTRODUCTION

#### Background

Sustaining and improving hospital quality of care continues to be an international challenge.<sup>1–3</sup> These challenges include an ageing population, with an associated increase in comorbidities, coupled with increasingly complex care due to advances in technology, pharmacology and clinical specialisation.<sup>4–6</sup> These demographic and societal changes have resulted in an increase in healthcare resource and expenditure.<sup>5–7</sup> Hence, there are competing demands for limited healthcare resources. UK policymakers and healthcare organisations have responded to these trends by shifting the balance of care from an over-reliance on hospitals to community-led and home-led services.<sup>8–11</sup> The aspiration has been, and is, for mutual health services; co-design and co-creation of services with patients, rather than for patients.<sup>11</sup> Despite these changes, there remains significant

BMJ

Beattie M, *et al*. *BMJ Open* 2016;6:e010101. doi:10.1136/bmjopen-2015-010101

1

pressure on hospitals to deliver high quality of care with limited resources.

Measurement is necessary to determine whether or not new interventions or approaches are indeed improving the quality of hospital care.<sup>12</sup> Healthcare providers translate strategic targets into local measurement plans for hospital quality of care. How hospital quality of care is measured matters, as limited resources are often directed towards what is being measured.<sup>13</sup> Hence, though narratives of the patient experience can provide powerful insights into quality of healthcare, more tangible and measurable aspects of quality, such as waiting times, are used as ongoing indicators of quality, and so attract resources to address them.

Measurement of the patient perspective is important. Investigations into high profile failures of care highlight a disregard for patients' concerns, leading to calls for the patients' voice to be heard.<sup>14 15</sup> Views on what constitutes quality of care differ between patients, clinicians and managers; the patient can provide an additional (and essential) perspective to professional views.<sup>16 17</sup> The complexity of hospital care ensures that the only consistent person in the patient's journey is the patient; their perspective is unique. Devising a global measure of hospital quality of care, from the patient perspective, has the potential to use the patients' voice to direct quality improvement efforts within hospital care.

Many tools exist to measure patient satisfaction of hospital quality of care,<sup>18-22</sup> but theoretical and methodological challenges limit their use as quality measures.<sup>23</sup> For example, there is evidence to show that patients over-rate satisfaction due to gratitude bias and fear of reprisal; therefore, results tend to show a high ceiling effect, which could prevent the measure from differentiating between poor and good quality of care.<sup>24</sup> Satisfaction tends to be influenced by patient expectations, which fluctuate over time, again limiting their use in measurement of hospital quality.<sup>25</sup> As an alternative, there is evidence to suggest that patient reports of their experiences of healthcare reduce these limitations.<sup>26 27</sup> Measuring patient experience requires questions to be designed around what and how often care processes or behaviours occurred, as opposed to patient ratings of care. For example, a satisfaction survey may ask patients to rate the care process of medicine administration, whereas a patient experience survey may ask how often they received the right medication at the right time.

Instruments already exist to measure the patient experience, such as Patient Reported Experience Measures (PREMS). However, PREMS are designed to measure the patient experience of care of a specific condition or treatment as opposed to a global measure of hospital quality of care experience.<sup>28</sup> There are also instruments measuring aspects of the patient experience of hospital quality of care, such as the Consumer Quality Index<sup>29</sup> which measures collaboration between general practitioners and medical specialists; the Patient Evaluation of Emotional Care during hospitalisation<sup>30</sup> measuring

relational aspects of care; and the Consultation and Relational Empathy<sup>31</sup> measure which quantifies the patient perception of clinicians' empathy. There remains a need for a global measure of hospital quality of care.

There is no 'best' tool when identifying an instrument to measure the patient experience of hospital quality of care. Rather, decision-making is largely dependent on the purpose and context in which the data will be used. For example, patient experience data used for national league tables would require the use of a highly reliable and valid instrument, while accepting the associated high cost of resource necessary for administration, whereas patient experience data used for local quality improvement may tolerate lower levels of reliability in favour of cost-effectiveness and user acceptability.<sup>12</sup> Although it is important to consider whether an instrument accurately measures a concept (validity) in a consistent manner (reliability), other factors, such as the brevity of the instrument, may also be important.

We previously conducted a systematic review which found 11 instruments measuring the patient experience of hospital care.<sup>12</sup> Six instruments, with extensive psychometric testing, were designed for the primary purpose of data being used for national comparisons.<sup>32-37</sup> The other five instruments, which were primarily used and designed for quality improvement purposes, demonstrated some degree of validity and reliability.<sup>38-42</sup> However, their utility was compromised either by having too many items or by a poor fit to the UK healthcare context. Those which were considered too lengthy included the Quality from the Patients' Perspective with 68 items,<sup>38</sup> the Quality from the Patients' Perspective Shortened<sup>39</sup> instrument with 24 items and the Patient Experience Questionnaire<sup>40</sup> consisting of 20 items. Quality improvement measurement involves repeated data collection over time, often displayed on statistical process control charts.<sup>43</sup> Data collection may happen on a daily, weekly or monthly basis; therefore, it needs to be brief to be sustained within clinical practice. The Patient Experiences with Inpatient Care (I-PAHC)<sup>41</sup> and Patient Perceptions of Quality (PPQ)<sup>42</sup> were developed in non-western, low-income healthcare settings, which limits transferability. For example, the instruments included items around medicine availability, which are irrelevant in more affluent countries.

A concise instrument which enables rapid completion is required for improvement purposes within clinical wards or units. Lengthy instruments are a burden to patients, who are likely to be in a period of convalescence at the time of hospital discharge. Similarly, lengthy instruments could divert resources from clinicians providing care to those measuring care, which might have negative effects on the very concept we are trying to improve: quality of care. There are, of course, challenges of designing an instrument to measure patient experience of hospital quality of care for improvement purposes. For example, the brevity of the instrument risks that the tool does not fully capture the



concept of interest (threat to validity). Considering that all aspects of utility will enable a balanced consideration of the important but often competing interests for instrument development and use, Van der Vleuten<sup>44</sup> suggests that instrument utility comprises five aspects, namely: validity, reliability, cost-efficiency, acceptability and educational impact. Using this wide conception of utility will aid the development of a brief instrument to measure the patient experience of hospital quality of care for local quality improvement which has high practicality. The complexity inherent in achieving this requires a series of investigations, the first of which is reported in this study. We will require to conduct a future generalisability study to determine the number of patient opinions needed to obtain reliable results.

#### AIM

To develop a structurally valid and internally consistent brief measure of hospital care experience, the Care Experience Feedback Improvement Tool (CEFIT), by:

1. Developing items from the literature and patient experience expertise;
2. Examining the factor structure of CEFIT with those who have had previous care experience;
3. Determining the internal consistency of CEFIT in a population with care experience;
4. Critiquing utility aspects of CEFIT.

#### INSTRUMENT DEVELOPMENT

##### Stage 1: theoretical development

The instrument was theoretically informed by a literature review exploring current definitions and domains of healthcare quality. The review found that quality of care was composed of six domains, namely care which is: safe, effective, timely, caring, enables system navigation and is person-centred.<sup>46</sup> The review informed development of a model of healthcare quality, which illustrates that person-centred care is foundational to the other five domains of quality (figure 1). For example, to ensure quality of care, effectiveness needs to be delivered in a person-centred way, that is, adjusting treatment regimes in accordance with individual circumstances and needs. Therefore, person-centred care is not a separate domain, but inherent within the five domains of quality of healthcare.

##### Stage 2: item construction

An item was constructed for each of the five domains. Items were worded to capture the patient experience of quality of care, as opposed to ratings of satisfaction. Behavioural or observational prompts were devised for each item to aid patient interpretation. For example, 'I received procedures and treatments within acceptable waiting times' is a prompt for the item 'I received timely care'. Prompts were designed to enable easy adaptation to differing contexts. For example, a prompt for timely care might be 'staff responded to my call bell within a reasonable time' for a hospital context or 'I waited an



Figure 1 Beattie's model of healthcare quality.

acceptable amount of time to be seen' in an outpatient setting. A five-point Likert response was devised from 'never' to 'always' to determine the frequency of care behaviours. Any response which does not indicate 'always' suggests there is room for improvement. A global question was designed to rate the patient's overall healthcare quality experience, where 1=the worst quality care and 5=the best quality of care.

##### Stage 3: content validity

The draft tool was subjected to a content validity index (CVI) to test for a statistically significant proportion of agreement between experts on the instrument construction.<sup>45</sup> This process requires between 5 and 10 experts to review the instrument and complete a four point rating scale to determine item relevance, where 1=irrelevant and 4=very relevant.<sup>45</sup> A section is available for reviewers to make suggestions for improvement. Content validity is achieved when items are rated as 3 or 4 by a predetermined proportion of experts.<sup>45</sup> The research and development manager deemed the CVI procedure to be service evaluation.

CVI was completed by two expert panels simultaneously. One panel consisted of five volunteers who had had a previous hospitalisation for more than 24 hours within the past 6 months. The public volunteer group was derived from a cardiac rehabilitation class. At the end of the class, attendees were asked by their clinician if they would be interested in giving their views on the comprehension and completeness of a draft tool designed to obtain patient feedback. Those remaining at the end of the class reviewed the draft tool and provided feedback. The other panel consisted of five 'experts' in patient experience. Professional experts met the following criteria: current or previous role providing direct patient care, and had either a specific role in patient experience

policy or practice (ie, public involvement officer), or published research in relation to patient experience or quality of care. The patient experience experts completed the CVI while attending an international conference on healthcare quality. Eight out of 10 experts were required to rate the items as 3 or 4 to achieve content validity beyond the 0.05 level of significance.<sup>45</sup>

Both volunteer and professional experts rated all items (five domains and one overall rating question) as content valid (rating responses as 3='relevant but needs minor alterations' or 4='very relevant'), with the exception of one item (2='needing revision'). Nine out of 10 experts rated all item contents valid to measure the patient experience of hospital quality of care. Their

aggregated score gave an overall CVI as 0.90, endorsing validity beyond the 0.05 level of significance.<sup>45</sup> Suggestions for improvement within the open text comments section included the use of colour to visually separate sections within the questionnaire. The item rated as 'needing revision' was the overall global rating scale as there was no yardstick or parameter to help patients respond appropriately. The expert feedback and literature supported removal of the global rating scale as CEFIT was designed to highlight key areas for improvement and action as opposed to an assessment for judgement.<sup>46</sup> Minor alterations were made to wording and prompts. The five items, with example prompts and rating response options, are displayed in figure 2.

<p>The five statements below have been identified as important aspects of care by patients. We would welcome your feedback on how often you experienced these aspects of quality whilst receiving care. Examples are given to help describe each of the statements. You may have different examples or interpretations of the statements, which is OK.</p> <p>Please tick one box like this <input checked="" type="checkbox"/> with a ball point pen. If you change your mind just cross out your old response and make a new choice. Please answer every statement.</p>					
	Never <sup>1</sup>	Occasionally <sup>2</sup>	Sometimes <sup>3</sup>	Often <sup>4</sup>	Always <sup>5</sup>
1. I received safe care...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p>Examples might include:                  Staff washed their hands before and after any direct contact with me                  I received the right medication dose</p>					
	Never <sup>1</sup>	Occasionally <sup>2</sup>	Sometimes <sup>3</sup>	Often <sup>4</sup>	Always <sup>5</sup>
2. I received timely care...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p>Examples might include:                  I waited an acceptable amount of time to be seen                  I received procedures and treatments within a reasonable time</p>					
	Never <sup>1</sup>	Occasionally <sup>2</sup>	Sometimes <sup>3</sup>	Often <sup>4</sup>	Always <sup>5</sup>
3. My care met my personal needs...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p>Examples might include:                  My care was tailored to my individual needs i.e. adjustments made for my other conditions and personal lifestyle                  I was involved in all decisions about my care</p>					
	Never <sup>1</sup>	Occasionally <sup>2</sup>	Sometimes <sup>3</sup>	Often <sup>4</sup>	Always <sup>5</sup>
4. Staff were caring to me...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p>Examples might include:                  I was treated pleasantly and with courtesy                  Staff really listened to me and took my concerns seriously</p>					
	Never <sup>1</sup>	Occasionally <sup>2</sup>	Sometimes <sup>3</sup>	Often <sup>4</sup>	Always <sup>5</sup>
5. I was able to get the care I needed...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p>Examples might include:                  I knew how to access the care I needed                  There were no barriers to me getting the treatment/services that I needed</p>					

Figure 2 Care Experience Feedback Improvement Tool (CEFIT).

## INSTRUMENT TESTING

### Design

CEFIT was administered via a telephone population survey of people dwelling in Queensland from July to August 2014. Embedding CEFIT within a large-scale survey provided an opportunity to test the internal consistency and structural validity of CEFIT with a random sample, to determine how well items were related and to check for any possible duplication or redundancy of final content. The Queensland Social Survey is an annual statewide survey administered by the Population Research Laboratory at Central Queensland University (CQU) Australia to explore a wide range of research questions relevant to the general public. The CQU survey consists of demographic, introductory and specific research questions. We embedded the five-item CEFIT instrument within the survey. Pilot testing of CEFIT in randomly selected households (n=68) suggested that no question changes were necessary.

### Sampling

A two-stage sampling procedure was used. First, two geographically proportionate samples were drawn from (1) South-East Queensland and (2) the rest of Queensland. A telephone database was used to draw a random sample of telephone numbers within postcode areas. Second, participants were selected per household on the basis of gender (to ensure an equal male and female sample) and age ( $\geq 18$  years). Where more than one male/female met the criteria in one household, the adult with the most recent birthday was selected. The questions were preceded with a screening question to identify participants with a recent healthcare experience (<12 months). Samples of patient experience surveys are usually estimated on rates of patient discharge and include patients within 5 months of hospital discharge.<sup>12</sup> A wider time frame and context was required for CEFIT due to the nature of a population survey.

### Data collection

The sample was loaded into a Computer-Assisted Telephone Interviewing (CATI) System held within the Population Research Laboratory. The system allocates telephone numbers and provides standardised text instructions to trained interviewers.

### Analysis

Data were entered into SPSS V.19. Descriptive statistics were calculated for questionnaire characteristics to establish the range of responses. Interitem (question) total correlations were calculated in order to identify the possibility of unnecessary or redundant questions. Internal consistency of the CEFIT was calculated using Cronbach's  $\alpha$  to determine the consistency in responses to the questions and examine the error generated by the questions asked.<sup>47</sup> Exploratory factor analysis was used to examine the factor structure of CEFIT and determine structural validity. The Kaiser-Meyer-Olkin test was conducted to measure the sampling adequacy and compare magnitudes of correlation. Bartlett's Test of Sphericity was used to ensure the study assumptions were met for factor analysis. Factor analysis using the principal component method (principal component analysis) and Eigenvalue >1 rule was performed. Results for structural validity are determined as positive if factor analysis explains at least 50% of the variance.<sup>48</sup>

## RESULTS

The overall CQU survey response rate was 35.9%. From the 1223 survey participants, 802 (66%) were eligible to complete CEFIT (healthcare experienced within the preceding 12 months). All eligible participants responded to the CEFIT questions (100% response). Of the 802, 50.5% (n=405) were male and 49.5% (n=397) were female. The mean age was 58 years, range 18–95 years. The range of question responses was towards the high end of the responses, indicating that the majority of quality care processes were occurring 'often' or 'always'. However, all rating responses were used highlighting necessary range (table 1).

Questions demonstrating interitem total correlations of 0.2–0.8 are considered as offering unique and useful content which is related to the overall inventory.<sup>47</sup> Correlations (0.28–0.74) between questions were all significant ( $p \leq 0.05$ ), and this provided reassurance that all questions were inter-related, yet unique enough to necessitate their inclusion within the inventory (table 2). Cronbach's  $\alpha$  for the final questionnaire was good (0.78) and confirmed that the scale as a whole was sufficient. This provided reassurance that CEFIT's overall inventory of questions did not include unnecessary questions

**Table 1** Descriptives of Care Experience Feedback Improvement Tool (CEFIT) questions

Question	Never n (%)	Occasionally n (%)	Sometimes n (%)	Often n (%)	Always n (%)	Missing n (%)
I received safe care	10 (1.3)	5 (0.6)	9 (1.1)	42 (5.3)	724 (91.6)	12 (1.49)
I received timely care	25 (3.1)	19 (2.4)	49 (6.1)	104 (13.0)	602 (75.3)	3 (0.37)
My care met my personal needs	18 (2.3)	8 (1.0)	27 (3.4)	58 (7.3)	689 (86.1)	2 (0.24)
Staff were caring towards me	7 (0.9)	4 (0.5)	17 (2.1)	49 (6.1)	724 (90.4)	1 (0.12)
I was able to get the care I needed	13 (1.6)	9 (1.1)	20 (2.5)	36 (4.5)	723 (90.3)	1 (0.12)

**Table 2** Correlation of Care Experience Feedback Improvement Tool (CEFIT) domains

Domain	System navigation			
	Safe	Timely	Effective	Caring
Safe	1.00			
Timely	0.300	1.00		
Effective	0.411	0.383	1.00	
Caring	0.328	0.399	0.679	1.00
System navigation	0.318	0.284	0.736	0.664

**Table 4** Component matrix for Care Experience Feedback Improvement Tool (CEFIT) single-factor solution

	Component 1
Safe	0.579
Timely	0.582
Effective	0.884
Caring	0.845
System navigation	0.836

Extraction method: principal component analysis.

\*1 means that components were only extracted for factors with an Eigenvalue of &gt;1.

( $\alpha > 0.9$ ), and neither was the overall inventory too narrow in its scope ( $\alpha > 0.7$ ).<sup>47</sup>

The Kaiser-Meyer-Olkin measure of sampling adequacy was good (0.792) and Bartlett's Test of Sphericity (1546.08, Df=10,  $p=0.000$ ) indicated that the study met the assumptions for factor analysis. Eigenvalues identified a one-factor solution which accounted for 57.33% of total variance (table 3). The 57.33% variance of the one-factor solution was shared by five components, namely: safety 0.579, timely 0.582, effective 0.884, caring 0.845 and system navigation 0.836 (table 4).

#### CEFIT QUALITY CRITIQUE

Five aspects of utility were critiqued for CEFIT. First, we applied the appropriate Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) checklist for content validity, structural validity and internal consistency to assess the quality of the study methods (criteria and results are available in tables 5–7, respectively).<sup>49</sup> Responses within individual checklists were given a methodological score by applying the COSMIN four-point checklist scoring system, namely: excellent, good, fair or poor. Where individual answers to checklist questions were of variable ratings (ie, some excellent, some poor), the overall score was determined by taking the lowest rating of any item. In other words, the worst score counted.<sup>49</sup>

Second, we applied the criteria developed by Terwee<sup>48</sup> to determine rating for the quality of the results of each psychometric test performed on CEFIT (see figure 3). This enabled study results to be categorised as positive (+), indeterminate (?) or negative

(–) according to the quality criteria for each measurement property. For example, positive ratings for internal consistency are given, using Terwee *et al* criteria, if the Cronbach's  $\alpha$  is  $>0.70$ . Studies with Cronbach's  $\alpha$  results of  $<0.70$  would be categorised as negative, or where Cronbach's  $\alpha$  was not determined the result would be categorised as indeterminate. A full explanation, with justification for all COSMIN criteria results, is available from Terwee.<sup>48</sup>

Third, we applied criteria developed and tested in our previous systematic review for additional aspects of instrument utility: cost-efficiency, acceptability and educational impact (detailed in figure 4). Further explanation of the criteria and scoring is available at Beattie *et al*.<sup>12</sup> Results from all three steps are presented in an adaptation of the Beattie and Murphy Instrument Utility Matrix for CEFIT (table 8).

The study quality for the content validity of CEFIT was rated as fair as there was no assessment of whether all items were relevant for the study population (eg, gender, disease characteristics, country and setting). The overall rating of the content validation results was positive as the target population considered all items in the questionnaire to be relevant and complete. The quality of the structural validity was rated as excellent as there was an adequate sample size and no major flaws in the study design. Results for structural validity were categorised as positive as the one-factor solution explained more than 50% of the variance (57.3%).<sup>48</sup>

The study quality for internal consistency was rated as excellent using the COSMIN checklist as all questions were answered positively and there were no major flaws identified in the study. The quality of the results for

**Table 3** Total variance from factor analysis

Component	Initial Eigenvalues			Extraction sums of squared loadings		
	Total	Per cent of variance	Cumulative per cent	Total	Per cent of variance	Cumulative per cent
1	2.867	57.336	57.336	2.867	57.336	57.336
2	0.834	16.670	74.006			
3	0.710	14.194	88.200			
4	0.339	6.786	94.986			
5	0.251	5.014	100.000			

**Table 5** Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) criteria and Care Experience Feedback Improvement Tool (CEFIT) results for content validity

COSMIN questions for content validity	Response	Rating
1. Was there an assessment of whether all items refer to relevant aspects of the construct to be measured?	Yes	Excellent
2. Was there an assessment of whether all items are relevant for the study population? (eg, age, gender, disease characteristics, country, setting)	No	Fair
3. Was there an assessment of whether all items are relevant for the purpose of the measurement instrument? (discriminative, evaluative and/or predictive)	Partial	Fair
4. Was there an assessment of whether all items together comprehensively reflect the construct to be measured?	Yes	Excellent
5. Were there any important flaws in the design or methods of the study?	No	Fair
6. Was there evidence that items were theoretically informed?	Yes	Excellent
Total rating	Lowest score counts	Fair

**Table 6** COSMIN criteria and CEFIT results for structural validity

COSMIN questions for structural validity	Response	Rating
1. Does the scale consist of effect indicators, that is, is it based on a reflective model?	Yes	Excellent
2. Was the percentage of missing items given?	Yes	Excellent
3. Was there a description of how missing items were handled?	Not necessary as none	Excellent
4. Was the sample size included in the analysis adequate?	Yes	Excellent
5. Were there any important flaws in the design or methods of the study?	No	Excellent
6. For CTT: Was exploratory or confirmatory factor analysis performed?	Yes	Excellent
7. For IRT: Were IRT tests for determining the (uni) dimensionality of the items performed?	NA	NA
Total rating	Lowest score counts	Excellent

CEFIT, Care Experience Feedback Improvement Tool; COSMIN, Consensus-based Standards for the Selection of Health Measurement Instruments; CTT, Classical Test Theory; NA, not available.

**Table 7** COSMIN criteria and CEFIT results for internal consistency

COSMIN questions for internal consistency	Response	Rating
1. Does the scale consist of effect indicators, that is, is it based on a reflective model?	Yes	Excellent
2. Was the percentage of missing items given?	Yes	Excellent
3. Was there a description of how missing items were handled?	Not necessary as none	Excellent
4. Was the sample size included in the internal consistency analysis adequate?	Yes	Excellent
5. Was the unidimensionality of the scale checked? That is, was factor analysis or IRT model applied?	Yes	Excellent
6. Was the sample size included in the unidimensionality analysis adequate?	Yes	Excellent
7. Was an internal consistency statistic calculated for each (unidimensional) (sub)scale separately?	NA	NA
8. Were there any important flaws in the design or methods of the study?	No	Excellent
9. For CTT: Was Cronbach's $\alpha$ calculated?	Yes	Excellent
10. For dichotomous scores: Was Cronbach's $\alpha$ or KR-20 calculated?	NA	NA
11. For IRT: Was a goodness of fit statistic at a global level calculated? For example, $\chi^2$ , reliability coefficient of estimated latent trait value (index of subject or item)	NA	NA
Total rating	Lowest score counts	Good

CEFIT, Care Experience Feedback Improvement Tool; COSMIN, Consensus-based Standards for the Selection of Health Measurement Instruments; CTT, Classical Test Theory; IRT, Item Response Theory; KR-20, Kuder-Richardson Formula 20; NA, not available.

internal consistency was rated as positive as the quality criteria for measurement properties suggest that positive ratings are applied when Cronbach's  $\alpha$  is  $\geq 0.70$  (Cronbach's  $\alpha$  for CEFIT was 0.78).<sup>48</sup>

Applying the additional aspects of utility scoring criteria found that CEFIT was rated as 'good' for cost-efficiency. While the majority of responses were good to excellent, it is not yet known how many CEFIT

Figure 3 Quality criteria for measurement properties (Terwee).<sup>48</sup>

Property	Rating <sup>†</sup>	Quality Criteria
<b>Reliability</b>		
Internal consistency	+	Cronbach's alpha(s) $\geq 0.70$
	?	Cronbach's alpha not determined or dimensionality unknown
	-	Cronbach's alpha(s) $< 0.70$
Reliability	+	ICC / weighted Kappa $\geq 0.70$ OR Pearson's $r \geq 0.80$
	?	Neither ICC / weighted Kappa, nor Pearson's $r$ determined
	-	ICC / weighted Kappa $< 0.70$ OR Pearson's $r < 0.80$
Measurement error	+	MIC $>$ SDC OR MIC outside the LOA
	?	MIC not defined
	-	MIC $\leq$ SDC OR MIC equals or inside LOA
<b>Validity</b>		
Content validity	+	All items are considered to be relevant for the construct to be measured, for the target population, and for the purpose of the measurement AND the questionnaire is considered to be comprehensive
	?	Not enough information available
	-	Not all items are considered to be relevant for the construct to be measured, for the target population, and for the purpose of the measurement OR the questionnaire is considered not to be comprehensive
Construct validity - Structural validity	+	Factors should explain at least 50% of the variance
	?	Explained variance not mentioned
	-	Factors explain $<$ 50% of the variance
- Hypothesis testing	+	Correlations with instruments measuring the same construct $\geq 0.50$ OR at least 75% of the results are in accordance with the hypotheses AND correlations with related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlations with instruments measuring the same construct $< 0.50$ OR $<$ 75% of the results are in accordance with the hypotheses OR correlations with related constructs are lower than with unrelated constructs
- Cross-cultural validity	+	No differences in factor structure OR no important DIF between language versions
	?	Multiple group factor analysis not applied AND DIF not assessed
	-	Differences in factor structure OR important DIF between language versions
Criterion validity	+	Convincing arguments that gold standard is "gold" AND correlation with gold standard $\geq 0.70$
	?	No convincing arguments that gold standard is "gold"
	-	Correlation with gold standard $< 0.70$
<b>Responsiveness</b>		
Responsiveness	+	Correlation with changes on instruments measuring the same construct $\geq 0.50$ OR at least 75% of the results are in accordance with the hypotheses OR AUC $\geq 0.70$ AND correlations with changes in related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlations with changes on instruments measuring the same construct $< 0.50$ OR $<$ 75% of the results are in accordance with the hypotheses OR AUC $< 0.70$ OR correlations with changes in related constructs are lower than with unrelated constructs

MIC = minimal important change, SDC = smallest detectable change, LoA = limits of agreement, ICC = intraclass correlation coefficient, DIF = differential item functioning, AUC = area under the curve

<sup>†</sup> + = positive rating, ? = indeterminate rating, - = negative rating

questionnaires will be necessary to ensure reliable data to differentiate between those experiencing different care experiences, which resulted in an overall score of good as opposed to excellent for cost-efficiency. Acceptability of CEFIT was scored as 'poor' as it has not yet been tested in the context in which it is intended (hospital). Educational impact was scored as 'good' as CEFIT was determined as easy to calculate the score and use results, but there is not yet evidence of it being used for quality improvement purposes.

## DISCUSSION

Our psychometric findings support CEFIT as a structurally valid instrument with positive internal consistency to measure patient care experience within an Australian

community population. The uniqueness of CEFIT is that it provides a brief yet structurally valid and reliable tool. The brevity of the instrument indicates its potential use as a quality improvement measure of patient experience. Quality improvement advocates ongoing measurement over time, as opposed to snapshot audits or before and after measures.<sup>43</sup> The sustainable use of CEFIT within busy hospital wards depends on its simplicity and brevity. However, there are ongoing challenges to ensure that such a brief measure captures the complexity of the patient experience of hospital quality of care (validity), that a small number of items within an instrument sufficiently captures the concept of interest. This study found that the CEFIT structure is measuring the patient experience of quality of care sufficiently (validity) and is doing so in a consistent manner (internal consistency

**A**

Questions for Cost Efficiency	Excellent (****)	Good (***)	Fair (**)	Poor (*)
1. What are the number of assessments (observations, questionnaires) needed to ensure reliable data?	Only a small sample needed (<30)	A moderate sample size (30-49)	Not explicit but can be assumed or (50-99 assessments needed)	No details given or (>100 assessments needed)
2. How long does an assessment take to complete?	≤15 mins	≤30mins	30-60min	>60mins
3. What are the administrative costs of completing the assessment?	Easily embedded within existing resource. Little additional support required.	Some administrative resource but no specialist resource required.	Large amount of resource required to assess and administer	Significant specialist expertise and administrative time required to assess and administer
4. What is the cost to complete a reliable sample?	Minimal	Moderate	Considerable	Extensive
Questions for Acceptability	Excellent (****)	Good (***)	Fair (**)	Poor (*)
1. Is there evidence of subjects understanding of the instrument/assessment?	Investigations of subjects understanding (i.e. cognitive testing of instruments)	Estimated evidence of subjects understanding (i.e. high number of questions missed)	Subject understanding not explicitly stated but some can be assumed (i.e. student guide to OSCE)	No evidence of subject understanding

**B**

2. How many assessments are not completed?	There are low numbers of missing items (<10%) and adequate response rates (≥40%)	There are a high number of missing items (≥10%) and an adequate response rates (≥40%)	There are low numbers of missing items or poor (<10%) and an inadequate response rate (<40%)	There are high numbers of missing items (≥10%) and poor response rates (<40%)
3. Has the instrument/assessment been tested in an appropriate context?	Evidence of successful administration/use within an appropriate setting	Tested in vivo and changes recommended would be achievable	Testing in vivo and changes recommended would be difficult or only partially tested in vivo	Testing has only been conducted in vitro/simulation
Questions for Educational Impact	Excellent (****)	Good (***)	Fair (**)	Poor (*)
1. There is evidence of the instruments intended purpose being achieved (i.e. if aim is to enable hospital ranking for patient selection, is there evidence that the results are actually influencing patient choice?)	Clear evidence of intended purpose being fulfilled.	Explanatory or theoretical link between intended and actual use, but no clear evidence.	Evidence of theoretical work but relationship between intended and actual purpose poorly or not described.	No evidence of intended purpose becoming actual.
2. The scoring system is easily translated or available in an easy to use format?	Explicitly stated and easy to calculate	Explicitly stated but not easy to calculate	Scoring only calculated by resource with statistical knowledge	Scoring not explained well enough to calculate.
3. The feedback from the results can be readily used for action where necessary?	Feedback is readily available in a format that enables necessary action.	Feedback is readily available but not drilled down enough to enable targeted action.	Minimal feedback available or delay results in limited impact.	No explanation to determine adequacy of feedback. No direct feedback could be readily used without additional expertise.

**Figure 4** Additional aspects of utility scoring criteria. OSCE, Objective Structured Clinical Examination.

reliability). Testing of internal consistency reliability and interitem correlations confirmed the need for five components, which were inter-related yet unique enough to require their own component. A Cronbach's  $\alpha$  of 0.78 provides reassurance of the reliability of the instrument and confirms that each of the five items is valuable. In addition, factor analysis revealed CEFIT as measuring a single factor, which we have called healthcare quality. We have therefore taken the first steps to ensure that CEFIT is structurally sound. This is an initial but essential step in instrument development.<sup>50</sup> Of course, validity and reliability are not all or nothing approaches. Rather, each study, with positive results, adds to the psychometric strength of the instrument.

Many instruments are criticised for not being derived from a theoretical model, which is an essential step in instrument development and subsequent validation.<sup>27</sup> CEFIT was derived from our theoretically informed model with factor analysis identifying one factor (quality of care), composed of five domains: safety, effectiveness, caring, system navigation and timeliness. To ensure that

patients experience high quality of care, these domains need to be expressed in a person-centred way. Hence, person-centred care is foundational and necessary for all other domains of healthcare quality. Our one-factor solution enables a brief yet structurally valid measure.

While there are patient experience instruments to measure national-level performance, the data are not timely, nor specific enough, to direct or measure local improvement efforts at the clinical ward level.<sup>35</sup> For example, when data are used for national comparisons, there are robust and lengthy procedures to ensure a reliable sample. While such criteria are necessary, the delay between sampling and analysis often creates a significant delay (up to 1 year) between data collection and results.<sup>12</sup> Given that much change can occur over that time within a hospital, such data would be of limited use for improvement purposes at a ward level. This is not to suggest that brief instruments are superior to lengthy evaluations, but rather that they serve different purposes.

Further, if hospital-level or national-level data identified a deterioration in patient experience around



Table 8 CEFIT results of Beattie and Murphy Instrument Utility Matrix

Aspect of utility	Measurement property/criteria checklist	Result	Rating of study quality	Rating of quality of results	Combined score of methods and results
Validity	Content validity (COSMIN checklist <sup>49</sup> and Terwee) <sup>48</sup>	Items derived from integrative review of the literature and associated theoretical model. Content validity index with a patient and expert group found items were relevant and comprehensive.	Fair	Positive	***
	Structural validity (COSMIN checklist <sup>49</sup> and Terwee) <sup>48</sup>	Principal component analysis confirmed a unidimensional scale (one-factor solution) accounting for 57.3% variance.	Excellent	Positive	****+
Reliability	Internal consistency (COSMIN checklist <sup>49</sup> and Terwee) <sup>48</sup>	Cronbach's $\alpha$ 0.78	Excellent	Positive	****+
Cost	Additional aspects of utility scoring criteria in figure 4	Number of CEFIT questionnaires needed to ensure reliable data is not yet known. Completion of CEFIT <15 min. Some administrative resource but no specialist resource required. Overall cost-efficiency calculated as moderate	NA	Good	***
Acceptability	Additional aspects of utility scoring criteria in figure 4	Investigations of participants' understanding. There are low numbers of missing items (<10%) and adequate response rates (>40%). Testing has not yet been conducted in context (ie, hospital setting).	NA	Poor	*
Educational impact	Additional aspects of utility scoring criteria in figure 4	Explanatory or theoretical link between intended and actual use, but no clear evidence yet. Scoring stated and easy to calculate. Feedback is readily available in a format that enables necessary action.	NA	Good	***

Ratings of study quality \*poor, \*\*fair, \*\*\*good, \*\*\*\*excellent ratings of quality of results: (+) positive rating, (-) negative rating, (?) indeterminate rating, mixed (+?).  
 CEFIT, Care Experience Feedback Improvement Tool; COSMIN, Consensus-based Standards for the Selection of Health Measurement Instruments.

privacy and dignity, for example, there is no way of knowing from which ward or unit within the hospital this problem originates. Similarly, episodes of positive patient experiences cannot be linked to specific wards or teams, thus limiting the ability to spread good practice. Since the intention of CEFIT is to gather data per ward, results will be directly applicable to that area. CEFIT will most likely be a useful indicator of areas of problem identification or to demonstrate improvement in key aspects of quality of care from the patient experience. It is also likely that the addition of an open question to CEFIT will help direct improvement activity given the known strengths of patient narratives to motivate clinical teams to improve.<sup>51</sup> The ability of CEFIT to drive improvement remains untested and is a matter for future work.

As well as the ongoing threat to validity, brief measures also present psychometric challenges. For example, the five-point rating response options assume that there are equidistant differences between each rating and that each of the five questions has equal importance in the patient experience of hospital quality of care.<sup>52</sup> These issues will be investigated further using statistical modelling in future studies. However, this needs to be balanced with brevity and simplicity. Use of the instrument by frontline staff should not require sophisticated statistics.

Application of the utility matrix and clarity of the instrument's primary purpose will continue to help inform the development of CEFIT. As evidence of instrument utility develops over time, it is important not to dismiss newer instruments with poorer scores.



For example, the content validity of CEFIT was rated as fair, as there is not yet evidence of the instrument being tested with hospitalised patients. We are also aware that the criteria for additional aspects of utility will need further development, but they offer a useful starting point to consider these other important aspects of instrument utility. Conducting a utility critique will help ensure the usefulness of CEFIT for frontline care.

A limitation of the study is that CEFIT was embedded within a telephone survey and tested with people who had had a healthcare experience within the preceding 12 months. While this enabled exploration of the CEFIT structure with a large sample, the findings are limited to an Australian community population, with a healthcare experience, as opposed to inpatients. However, the promising psychometric results suggest that it would be worthwhile to test CEFIT with recently hospitalised patients to determine validity in this population, as well as identifying the numbers needed to differentiate between good and not so good patient experience of hospital quality of care.

We recognise the potential limitation of range due to CEFIT responses being mostly towards the high end of response options. There is a risk that data grouped towards the high end of option responses limit the ability to distinguish between 'good' and 'poor' care. However, although CEFIT responses were mostly towards the high end of the scale, all responses were used, indicating the potential range necessary to differentiate between varying care experiences. Other instruments with high ceiling effects have been able to differentiate between aspects of good and not so good care.<sup>31</sup> There also remains debate as to the accuracy of using Cronbach's  $\alpha$  and factor analysis with non-normally distributed data, although large samples can reduce the effect.<sup>53 54</sup> We will remain vigilant of the potential threat, but will not know whether CEFIT can differentiate between different experiences of quality of care until we conduct our future generalisability study.

## CONCLUSIONS

Measuring the quality of hospital care from the patient perspective is a vital component of healthcare measurement and improvement plans, but the effectiveness of the data collected is dependent on a balanced consideration of all aspects of instrument utility. This study has established a structurally valid and internally consistent measure of the patient experience of hospital quality of care, namely the CEFIT. The next steps are to validate within an inpatient context and establish its reliability to discriminate between different patient experiences at ward/unit level to direct improvement efforts in clinical practice.

### Author affiliations

<sup>1</sup>School of Health Sciences, University of Stirling, Centre for Health Science, Inverness, UK

<sup>2</sup>School of Health Sciences, University of Stirling, Stirling, UK

<sup>3</sup>College of Nursing, University of South Florida, South Florida, Florida, USA

<sup>4</sup>School of Nursing, Midwifery and Social Care, Edinburgh Napier University, Edinburgh, UK

<sup>5</sup>Quality, Safety and Informatics Research Group, University of Dundee, Dundee, UK

**Twitter** Follow Michelle Beattie at @BeattieQi

**Acknowledgements** The authors thank their colleagues at Population Research Laboratory, Central Queensland University, Australia for embedding the CEFIT questions within their Annual Social Survey.

**Contributors** MB and DJM conceived and designed the CEFIT instrument. MB designed the theoretical model of healthcare quality. WL and IA contributed to the thinking and development of the work in their role as MB supervisors. MB and DJM designed the study. AS facilitated acquisition of data via the Queensland survey. MB and DJM designed and collected data for the CVI. WL and IA conducted statistical analysis and interpretation. JC helped in result interpretation and statistical revision. MB drafted the manuscript which was critically revised by all authors before agreeing on the final version of the manuscript.

**Funding** The University of Stirling provided funding to embed the CEFIT questions within the Queensland Annual Social Survey.

**Competing interests** None declared.

**Ethics approval** Human Ethics Review Panel Central Queensland University (Reference H13/06-120).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** The CEFIT measure is available free of charge for NHS and University staff for the purpose of research or improvement. Anyone wishing to use the measure should contact and register their request with MB ([michelle.beattie28@gmail.com](mailto:michelle.beattie28@gmail.com)) and/or DJM ([douglasmurphy27@gmail.com](mailto:douglasmurphy27@gmail.com)).

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

## REFERENCES

1. Institute of Medicine. *Crossing the quality chasm: a new health system for 21st century*. Washington DC: National Academy Press, 2001.
2. Department of Health. High Quality Care for All. Gateway Ref, 2008:10106.
3. Scottish Government. *Achieving sustainable quality in Scotland's healthcare: a '20:20' Vision; 2012*. <http://www.gov.scot/Topics/Health/Policy/2020-Vision/Strategic-Narrative> (accessed 19 Jun 2015).
4. World Health Organisation. *Innovative care for chronic conditions. Building blocks for action*. Geneva: WHO, 2002.
5. Department of Health. *Equity and excellence: liberating the NHS*. London: Stationary Office, 2010.
6. Ham C, Dixon A, Brooke B. *Transforming the delivery of health and social care: the case for fundamental change*. London: The Kings Fund, 2012.
7. World Health Organisation. *World Health Day 2012—ageing and health-toolkit for event organizers*. WHO, 2012. [http://apps.who.int/iris/bitstream/10665/70840/1/WHO\\_DCO\\_WHD\\_2012.1\\_eng.pdf](http://apps.who.int/iris/bitstream/10665/70840/1/WHO_DCO_WHD_2012.1_eng.pdf) (accessed 26 Aug 2015).
8. Department of Health. *Our health, our care, our say: a new direction for community services*. Her Majesty Government, 2006.
9. Scottish Executive. *Better health, better care: discussion document*. Edinburgh: Scottish Executive, 2007.
10. Scottish Executive. *Building a health service fit for the future*. Edinburgh: Scottish Executive, 2006.
11. Scottish Government. *The healthcare quality strategy for NHS Scotland*. Scottish Government, 2010.
12. Beattie M, Murphy D, Atherton I, et al. Instruments to measure patient experience of health care quality in hospitals: a systematic review. *Syst Rev* 2015;4:97.



13. Berry N, Gardner T, Anderson I. *On targets: how targets can be most effective in the English NHS*. London: Health Foundation, 2015.
14. Francis R. *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. London: Stationary Office, 2013.
15. MacLean L. *The Vale of Leven Hospital Inquiry Report*. Scotland: The Vale of Leven Hospital Inquiry Team APS Group, 2014.
16. Beattie M, Shepherd A, Howieson B, et al. Do the Institute of Medicines' (IOM) dimensions of quality capture the current meaning of quality in health care?—An integrative review. *J Res Nurs* 2012;18:288–304.
17. Donabedian A. *Explorations in quality assessment and monitoring, vol. 1, the definition of quality and approaches to its assessment 1980*. Ann Arbor, MI: Health Admin Press.
18. Rubin H, Ware J, Nelson E, et al. The Patient Judgements of Hospital Quality (PJHQ) questionnaire. *Med Care* 1990;28:S17–18.
19. Parasuraman A, Berry L, Zeithaml V. Refinement and reassessment of the SERVQUAL Scale. *J Retail* 1991;67:420–50.
20. Mountzoglou A, Dafogianni C, Karra V, et al. Development and application of a questionnaire for assessing parent satisfaction with care. *Int J Qual Health Care* 2000;12:331–7.
21. Henriks A, Oort F J, Vriellink M R, et al. Reliability and validity of the Satisfaction with Hospital Care Questionnaire. *Int J Qual Health Care* 2002;14:471–82.
22. Smith LF. Postnatal care: development of a psychometric multidimensional satisfaction questionnaire (the WOMBPNSQ) to assess women's views. *Br J Gen Pract* 2011;61:628–37.
23. Williams B. Patient satisfaction: a valid concept? *Soc Sci Med* 1994;38:509–16.
24. Moret L, Nguyen J, Pillet N, et al. Improvement of psychometric properties of a scale measuring inpatient satisfaction with care: a better response rate and a reduction of the ceiling effect. *BMC Health Serv Res* 2007;7:197.
25. Williams B, Coyle J, Healy D. The meaning of patient satisfaction: an explanation of high reported levels. *Soc Sci Med* 1998;47:1351–9.
26. Luxford K. What does the patient know about quality? *Int J Qual Health Care* 2012; 24:439–40.
27. Foundation H. *Measuring patient experience: no. 18, evidence scan*. England: Health Foundation, 2013.
28. Black N, Varaganum M, Hutchings A. Relationship between patient reported experience (PREMs) and patient reported outcomes (PROMs) in elective surgery. *BMJ Qual Saf* 2014;23:534–42.
29. Ikkersheim DE, Koolman X. Dutch healthcare reform: did it result in better patient experiences in hospitals? A comparison of the consumer quality index over time. *BMC Health Serv Res* 2012;12:76.
30. Murrells T, Robert G, Adams M, et al. Measuring relational aspects of hospital care in England with the Patient Evaluation of Emotional Care during Hospitalisation (PEECH) survey questionnaire. *BMJ Open* 2013;3:e002211.
31. Mercer S, Murphy D. Validity and reliability of the CARE measure in secondary care. *Clin Govern Int J* 2008;13:261–83.
32. O'Malley AJ, Zaslavsky AM, Hays RD, et al. Exploratory factor analyses of the CAHPS hospital pilot survey responses across and within medical, surgical and obstetric services. *Health Serv Res* 2005;40(6 Pt 2):2078–94.
33. Jenkinson C, Coulter A, Reeves R, et al. Properties of the Picker Patient Experience questionnaire in a randomized controlled trial of long versus short form survey instruments. *J Public Health Med* 2003;26:197–201.
34. Picker Institute Europe. *Guidance manual for the NHS Adult Inpatient Survey 2012*. Oxford: Picker Institute Europe, 2012.
35. Scottish Government. *Scottish Inpatient Patient Experience Survey 2012, volume 2: technical report*. <http://www.scotland.gov.uk/Publications/2012/08/5616/0>.
36. Wong EL, Coulter A, Cheung AW, et al. Validation of inpatient experience questionnaire. *Int J Qual Health Care* 2013;25:443–51.
37. Olteidal S, Garrat A, Bjertnaes O, et al. The NORPEQ patient experience questionnaire: data quality, internal consistency and validity following a Norwegian inpatient survey. *Scand J Caring Sci* 2007;35:540–7.
38. Wilde B, Larsson G, Larsson M, et al. Quality of care: development of a patient-centred questionnaire based on a grounded theory model. *Scand J Caring Sci* 1994;8:39–48.
39. Larsson BW, Larsson G. Development of a short form of the Quality from the Patient's Perspective (QPP) questionnaire. *J Clin Nurs* 2002;11:681–7.
40. Pettersen KL, Veenstra M, Guldvog B, et al. The Patient Experiences Questionnaire: development. *Int J Qual Health Care* 2004;16:463–63.
41. Webster TR, Mantopoulos J, Jackson E, et al. A brief questionnaire for assessing patient healthcare experiences in low-income settings. *Int J Qual Health Care* 2011;23:258–68.
42. Rao KD, Peters DH, Banteen-Roche K. Towards patient-centered health services in India—a scale to measure patient perceptions of quality. *Int J Qual Health Care* 2006;18:414–21.
43. Benneyan JC, Llyod R, Plesk P. Statistical process control as a tool for research and healthcare improvement. *Qual Saf Health Care* 2003;12:458–64.
44. Van der Vleuten C. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ* 1996;1:41–67.
45. Lynn MR. Determination and quantification of content validity. *Nurs Res* 1986;35:382–5.
46. Freeman T. Using performance indicators to improve health care quality in the public sector: a review of the literature. *Health Serv Manag Res* 2002;15:126–37.
47. Streiner DL, Norman GR. *Health measurement scales*. 3rd edn. Oxford Medical Publications, 2003.
48. Tenwee CB. *Protocol for systematic reviews of measurement properties*. KnowledgeCenter Measurement Instruments, 2011:13. <http://www.cosmin.nl/images/upload/files/Protocol%20klinimetriche%20review%20version%20nov%202011.pdf>
49. Mokkink L, Terwee C, Patrick D, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–49.
50. Coaley K. *An introduction to psychological assessment and psychometrics*. London: Sage publications, 2014.
51. Sargeant J. Reflecting upon multisource feedback as 'assessment for learning.' *Perspect Med Educ* 2015;4:55–6.
52. Larsson G, Larsson B W. Quality improvement measures based in patient data: some psychometric issues. *Int J Nurs Pract* 2003;9:294–9.
53. Sheng Sheng Y, Sheng Z. Is coefficient alpha robust to non-normal data? *Front Psychol* 2012;3:34.
54. Norman G. Likert scales, levels of measurement and the "laws" of statistics. *Adv Health Sci Educ Theory Pract* 2010;15:625–32.

BMJ Open

## Development and preliminary psychometric properties of the Care Experience Feedback Improvement Tool (CEFIT)

Michelle Beattie, Ashley Shepherd, William Lauder, Iain Atherton, Julie Cowie and Douglas J Murphy

BMJ Open 2016 6:  
doi: 10.1136/bmjopen-2015-010101

---

Updated information and services can be found at:  
<http://bmjopen.bmj.com/content/6/6/e010101>

*These include:*

- |                               |  |
|-------------------------------|--|
| <b>References</b>             | This article cites 33 articles, 12 of which you can access for free at:<br><a href="http://bmjopen.bmj.com/content/6/6/e010101#BIBL">http://bmjopen.bmj.com/content/6/6/e010101#BIBL</a>   |
| <b>Open Access</b>            | This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <a href="http://creativecommons.org/licenses/by-nc/4.0/">http://creativecommons.org/licenses/by-nc/4.0/</a> |
| <b>Email alerting service</b> | Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.   |

- 
- |                          |  |
|--------------------------|--|
| <b>Topic Collections</b> | Articles on similar topics can be found in the following collections<br><a href="#">Health services research (995)</a><br><a href="#">Medical management (163)</a> |
|--------------------------|--|

---

### Notes

---

To request permissions go to:  
<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:  
<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:  
<http://group.bmj.com/subscribe/>

### **5.3 Overview of Paper Five**

The following paragraphs provide additional details of the development and testing of CEFIT not covered in the associated publication, due to the limitations of the journal word count. This is followed by a critical reflection of the limitations of the study.

#### **5.3.1 CEFIT Item Construction**

A common approach to item construction is to produce multiple items (questions) in an attempt to adequately measure the construct of interest (Streiner et al 2015). Items are usually derived from existing instruments, literature reviews, expert opinion (gathered through focus groups or interviews) or clinical observations (Polit and Yang 2016, Streiner et al 2015). CEFIT was derived from the integrative literature review exploring current perceptions of quality of care (Paper Two) embedded within Chapter 2. Instrument developers usually begin with many more items than are thought to be required, which are subsequently reduced through various validity testing. The more items within an instrument, the more the potential that the construct of interest is being fully measured. However, CEFIT was constructed from the outset as one item for each of the five domains of quality identified in Beattie's Model of Healthcare Quality (see Appendix 5 for a copy of CEFIT). Developing as few items as possible assisted in the aim of creating a brief instrument. The integrative review actually found six domains constituting quality of care, namely, care which is; safe, effective, timely, caring, enables system navigation and is person-centred. However, person-centred care was inherent within every domain. Therefore, quality of care could be modelled as being unidimensional. Person-centred care was the central dimension, with the five remaining domains being components of it. So, there are only five domains within Beattie's model; to represent the patient experience of healthcare quality these five domains needed to be represented or enacted in a person-centred way.

Note that the decision to examine patient 'experience' to gather the patient perspective was informed by Paper Two (the ED Study) and the subsequent discussion in Chapter 3. For example, a prompt for determining timely care is "staff responded to my call bell within a reasonable time," as opposed to asking patients to evaluate the promptness of their care by rating how satisfied they were with waiting time. Items obtaining reports of whether, or the degree to which, patients experienced certain care processes or behaviour are thought to provide a more accurate account of care compared to patient satisfaction with care (Health Foundation 2013, Luxford 2012, Salisbury et al 2010).

Given that the items within Beattie's model were broad domains, prompts were designed to assist patients in interpreting them. For example, "I received procedures and treatments within acceptable waiting times" is a prompt for the item "I received timely care". Prompts or examples could potentially be adapted to fit different contexts. For example, using CEFIT in a general population survey may alter the timely prompt to "I waited an acceptable length of time for my appointment", or "I waited an acceptable length of time to be seen within the department". Hence, CEFIT has a unique design in that it has fixed domains (required for validity) with the flexibility of re-wording prompts, thus enabling simple adaptation to context.

### **5.3.2 CEFIT Rating Scale**

The type of rating scale designed will be dependent on the theory or attribute that is being measured (Streiner et al 2015). Rating responses for CEFIT were designed to measure the frequency of quality of care experienced by patients. One of the key aims for health services in Scotland is to have high quality of care for "every patient, every time" (Healthcare Improvement Scotland 2014, Scottish Government 2010). Therefore, CEFIT was devised with a five-point ordinal rating response scale (Never, Occasionally, Sometimes, Often, Always) in an attempt to measure the consistency of high quality of care, from the patient experience.

CEFIT was designed with a five-option response choice to ensure brevity and simplicity. The number of response options remains subject to debate. Statistically, there is evidence that reliability reduces as fewer categories are used, with little difference in reliability between seven to ten options (Streiner et al 2015). Participants have also reported preferring between five and nine response options for simplicity and brevity (Preston and Coleman 2000). This also fits with the findings of Miller's (1956) cognitive experiments where people were best able to judge difference if there was seven (plus or minus two) options. To some extent, brevity was chosen to favour user acceptability over reliability, with such losses expected to be minimal.

### **5.3.3 CEFIT Scoring**

Responses from each item should be scored as; never 1, occasionally 2, Sometimes 3, often 4, and always 5. All five items of CEFIT should be added together to obtain a patient experience quality of care score, as the domains of Beattie's model (safe, effective, timely, caring, system navigation and person-centred) are components of healthcare quality. Therefore, there would be a maximum score of 25 and a minimum

score of 5. Once data have been collected, scores should be summed for each item across the number of questionnaires. For example, Patient 1, question one might score 3, Patient 2, question one might score 5, and Patient 3, question one might score 4, therefore the total score for question one would be 12. Scoring CEFIT items in this way ensures that the scores for each question are normally distributed (assuming an adequate sample), regardless of individual distribution of scores. This is particularly important when using ordinal variables (as used in CEFIT) where the mean value is ambiguous (Carifio and Perla 2008). To ensure simplicity, any score less than 'always' requires attention as the aim is for reliable care for every patient, every time. Items with the lowest rating will likely be prioritised for improvement.

The initial drafts of CEFIT had an overall global rating scale of eleven possible values; where 0 = poor care experience and 10 = excellent care experience. However, feedback highlighted that users found the global rating difficult to interpret. Also, because CEFIT was designed to highlight key areas for improvement and action as opposed to an individual global score, the overall rating question was removed. Responses to individual items would be more useful for targeting areas for improvement than an overall assessment score. Also, having an overall score may suggest a 'grade' for judgement or scrutiny as opposed to highlighting areas for improvement.

Beattie's Model of Healthcare Quality (introduced in Chapter 2) suggests that quality of healthcare is defined by the domains of safe, effective, timely, caring, system navigation and person-centred. Where a construct is defined by the items themselves it is unlikely a single global question or item will adequately capture the construct of interest (Streiner et al 2015). This reaffirms that the use of a global rating scale to measure patient experience of hospital quality of care would be of little value, hence justifying its removal from the CEFIT. Instead, the construct of healthcare quality was composed of five domains, each represented by one question and its associated prompts.

#### **5.3.4 Content Validity Index Procedure**

Content validity is a judgement of whether or not an instrument adequately reflects the construct of interest, in this case the patient experience of hospital quality of care. There is suggestion in the literature that this judgement could be made more robust by applying quantification using the index of content validity (Lynn 1986). Once the instrument is assembled a pre-specified number of experts critique it, and the amount of agreement can be ascertained using a table devised by Lynn (1986, p. 384); it estimates the level of agreement required for content validity of the instrument to be confirmed. Eight out

of ten experts were required to rate the items as 3 or 4 (3 = relevant but needs minor alteration; 4 = very relevant and succinct) to achieve content validity beyond the 0.05 level of significance (Lynn 1986).

Five volunteers who had had a previous hospitalisation of more than 24 hours within the last six months completed the CVI procedure to provide a patient perspective of the CEFIT items. The CVI procedure was discussed with the local NHS Research and Development Office who judged the process to be service evaluation, as the NHS Board was currently seeking to improve the mechanism of patient feedback (see Appendix 6 for Patient Feedback Tool Ethical Decision Letter). Permission was sought from a hospital Consultant prior to the cardiac rehabilitation nurse asking patients if they would be interested on commenting on the feedback tool. Volunteers were asked at the end of their cardiac rehabilitation class if they would be willing to give their views on the CEFIT tool. Those willing to participate (five) completed the CVI at the end of their cardiac rehabilitation class (see Appendix 7 for Expert Feedback Form: Content Validity Index).

Experts in patient experience were identified using the list of attendees at the International Forum on Quality and Safety in Healthcare in London in 2013. Five experts working in the field of patient experience (researchers and practitioners) completed the same CVI procedure as public volunteers. Feedback from both patient and expert groups found that the CEFIT content was valid. The exercise also provided useful feedback for minor modifications to the instrument.

### **5.3.5 Queensland Social Survey**

The Queensland Social Survey (QSS) is an annual state-wide survey administered by the Population Research Laboratory at Central Queensland University (CQU) in Australia to explore a wide range of research questions relevant to the general public. The School of Health Sciences at the University of Stirling has a history of submitting research questions for inclusion into that survey. The CEFIT questions were proposed for inclusion to enable a quick, yet robust method to determine the internal consistency and structural validity of the instrument. Funding was obtained from the School of Health Sciences to submit the questions to the QSS. Pilot testing confirmed the suitability of the questions and the likelihood of obtaining an adequate sample for analysis.

### 5.3.6 Analysis and Results

To ensure that CEFIT is actually measuring the patient experience of hospital quality of care (validity) in a consistent manner (reliability), it is usual to examine this statistically, as well as theoretically (De Vet et al 2011). As discussed in Chapter 4, validity and reliability are not 'all or nothing' approaches, rather, they are cumulative, with each study furthering the robustness or otherwise of an instrument to quantify the construct of interest consistently (Bannigan and Watson 2009, Streiner et al 2015). Given that CEFIT had been designed theoretically and tested by patient experience experts using the CVI, the next step was to test the validity and reliability of the CEFIT structure, statistically.

The reliability (internal consistency) of the CEFIT structure was examined using Cronbach's alpha. Cronbach's alpha is used to calculate the consistency, or otherwise, of how respondents collectively answer the items/questions within an instrument (as discussed in Chapter 4). Item-correlations are calculated to examine how well each question relates to another (De Vet et al 2011, Streiner et al 2015). CEFIT data were entered into SPSS and Cronbach's alpha was calculated. The Cronbach's alpha is calculated for the structure as a whole and then SPSS determines the Cronbach's alpha if one of the items/questions were deleted. Item-correlations determine the usefulness, or otherwise, of individual items and Cronbach's alpha provides a measure of the consistency in which the instrument is responded to as a whole. Inter-item correlations between 0.2 and 0.8 signify a reasonable relationship, while correlations  $<0.2$  indicate the items are not related and  $>0.8$  suggest the item may be too similar (Streiner and Norman 2003). Items calculated as  $<0.2$  and  $>0.8$  would prompt developers to consider removal of the item from the measure. However, results need to be considered alongside the instrument's purpose and validity (Streiner et al 2015). CEFIT's five items were within the 0.2-0.8 range suggesting that all items were unique enough to justify retaining them. As indicated in Chapter 4 a Cronbach's alpha  $> 0.70$  indicates a positive result for internal consistency reliability (Terwee et al 2007). The Cronbach's alpha for CEFIT was 0.78; a positive result.

The validity of the CEFIT structure was then examined using factor analysis. As explained in Chapter 4, structural validity is examined statistically by conducting factor analysis to explore how many 'factors' are within an instrument or questionnaire. It is usually used when an instrument has too many items or to examine whether items could be grouped into domains. However, Beattie's Model of Healthcare Quality (described in Chapter 2) proposed that healthcare quality was one construct, containing five domains, articulated through 5 items (one for each component). Exploratory factor analysis was



not conducted to reduce or revise the items of CEFIT; rather, it was used to explore the factor structure. In statistics, factors are orthogonal, in other words, unrelated. The items within CEFIT are related; safe, effective, timely, caring, and system navigation are components of one construct (patient experience of quality of care), which need to be enacted in a person-centred way. It is predicted, based on the Beattie Model of Healthcare Quality, that CEFIT will have one factor (patient experience of quality of care) which is composed of aspects of safe, effective, timely, caring, and system navigation, with person-centeredness at the core of these domains.

As explained in Chapter 4, to be structurally valid, factors should explain at least 50% of the variance within an instrument (Terwee et al 2007). That is to say that at least 50% of the items within a questionnaire should be measuring the construct intended, such as patient experience of hospital quality of care. Eigenvalue is the statistic used to calculate the variance. For example, as CEFIT has five items and the first eigenvalue is 2.85 (detailed in Paper Five), that factor accounts for 2.85 of the variance. This is calculated as 2.85 divided by 5 (number of items within the questionnaire) which is 0.57 (this is usually represented as a percentage, therefore multiplied by 100) presented as 57%. This result is articulated in Paper Five as 57.33% variance of the one factor solution. Given that an instrument is said to be structurally valid if the variance is at least 50%, the eigenvalue result shows that CEFIT tests positive for structural validity. The 57.33% variance is shared by five the domains, namely; safety, timely, effective, caring and system navigation.

### **5.3.7 CEFIT Quality Critique**

A quality critique of all five aspects of instrument utility (validity, reliability, cost efficiency, acceptability and educational impact) was conducted for two reasons. Firstly, taking a holistic view of instrument utility will help direct efforts to design a measure of patient experience of hospital quality of care which is practical at ward-level use. Secondly, it enables a transparent review of the quality of CEFIT development and the necessary next steps. The COSMIN checklists for content validity, structural validity and internal consistency were used to assess the quality of the study *methods* used for the development of CEFIT (see Appendix 8, 9 and 10 for copies of COSMIN checklists and CEFIT results). The quality criteria for measurement properties was applied to judge the quality of the *results* of each psychometric test performed on CEFIT (see Appendix 11). The criteria for additional aspects of instrument utility (cost efficiency, acceptability and educational impact) were described in Chapter 4 and Paper Four. These were then applied to CEFIT (see Appendix 12 for criteria and CEFIT results). The results are

available in Table 5 of Paper Five; Cost was rated as Good, Acceptability was rated as Poor and Educational Impact was rated as Good. Acceptability was rated as poor as CEFIT has not yet been tested within an inpatient context.

## **5.4 Critical Reflection of Paper Five**

The development and preliminary psychometric testing of CEFIT was informed by all other aspects of this collection of works. Chapter 2 and Paper One informed the development of Beattie's Model of Healthcare Quality which provided the theoretical foundation for the construction of CEFIT. Chapter 2 also highlighted the evolving nature of the the concept of healthcare quality and hence the need to design an instrument which could be easily adapted to context. The fixed components and flexible prompts will enable relative ease of adaptation of CEFIT (highlighted as a necessity in Chapter 2 due to the evolving nature of what constitutes quality), although further tests of validation would be necessary. Chapter 3 and Paper Two (the ED study) informed the requirement to measure patient experience as opposed to satisfaction and hence CEFIT questions and responses were designed to measure patient *experience* of hospital quality of care. Chapter 4 and Papers Three and Four enabled the development of a wide view of instrument utility and an understanding of psychometrics which were used to inform the development and testing of a brief, yet structurally valid and reliable, CEFIT. Paper Four also provided systematic evidence that a gap existed for a brief measure of patient experience of hospital quality of care for use at the clinical microsystem level (i.e. the hospital ward) for the purpose of quality improvement. The following paragraphs provide a critical reflection of the learning and methods used in the development and preliminary psychometric testing of CEFIT.

### **5.4.1 Limitation Between Healthcare and Hospital Context**

The main limitation of the CEFIT study was that the structural validity and internal consistency were tested within an Australian population survey, which highlights the fact that the findings are limited to an Australian population with healthcare experience as opposed to inpatients. Streiner et al (2015) remind us of the limitations of context by stating that an instrument can never be truly said to be valid, but rather valid for the population and context in which it was tested (Streiner et al 2015). However, the QSS presented an opportunity to test the validity and reliability of the CEFIT structure with a large, random sample. Conducting the same tests utilising inpatients in a UK context would have been resource-intensive and outwith the scope of this collection of works without additional resource. However, it is more likely a future large-scale study will

obtain funding given that CEFIT now has some psychometric credibility. Validity is cumulative, therefore the more positive results for validity studies an instrument has the more trust users can have that the instrument is measuring what it intends to measure (Streiner et al 2015). The positive findings for internal consistency and structural validity suggest the resource for a UK-based inpatient study would be justified.

#### **5.4.2 Limitation of Range**

The descriptive statistics of CEFIT highlight that respondents' answers were mostly towards the high end of the response options, indicating that the majority of quality care processes were occurring 'often' or 'always.' There is a potential that the reduced range of responses could be similar to the high ceiling effect found in satisfaction surveys, which would limit the ability of CEFIT to differentiate between excellent and poor care experiences (Ahmed et al 2014, Fenton et al 2012, Greaves et al 2012, Haggerty 2010, Leonard 2008, Moret et al 2007, Salisbury et al 2010). However, all rating options were utilised in CEFIT, indicating the possibility of limiting the high ceiling effect. Also, other instruments with high ceiling effects have been able to differentiate between aspects of good and not-so-good care (Larsson and Larsson 2002, Mercer and Murphy 2008, Oltedal et al 2007, Rao et al 2006). While it is important to remain vigilant to the potential threat, the ability of CEFIT to differentiate between different experiences of quality of care will remain unknown until a future generalisability study is conducted.

#### **5.4.3 Statistical Methods**

There is some debate as to the use of Cronbach's alpha and factor analysis with skewed distribution, despite such use being common practice (Larsson and Larsson 2002, Mercer and Murphy 2008, Oltedal et al 2007, Polit and Yang 2016, Rao et al 2006). Given that CEFIT responses were mostly towards the high end of rating options, data were not normally distributed. A study by Sheng and Sheng (2012) used various statistical simulations to observe the effect of different distributions on Cronbach's alpha. Their findings suggest that Cronbach's alpha is affected by skewed distributions but that increased sample sizes help improve the accuracy of non-normal data. Similarly, Norman (2010) suggest that if we were to assure the assumption of normally distributed data for factor analysis (and some other tests) we would effectively dismiss about 75% of educational, health status and quality-of-life assessment. Norman (2010) also conducted various modelling with skewed data and concluded that parametric statistics can be used with Likert data with non-normal distributions.

In summary, it would appear that skewed distributions can affect the robustness of Cronbach's alpha and factor analysis but this has not been well investigated and remains subject to some debate (Gadernann et al 2012, Norman 2010, Sheng and Sheng 2012, Sullivan and Artino 2013). Whilst Cronbach's alpha and factor analysis are probably effected by non-normal data, the large sample ( $n=802$ ) used in the CEFIT study likely mitigates the effect on the results. Therefore, the statistical analyses used in the CEFIT study support the hypothesis that CEFIT questions are related and tap into a construct to measure patient experience of healthcare quality. The challenge for the future will be to ensure that skewed data and ceiling effects will not limit the capacity of CEFIT to discriminate reliably between care providers (as mentioned in above).

## **5.5 Study Contribution to the Research Question**

The development and psychometric testing of CEFIT presented in this Chapter and Paper Five has been the cumulative result of a series of investigations exploring measurement of the patient perspective of hospital quality of care. The work was triggered by the discrepancy between reported metrics of national and board-level hospital quality of care and the experiences of individual patients. A research gap was identified to develop a measure of patient experience of hospital quality of care for use at the clinical microsystem level (i.e. ward/unit). Paper Five presents a robustly developed tool designed to address this gap. CEFIT has been theoretically informed and developed from patients, patient experience experts and the literature. The approach to instrument design has been informed by empirical and theoretical knowledge. Paper Five has established a structurally valid and internal consistent measure of the patient experience of hospital quality of care, namely the Care Experience Feedback Improvement Tool (CEFIT). The uniqueness of CEFIT is the brevity and simplicity of the instrument, whilst so far meeting psychometric standards for validity and reliability. Hence CEFIT fills a current gap by devising a timely and relevant measure of hospital quality of care, from the patient experience of healthcare quality. Of course, instrument validation is an ongoing process and further studies are required to determine whether CEFIT can be used as a valid and reliable measure in a hospital context to measure ongoing improvement in clinical practice. However, the brevity and simplicity of CEFIT will increase the likelihood of this being a useful metric for improvement purposes. Criteria for all aspects of instrument utility will be used for the continued development of CEFIT.

# Chapter 6

## Discussion

### 6.1 Introduction

This PhD by publication and its associated narratives aimed to provide insight into the complexity of measuring the patient perspective of hospital quality of care in the NHS in Scotland. This collection of works began from an increasing awareness of, and discomfort about, the apparent disconnect between the reported metrics of national and board-level hospital quality of care and the experiences of individual patients. While most national and board-level reports were suggesting hospital quality of care was good, there was an increase in poor patient hospital care experiences, as related by patients themselves (Department of Health 2013a, Francis 2013, ISD 2014, Jha et al 2005, Right Care 2011). This insight resulted in the identification of a research gap for a timely and relevant measure of hospital quality of care to drive improvements, culminating in the development of CEFIT.

The complexity of healthcare creates a threat and challenge to ensuring patients receive high quality hospital care, recently accentuated by reducing NHS resource (Bevan 2016). One response has been an increase in measurement of healthcare quality to assure and to determine whether interventions are improving quality of care (Raleigh and Foot 2010). The last decade has also seen an increased focus on measuring aspects of healthcare quality from the patient perspective, in recognition of their unique perspective, and to direct efforts towards co-production and mutual health service design. The net result has been a proliferation in instruments (questionnaires) to measure the patient perspective of hospital care, each with varying degrees of utility (validity, reliability, cost efficiency, acceptability and educational impact).

Chapter 1 set out the current NHS Scotland Measurement Framework, revealing that, at the time of writing, quality was measured at the macro (National), meso (NHS Board) and micro (e.g. ward) levels of the healthcare system. The patient perspective of hospital quality of care was measured at the macro level via the national Inpatient Patient Experience Survey. There were no specific patient experience measures at the meso level, although the National Survey data were available at NHS Board and Hospital level. Finally, there were no specified measures of patient experience at the micro level; despite quality of care being subject to much scrutiny, there was a gap in measuring the

patient perspective of hospital quality of care, for quality improvement purposes, at the micro ward/unit level. This was despite the proliferation of instruments, mentioned above.

In this collection of works, Paper One was an integrative review of the literature to explore a contemporary meaning of healthcare quality. Paper Two was a cross-sectional survey in the ED to determine whether empathy (as an indicator of caring) could be effectively used as a measure of quality from the patient perspective. Paper Three was a protocol of the methods for a systematic review to identify and critique the utility of existing instruments which measure the adult inpatient experience of hospital quality of care. Paper Four was the systematic review to identify and critique the utility of existing instruments which measure the adult inpatient experience of hospital quality of care. Paper Five was the development and preliminary psychometric testing of the Care Experience Feedback Improvement Tool (CEFIT). This Chapter (Chapter 6) illustrates how the research objectives were met by summarising the main findings of this PhD by publication and its associated narratives and how the findings relate to the existing literature. The contribution to the field of improving hospital quality of care will be discussed. Limitations of the research will also be highlighted.

## **6.2 Summary of Findings and Contribution to Improving Hospital Quality of Care from the Patient Experience**

Chapter 1 detailed the four thesis objectives which the five Papers and their associated narratives set out to achieve. This section summarises and discusses, with the aid of the literature, the subsequent findings from each Paper, as detailed in Chapters 2 to 5, and how they addressed the four thesis objectives.

### **6.2.1 Objective 1: To determine what domains capture the contemporary meaning of healthcare quality to inform the development of a theoretical model of quality of healthcare.**

Paper One is presented in Chapter 2. It aimed to synthesise how healthcare quality is currently defined. Before anything can be measured it needs to be clearly defined. It had long been accepted that healthcare quality was so complex and diverse that it could not be defined by one simple sentence, rather, the concept was composed of multiple aspects or domains (Donabedian 1980, Maxwell 1984, Ovretveit 1992, IOM 2001, World Health Organisation (WHO) 2006). As discussed in Paper One and Chapter 2, the domains of healthcare quality defined by the Institute of Medicine's (IOM) remained the most widely accepted and utilised; safe, timely, efficient, effective, equitable and patient-

centred (IOM 2001). The historical overview of the concept of healthcare quality explored in Chapter 2 highlighted that, although similarities existed in all definitions (e.g. the necessity for caring behaviours), the concept of healthcare quality was ever evolving, dependent on changing contexts over time (Meyer and Bishop 2007). This necessitated a re-examination of the IOM domains of healthcare quality to ensure they remained representative of the current context. There was no pre-existing study synthesising various understandings of healthcare quality through an integrative review.

The review in Paper One found that two of the IOM domains were really outcomes of the other domains; two more domains could replace these now re-defined domains, and one domain was a key foundational concept, of which the other five were components. The two domains which were really outcomes were “efficient” and “equitable”. The two additional domains of healthcare quality were “caring” and “system navigation”. The study also found that person-centred care was foundational to all other domains. The five domains of healthcare quality therefore became care which is; safe, effective, timely, caring and allows system navigation. These were all components of a unidimensional conceptualisation of quality in healthcare as one that is person-centred.

Although person-centred care has been highlighted as a key aspect of quality of healthcare, no models of healthcare quality have acknowledged its foundational nature (Donabedian 1998, IOM 2001, Wilde et al 1994, Scottish Government 2010. Professor Don Berwick, a leader in healthcare improvement, described the polarised views on person-centred care when agreeing the domains of healthcare quality with the IOM Committee over a decade ago (Berwick 2009). He described the tensions between professional control and patient needs. The IOM eventually agreed to include person-centeredness as a key domain of healthcare quality. Yet Berwick (2009) highlighted the centrality of the concept: “Call it person-centeredness, but I suggest, this is the core: it is that property of care that welcomes me to assert my humanity and my individuality” (p. 564). Despite this importance, the IOM represented the domain of person-centeredness equally, alongside the other dimensions of healthcare quality (IOM 2001).

According to the findings of the integrative review, all domains must be enacted in a person-centred way to achieve high quality of care. The domains and relationship between the domains were articulated in Beattie’s Model of Healthcare Quality (as described in Chapter 2). To acknowledge the centrality of person-centeredness, this aspect was not represented as being equal to the other domains, but as the foundational central concept from which the others radiated. Whilst all other models of quality of care

included aspects of technical and interpersonal domains, no other model was found which represented the foundational nature of person-centred care. In addition, no other model had a separate domain for system navigation (Donabedian 1980, IOM 2001, Maxwell 1984). Caring had previously been included in some models as an essential domain of quality of care (Berg et al 2012, Coyle and Williams 2001, Wilde et al 1993).

There was also an acknowledgement in the Chapter that the concept of healthcare quality is continually evolving and this is likely to continue, potentially at an accelerated pace (NHS Confederation 2013, Scottish Government 2010). This has important implications for the validity of measures of the patient experience of hospital quality of care. Whilst many acknowledged the evolving nature of healthcare quality, there was a paucity of literature articulating the impact of this evolution on defining domains of quality of care (Donabedian 1980, IOM 2001, Maxwell 1984, Ovretveit 1992). Both Beattie's Model of Healthcare Quality and knowledge of the evolving nature of healthcare quality were used later in the thesis to develop CEFIT (Chapter 5). Paper One and Chapter 2 therefore achieved the objective of capturing the contemporary domains of healthcare quality and using these to develop a model of quality of care.

### **6.2.2 Objective 2: To test whether empathy, as an indicator of caring, can be measured in an acute hospital setting, from a theoretical and practical perspective**

Paper Two aimed to establish whether empathy could be used as a measure of quality of care from the patient perspective within the ED. The two new domains of Beattie's Model of Healthcare Quality, namely; *system navigation* and *caring*, are arguably more difficult to quantify than the other domains (safe, effective and timely). It was important to establish whether a healthcare quality domain that is more challenging to quantify can be measured in practice (Carr 2013). Being unable to measure these important aspects of healthcare quality could be part of the reason why patients' experiences of hospital quality of care differ from those reported in National surveys.

To test whether it was possible to measure an aspect of healthcare quality less amenable to measurement, empathy as an indicator of caring was selected for testing in a busy hospital environment (the ED). Specifically, the study aimed to determine whether patients' perceptions of clinicians' empathy could be effectively used as a measure of quality by assessing, firstly, whether CARE measure scores correlated with a measure which rated patient satisfaction and, secondly, whether this correlation was greater than any found for a measure of waiting time. No previous study was found which measured



empathy in an ED setting. Some studies were found exploring aspects of caring behaviours within the ED, but none investigated the potential of using caring as an indicator of quality of care from the patient perspective (Gordon et al 2010, Nerney et al 2001, Nystrom et al 2003, Perez-Carceles et al 2010, Wiman and Wikbladm 2004). There were many studies measuring aspects of waiting time in the ED (Booth et al 1992, Jolly and Clancy 2009, Jones and Schimanski 2010, Pitrou et al 2009, Storm-Versloot et al 2014, Woodcock et al 2013).

Paper Two found that the majority of patients reported care to be good (21%) or very good (75%). Waiting times varied between 11 minutes and 5 hours 17 minutes. CARE scores ranged from 12 to 50 (mean 41.1). The study found a statistically significant relationship between ratings of patient satisfaction and CARE measure scores with a moderate correlation (Spearman's  $\rho = 0.55$ ,  $p < 0.001$ ), whereas no statistically significant correlation was found between satisfaction and waiting time (Spearman's  $\rho = -0.07$ ,  $p = 0.56$ ). The findings indicate that CARE measures scores may be a useful indicator of hospital quality of care from the patient perspective. Conversely, waiting time was found to be of little value as an indicator of healthcare quality from the patient perspective.

High ratings of patient satisfaction with quality of care in the ED have been consistently reported in other studies (Boudreaux et al 2000, Perez-Carceles et al 2010). Similarly, studies using the CARE measure in other environments have reported mostly high scores (Mercer et al 2005, Mercer and Murphy 2008, Mercer et al 2008). Other literature reporting waiting times in the ED is variable. As found in this study, busier time periods do not necessarily result in less satisfied patients. Previous research has also indicated that perceived waiting time is a stronger predictor of patient satisfaction than actual waiting time (Boudreaux et al 2000, Boudreaux and O'Hea 2004, Pitrou et al 2009, Toma et al 2009). A likely explanation for the reduced relationship between waiting time and patient satisfaction is the improvements that have been made in wait time over the last decade. Patients are mostly seen, treated and discharged or admitted to hospital within 4 hours in the UK (Scottish Government 2011b, Department of Health 2011). Because what constitutes healthcare quality continually evolves, the importance of different domains of quality also changes. However, it would be short-sighted to banish the domain of time, as there is a risk that this would eventually result in increased waits for patients.

From a practical perspective, measuring empathy in the ED was relatively straightforward. This study required a 'red flag' tracking process to enable patients' empathy scores to be correlated with their waiting time, but this would not be necessary if the CARE measure became part of routine data collection. Resource demands for data collection would be minor, therefore, patients could be given the CARE measure when 'booking in' and completion boxes could be available within the department for patients to post completed questionnaires. Resource would be necessary, however, for data input and analysis.

The ED study also found a high ceiling effect of responses of patient satisfaction. This led to further reading around using patient satisfaction as a measure of quality of care (as detailed in Chapter 3). Although the limited range of scores had little effect on the results of the ED study as the question was related to the relationship between empathy scores and patient satisfaction, it was an important consideration in developing a timely and relevant measure of hospital quality of care from the patient perspective. This important finding re-focused the direction of this collection of works to measuring patient 'experience' as opposed to 'satisfaction' in relation to findings in the literature.

Overall, Paper Two demonstrated that those who considered their care to have been of high quality were also more likely to have perceived staff as being more empathetic. This finding suggests that empathy (CARE measure) is likely to be a valid indicator of healthcare quality from the patient perspective in the ED. Therefore, aspects of quality, which are more difficult to quantify, can be measured in ED and are therefore more likely to be measurable in other inpatient areas. The findings demonstrate that the same domain in Beattie's Model of Healthcare Quality (empathy as an indicator of caring behaviour) can indeed be measured within a busy hospital environment (ED). The study also highlighted the theoretical and statistical limitations of using satisfaction as a valid and reliable measure and redirected efforts to measure patient 'experience' as opposed to 'satisfaction' in order to measure hospital quality of care, from the patient perspective. The findings also demonstrate that the domain of Beattie's Model of Healthcare Quality (empathy as an indicator of caring behaviour) can indeed be measured within a busy hospital environment (ED).

### **6.2.3 Objective 3: To identify and critique the utility of existing instruments which measure the adult inpatient experience of hospital quality of care.**

Whilst there was a proliferation of instruments aiming to measure the patient perspective of hospital quality of care the psychometric properties of existing instruments had not

been systematically reviewed. There remained a need to establish whether an instrument to measure the patient experience of hospital quality of care was available for use at the clinical microsystem (i.e. ward level). A systematic review was conducted to achieve Objective three. The methods were published via a protocol (Paper Three) and the results via Paper Four.

The systematic review, reported in Paper Four, found 1,157 records within the health-related databases. Many instruments were excluded on the basis that they were measuring satisfaction as opposed to experience; other exclusion criteria were discussed in Paper Three. The process resulted in 26 papers being retained in relation to 11 instruments measuring the patient experience of hospital quality of care. The retained instruments had various psychometric tests conducted, and, although the quality of the methods and results was variable, they were mostly of a high standard. Every instrument had evidence of being examined for at least one aspect of validity and of reliability. Every instrument had tested content validity by exploring which aspects of hospital quality care mattered most to patients. All instruments had published other types of validity, except NHSIP and SIPES. All instruments studied internal consistency to determine the reliability of the instrument structure. However, similar literature reviews have found that studies do not report sufficient psychometric information to enable a full critique of instrument utility, although this has improved over the last ten years (Castle et al 2005, Groene et al 2013). A Utility Matrix Tool was developed as part of the review, to enable all aspects of utility to be weighted (validity, reliability, cost efficiency, acceptability and educational impact). Paper Four found enough reported psychometric information to critique the retained instruments, although some missing data may have resulted in studies being apportioned a lower score for study quality. For example, the NHSIP publication referred to previous structural validity work, but the detail required to judge criteria was unavailable (Sizmur and Redding 2012).

The systematic review found cost efficiency was rated as good for QPPS, NORPEQ and I-PAHC (Larsson and Larsson 2002, Oltedal et al 2007, Webster et al 2011). All other instruments were rated as poor or fair, highlighting that considerable or extensive resource would be required to obtain an adequate sample. All instruments, except QPP, were rated excellent or good for the utility component of acceptability. Only five instruments (HCAHPS, SIPES, NORPEQ, I-PAHC, PPQ) were rated as good for educational impact (Keller et al 2005, Levine et al 2005, Oltedal et al 2007, Rao et al 2006, Scottish Government 2012, Scottish Government 2010, Sofaer et al 2005, Webster et al 2011). No other studies were found for comparison that critiqued these

additional aspects of instrument utility. Castle et al (2005) conducted a literature review of instruments measuring patients' perceptions of hospital quality of care, but they included those measuring patient satisfaction, located minimal psychometric data and did not include other important aspects of instrument utility. They concluded that it would be beneficial to use a standardised survey and data collection procedure, but they did not highlight the necessity to utilise different instruments for different purposes. The Utility Matrix developed for the Systematic Review facilitates the choice of different instruments for different purposes.

Although the psychometric standard of instruments was generally of a high standard, those which were brief (<20 questions) were unsuitable for use at the micro level of the system (e.g. ward) for other reasons. For example, the Picker Patient Experience Questionnaire (PPE-15) has 15 questions, but is not intended to be used as a stand-alone instrument (Jenkinson et al 2002b). The PPE is a summary measure taken from an existing bank of questions. There was no instrument measuring patient experience of hospital quality of care as represented in Beattie's Model of Healthcare Quality, namely, care that is; safe, effective, timely, caring, enables system navigation, and which delivers all of these elements in a person-centred manner. The objective, to identify and critique the utility of existing instruments which measure the adult inpatient experience of hospital quality of care, was achieved, and results suggested that there are instruments available for use. The choice would be dependent upon the purposes for which the data would be used and the context in which it would be used. There remained a gap to devise a timely, relevant and brief measure of the patient experience of hospital quality of care, for use at the clinical microsystem level (e.g. hospital ward).

#### **6.2.4 Objective 4: To develop a valid, reliable and brief measure of patient experience of hospital quality of care.**

Paper Five described the development and preliminary testing of a brief measure of patient experience of hospital quality of care; the Care Experience Feedback Improvement Tool (CEFIT). CEFIT was devised from Beattie's Model of Healthcare Quality (as described in Chapter 2). Beattie's model contained five domains, care which is; safe, effective, timely, caring and enables system navigation. These domains need to be enacted in a person-centred way to achieve high quality of care. That is, quality of care can be defined as a unidimensional concept with five behavioural domains.

Initially, CEFIT was found to be positive for content validity using a content validity index procedure with patient expertise (previous inpatients and academics or leaders with

patient experience expertise). It was then tested in a telephone survey. Responses from the survey of 802 eligible participants (healthcare experience within the previous 12 months) were used to assess the internal consistency and structural validity of CEFIT, which were both found to be positive. Specifically, Cronbach's alpha coefficient for internal consistency indicated high reliability (0.78). Factor analysis confirmed a unidimensional scale (one factor solution) accounting for 57.3% variance. The 57.33% variance was shared by the five domains in Beattie's Model of Healthcare Quality; safety, timely, effective, caring, and system navigation, with inter-item total correlations suggested their necessity in measuring the patient experience of hospital quality of care (0.28–0.73). Using the COSMIN standards to judge the quality criteria of psychometric methods and results, CEFIT was found to be fair and positive for content validity, excellent and positive for structural validity, and excellent and positive for internal consistency reliability. Applying the criteria for additional aspects of instrument utility (designed and applied in the systematic review, detailed in Paper Four), CEFIT was rated as good for cost, poor for acceptability (as not yet tested in a hospital context) and good for educational impact.

The acknowledgement that healthcare quality is continually evolving and the likelihood of accelerated change (highlighted in Chapter 2) suggests that theoretical models of what constitutes healthcare quality require regular re-examination (NHS Confederation 2013, Scottish Government 2010). Some instruments have been criticised for their lack of theoretical development (Health Foundation 2013), and some current patient experience measures, such as the PPE-15, have relied on theoretical models of healthcare quality from decades ago (Jenkinson et al 2002b).

The CEFIT was developed from a current theoretical model of quality of healthcare. It has also been designed to take account of the evolving nature of healthcare quality, by creating core domains with flexible prompts within the instrument design, therefore enabling interpretation of the quality domains to suit changing contexts. Adaptation of instruments to suit varying contexts is not new; however, most questionnaires are adapted following initial validation and use, as opposed to building this in as an original design feature (Harkness 2010). Few instruments are designed with an opportunity for easy adaptation; an example of one which has been is the Household Food Insecurity and Access Scale (HFIAS), although this is not within the field of healthcare quality (Gebreyesus et al 2015). The systematic review (Paper Four) found no measure of patient experience of quality of hospital care designed with both fixed and flexible components. Therefore CEFIT is a unique design within the field; created with an

adaptable feature to suit the evolutionary nature of conceptualising quality of healthcare. Of course, any adaptation would still require validity testing, but the design feature of CEFIT makes the initial changes easy.

In summary, the development and preliminary testing of CEFIT described in Paper Five achieved objective 4 of the thesis by developing a valid, reliable and brief measure of patient experience of hospital quality of care. Of course, at this stage, CEFIT only has structural validity and reliability. Further psychometric testing would be required to establish whether CEFIT could reliability distinguish between those reporting different experiences of hospital quality of care.

### **6.3 Collective Contribution and Implications**

This collection of works has implications for the field of healthcare quality, specifically in relation to the patient perspective of quality of care and how it can be measured at the clinical microsystem level within hospitals. There are also wider implications in relation to the research methods. The following sections consider the contribution and implications of the works for measuring hospital quality of care and beyond. Implications for practice, policy and research are considered.

#### **6.3.1 Practice Implications**

The main contribution of this thesis for practice is the development and preliminary testing of a brief, yet valid and reliable, instrument to measure the patient experience of hospital quality of care; CEFIT. The evolution of CEFIT can be traced across a series of investigations reported in this PhD by publication. Given the increasing scrutiny of clinical practice and the increasing pressure to balance quality of care with cost, the need to measure the quality of hospital care will likely remain for the foreseeable future. An evolving society, with an increase in mutual decision-making and service design will demand inclusion of the patient perspective. Acting on quality of care issues, those raised by patients, could contribute to solving aspects of the protracted problem of poor hospital care.

Nurses, and other healthcare professionals, face many challenges when attempting to measure quality of hospital care for improvement purposes. Firstly, what is measured may be perceived as not measuring aspects of quality of care which are important to patients and nurses (valid and acceptable). The ways in which aspects of care are measured may not be trusted (reliable and acceptable). There are finite resources for

delivering and evaluating care, an ongoing tension in practice (cost efficiency). Finally, the feedback loop of findings informing practice is not always adequate (educational impact).

CEFIT identifies and partly addresses these practice challenges. CEFIT has been developed from what patients and clinicians think are important aspects of quality of care. The five domains of quality of care in CEFIT would intuitively connect with patients and nurses. The flexible prompts for each domain would enable further adaptation to context, thus enabling local staff to further develop and take ownership of the measure. The evidence of the robust development and the psychometric results would also give nurses and patients assurance of the validity of CEFIT. Ensuring clinicians and patients have trust in the measure is an essential prerequisite for effective use of the data (Davies 2005).

Similarly, once there is evidence to suggest that CEFIT can reliability differentiate across a scale of different quality of care experiences, clinicians will be more trusting of the tool. For example, it will be possible to stipulate the numbers of completed CEFITs needed to produce reliable data. Some staff are dubious of the reliability of some quality improvement measures, which is likely due to the variation in how samples are collected and analysed. Systematic and explicit methods to devise measures are necessary for credibility (Davies 2005).

The tensions of competing resources in clinical practice will likely continue. Nurses working in hospital wards have highlighted the challenge of the array of care processes that require regular measurement (Personal Communication 2015). It is essential not to add to this burden. The brevity and patient-completion mode of CEFIT will alleviate this challenge. Some resource will be required for data entry and analysis. However, once the adequate sample numbers of CEFIT are known, data input could be conducted on a monthly basis. Also, the simplicity of CEFIT could enable identification of areas of improvement without any complex analysis. Given that the aim of high quality care is for every patient, every time, then any results not achieving 'always' will require action.

Also, in order for CEFIT data to be used for improvement purposes it needs to be incorporated into an appropriate feedback loop. Systems for feedback for improvement will likely differ across contexts, but the simplicity of CEFIT will enhance the ability to achieve this. This is timely given the current policy ambition of creating care assurance systems in Scotland (discussed further under contributions to policy).

The use of numerous QI measures in practice has caused data fatigue as healthcare staff utilise limited resources for data collection, thus limiting the resource available to make improvements to the quality of patient care (Smith et al 2008). This is likely due to the fact that most QI measures have not been implemented using QI principles. Compliance with authority has been the order of the day, with management requesting that clinical staff (mostly nurses) collect and interpret an increasing amount of QI data (Giraud 2001, RCN 2016). Organisations who utilise this approach to quality improvement have been categorised as ‘prod organisations’ (Allcock et al 2015).

Many improvement measures are collected using a random small sample (up to 20) of patients to calculate a monthly percentage reliability to document on a run chart. For example, the peripheral venous cannula (PVC) maintenance bundle stipulates that nurses carry out and document five interventions (e.g. remove PVC where there is signs of extravasation or inflammation) on a daily basis for each patient fitted with a PVC. Bundles are evidence-based interventions which, when initiated collectively, have been shown to improve outcomes for patients (Resar et al 2005). Every month these data are used to calculate a percentage reliability. These data are then plotted on a run chart monthly to observe patterns over time. This is one of many measures that accumulate to form the existing burden of measurement. If compliance with measurement falls, a frequent management response is to demand more frequent auditing. The reasons for low compliance are usually not explored. It is of no surprise that gaming ensues to achieve an acceptable compliance rate. The measure then becomes an end in itself, as opposed to an improvement in care.

CEFIT does not require the same data collection and interpretation procedure. Given that one of the key aims for health services in Scotland is to provide a high quality of care for “every patient, every time”, patients who score ‘always’ for each of the five items of CEFIT would indicate a positive patient experience of quality of care (Healthcare Improvement Scotland 2014, Scottish Government 2010). It might operate thusly: the aim could be to have CEFIT completed by patients at the time of hospital discharge. Completed questionnaires could be collected anonymously via a collection box, placed at the exit of the ward. The Charge Nurse could rapidly review the CEFIT questionnaires to identify any patient not scoring ‘always’ for every item. Those scoring anything less than ‘always’ should prompt the team to reflect on the care provided. Reflections and plans for improvement could be embedded within existing practices, such as staff meetings, learning sessions, monthly case reviews, among others. Embedding new processes within existing systems is a recognised QI strategy to integrate and sustain



improvement (Leatherman et al 2010). The facilitation of reflective events are essential to ensure psychological safety, within an open and transparent culture focused on learning and improvement as opposed to judgement and scrutiny (Dewar et al 2010). This process would be enhanced by including two qualitative open questions to CEFIT to explore the reasoning behind patient ratings of care.

The example above is only one potential way of using CEFIT. Using a QI approach to testing and implementing CEFIT would be key to its success or otherwise. Rather than enforcing how a measure should work within hospital wards, it would be necessary to work with patients and nurses to establish the best way of using CEFIT. QI principles include engagement and involvement of key stakeholders at an early stage (Hughes 2008). Although CEFIT is not a 'finished product' it has sufficient validity and reliability to engage patients and staff in its further development. Identifying a ward willing to test and adapt CEFIT within their improvement activity is an essential next step. Although the actual CEFIT items would remain fixed (unless further psychometric testing warranted otherwise) all other aspects of the instrument use can be adapted to the local context, such as the flexible prompts to aid interpretation, how data are collected and analysed and, importantly, how the findings from CEFIT will direct quality improvement efforts. Evidence suggests that changing practitioners' behaviour requires data to provide evidence of the problem combined with an altruistic drive to make a difference. The altruistic drive of practitioners can be triggered by narratives of patient experience (Dewar et al 2010). Therefore, combining both quantitative data from CEFIT and the addition of narrative feedback via open-ended questions could help connect with the practitioners to initiate behaviour change.

It is also suggested that quality improvement demands the interaction between technical (QI methods such as Plan, Do, Study, Act cycles) and rationale (psychosocial) elements of change (James et al 2016). For example, asking nurses to test CEFIT using PDSA cycles (technical) and adapting their use in accordance with their findings (psychosocial) would draw nurses into the process of using CEFIT before its full implementation. Enabling ownership of the data and how they will be used is far more likely to result in a usable tool that would be effectively used in practice. There needs to be a QI approach to test and implement CEFIT into hospital wards. These local uses of CEFIT do not preclude using CEFIT as an improvement measure of overall quality standards, as the actual CEFIT items remain fixed.

The fact that the NHS Scotland measurement framework does not include a measure of the patient perspective of hospital quality at the micro level is likely contributing to the disparity between the reported metrics of national and board-level hospital quality of care and the experiences of individual patients. CEFIT offers the potential to reduce this gap by providing a timely and relevant measure of patient experience of hospital quality of care, for use at the microsystem level of healthcare. The unique design of CEFIT increases the likelihood of being a useful measure at the clinical interface due to its brevity, simplicity and ability to adapt to context.

As well as informing the development of CEFIT, Beattie's Model of Healthcare Quality offers a contemporary model of healthcare quality, which would be of interest to those who deliver, monitor and/or manage healthcare. As noted in Chapter 1; what gets measured matters in healthcare as domains of healthcare quality are often translated into measurement plans to improve the quality of hospital care, and what gets measured tends to attract resource. Beattie's Model of Healthcare Quality offers a new perspective to consider what gets measured in hospitals. It presents person-centred care as fundamental, with the new domains of caring behaviours and enabling system navigation possibly influencing a directional change in priorities for hospitals aiming to improve quality of care. To date, measurement plans have focused on easy-to-measure domains of quality of care (e.g. waiting time), but there is a need to consider consistent data capture of aspects of quality previously not included. CEFIT offers the potential for patient experience of hospital quality of care to be measured and weighed equally with other aspects of healthcare quality.

### **6.3.2 Policy Implications**

As detailed in Chapter 1, healthcare policy in Scotland influences the approach to measuring and improving the quality of hospital care. To date, Scottish healthcare policy on quality of care has largely been influenced by the IOM domains (Scottish Government 2010, Scottish Government 2011a). However, Paper One and Chapter 2 highlight the evolutionary nature of contemporary domains of healthcare quality. Beattie's Model of Healthcare Quality provides a contemporary framework of what constitutes quality of care. Adopting the Beattie Model of Healthcare Quality for NHS Scotland Policy would redirect the focus on measuring *caring* and *system navigation* as inclusive aspects of healthcare quality. The evolving nature would also require that NHS Scotland policymakers commission an analysis of what constitutes quality every few years in order for the measure to remain valid.

The necessity to measure all aspects of healthcare from a person-centred approach would redirect what and how aspects of care are measured. This differs from the current policy approach which is focused on the measurable and potentially outdated domains of healthcare quality. For example, if waiting time (say, for surgery) was to be enacted in a person-centred way then there would be a need to measure not only wait time but other potential unintended consequences (known as a balancing measure). This might include how long individuals who have already breached the acceptable limit continue to wait (not been seen/treated within the target time frame). Currently, patients who wait beyond the 12-week wait for surgery can then wait many more months for their operation. Once the target has been missed the patient is no longer a priority to be seen; instead, efforts are directed to reduce the likelihood of others breaching the 12-week wait. However, if Beattie's Model of Healthcare Quality were considered, wait time measures would need to include aspects of person-centred care, for example, what is the impact on the person waiting months for surgery? This might include being off work, with potential loss of earning, among other factors. The movement and flow of waiting lists could involve a different approach, if the "timely" element was properly considered from a person-centred perspective.

"Excellence in Care Deliverables" (Scottish Government 2015) is the NHS Scotland policy response to the Vale of Leven Inquiry, with regard to the future direction of nursing and midwifery care. The report commits to several "deliverables", one of which is a nationally agreed set of indicators for high quality of nursing care, inclusive of a measure of patient experience. The report also requires Health Boards to devise robust processes and systems for measuring, assuring and reporting quality of care (Scottish Government 2015). Various systems and dashboards have been devised to embed the measures within electronic databases. The Care Assurance and Accreditation System is one approach being tested in two Health Boards in Scotland, which aims to join up disparate measures into a robust assurance system (Ford 2015). There is a statement in the Report from the Chief Nursing Officer for Scotland which echoes calls for a practical, brief measure: "We mustn't squeeze the life out of people by imposing impossible bureaucratic burdens" (Scottish Government 2015, p. 5). The brief and relevant CEFIT offers a timely contribution to this important agenda.

An opinion piece published by the Royal College of Nursing (RCN 2016) questions the currency and sustainability of the NHS Scotland measurement framework and a review of a target driven approach. Whilst some interesting perspectives are shared, no simple solution is offered. The piece does not differentiate measurement at different levels of

the healthcare system, so does not address the disconnection between measures of patient experience at different levels of the healthcare system. Findings from the systematic review (Paper Four) highlight the need to select different patient experience measures of hospital quality of care for different purposes. The results of the utility matrix in the systematic review would help policymakers and practitioners to select the right tool for the right purpose. Although different measures are needed at different levels of the healthcare system, they also need to be connected. For example, improvements at the micro level should be feeding into hospital reports (meso level), which directly influence national results (macro level). Selecting appropriate tools and connecting the measures would help to bridge the gap between the reported metrics of national and board-level hospital quality of care and the experiences of individual patients.

### **6.3.3 Research Implications**

There are several research implications arising from this body of work, which require a mixed methods approach. Firstly, CEFIT must be tested with inpatients to establish how many CEFIT questionnaires need to be completed to obtain a reliable sample. This work should include the addition of two open-ended qualitative questions to the CEFIT to examine whether narratives captured can help direct local improvement efforts. This would require data to be collected from several hospital wards to conduct a generalisability study to determine whether CEFIT can differentiate between different experiences of quality of hospital care at a ward level. The same data would also be used to test the validity of the measure within an inpatient context. The next steps of CEFIT development would be largely depend on the results of that study. Testing CEFIT in this way has already been discussed with the local NHS Board, who are supportive of the study. A detailed proposal is necessary to establish costs, although it is expected that study costs will be minimal as CEFIT will be completed by inpatients and returned to a collection box. Resource will be necessary to input and analyse the data. The foundational work of this PhD by publication will increase the likelihood of obtaining funding to conduct the study within a hospital context. If CEFIT is found to be able to differentiate between different care experiences on a continuum, then there is a possibility of a wider array of research applications; for example, testing and adapting CEFIT prompts in different contexts and checking cultural validity. Of course, the ongoing nature of validity and reliability and the evolution of healthcare quality suggest the need for ongoing psychometric studies to develop CEFIT. The challenge will remain of balancing all items of instrument utility to ensure CEFIT remains a practical, usable tool in practice.

Second, there needs to be continual monitoring and evaluation of the relevance of the domains of CEFIT. This could be conducted by an expert working group on a routine (for example, 5-year) cycle. Again, tentative discussions have taken place with other researchers and those working in policy at the Scottish Government, but a detailed proposal and costings must be further developed.

Professor Don Berwick recently explored past conceptualisations of healthcare quality, describing these as eras (Berwick 2016). The first era was the assumption of quality derived from the privileged position of medicine, that is to say, clinicians and other healthcare professionals. The second era was the domination of clinical scrutiny, audit and judgement, which was driven by a market approach to healthcare. Both of these concepts were explored in *The Evolving Definitions of Healthcare Quality* (section 2.3 of this thesis). Berwick (2016) calls for a third era – the moral era – which includes the use of improvement science and a reduction of mandatory measures. There is a risk that CEFIT, if used incorrectly, may fall into era two, being reduced to measurement as the outcome as opposed to improving healthcare quality for patients. This thesis has argued throughout for the necessity to differentiate between measures for improvement and measures for scrutiny. It is far more likely that CEFIT will guide improvement if qualitative questions are embedded in it, to reduce the likelihood of a reductionist approach to measuring the patient experience of hospital quality of care. The programme of work necessary to further develop CEFIT must include this important qualitative aspect to ensure that the results are specific enough to drive local improvements.

In relation to some of the methods developed within this body of work, the systematic review (Paper Four) was the first of its type to identify and critique the utility of instruments measuring the patient experience of hospital quality of care, hence offering a unique contribution in the field of critiquing the quality of instruments within systematic reviews. Within the systematic review (Paper Four), additional criteria were devised, tested and applied to critique the cost, educational impact and acceptability of existing instruments. Further testing and development of these criteria are necessary to ascertain the reliability between raters to apply and score the criteria. The criteria could be tested on other systematic reviews of instruments. Establishing their use in other subject areas would extend the contribution of the additional criteria beyond the field of quality of healthcare. Similarly, no established method of synthesising the quality of the methods and results of psychometric studies existed; therefore, a method was devised and represented as the Beattie and Murphy Utility Matrix. The matrix is a unique contribution in terms of offering a method to critique and synthesise psychometric studies

within systematic reviews, which will potentially aid users to select the right instrument for the right purpose.

## **6.4 Limitations**

Details of limitations from individual Papers were given in the relevant sub-sections on critical reflection in each Chapter and are therefore not repeated here. The purpose of this section, which forms part of the overall discussion of the thesis, is to identify the broad limitations of this body of work and suggest how these limitations can be addressed in future research.

### **6.4.1 Testing CEFIT in an Australian Survey**

The main limitation of this collection of works is that CEFIT has not yet been tested within a hospital context. Of course, context matters, therefore this limits the findings of the internal consistency and structural validity to an Australian population with a healthcare experience. There is no guarantee therefore that inpatients within Scotland would respond to CEFIT in the same way. Streiner et al (2015) remind us of the limitations of context by stating that an instrument can never be truly said to be valid, but rather, is valid only for the population and context in which it was tested.

However, the initial development and content validity index procedure was completed by patient experience experts from Scotland. The Queensland Survey presented a good opportunity to test the structural validity and reliability of the CEFIT with a large, random sample. The positive results indicate the potential for large-scale testing in a Scottish hospital context. Preliminary testing of all new instruments usually starts with testing the internal consistency reliability and structural validity before embarking on further psychometric testing (Hesselink et al 2013). In other words, the structure of the instrument needs to be valid and reliable before further psychometric testing. Given that validity and reliability are cumulative, the results of the structural testing of CEFIT provide a positive foundation on which to build and develop the instrument. Information was given in the Research Implications (section 6.3.1) of the necessary next steps.

### **6.4.2 CEFIT Scoring on Additional Aspects of Instrument Utility**

All elements of instrument utility (structural validity, internal consistency reliability, cost efficiency, acceptability and educational impact) were rated as 'good' or 'excellent', except for acceptability. Some response options for the acceptability critique rated CEFIT as 'excellent', but as CEFIT had not yet been tested in the context for which it

was designed (hospital ward), the overall rating of acceptability was 'poor'. Overall, scores for each category are determined by taking the lowest rating score of all questions, hence an overall rating for acceptability as poor. It will not be known whether CEFIT will be acceptable to users (patients, clinicians and managers) until it has been tested in this context.

Similar to validity and reliability, application of the additional aspects of instrument utility have the potential for giving higher ratings for 'older' instruments which have a longer history and accurate reporting of development. Although mature instruments have potential advantages, it is important not to dismiss newer instruments with only early development. Also, whilst instruments with extensive histories can be a strength, there are also potential limitations to be aware of. The systematic review found evidence of some items being added to instruments to measure areas of interest within healthcare policy (Paper Four). For example, the NHS Inpatient Survey included questions which were not rated as important by patients, but useful for other purposes. Questions on 'noise at night' were included because they were thought to be useful for the Healthcare Commission reviews of hospital performance, despite patients evaluating the item as having low importance (Boyd 2007). There is a risk that the instrument becomes an evaluation of policy implementation as opposed to the patient experience of hospital quality of care. It remains imperative that a measure of patient experience of quality of care is derived from what matters most to patients (Coulter et al 2009, LaVela and Gallan 2014). To reduce any threat to the validity of the instrument, those elements which constitute quality of care from the patient perspective needs to be re-explored every few years, as detailed in the Research Implications section (6.3.3).

## **6.5 In Summary**

In summary, the objectives of the research were met. That is, domains representing contemporary patient experiences of quality of care were identified, a domain of quality difficult to quantify was measured in practice, and the utility of instruments available to measure the patient experience of hospital quality of care were critiqued. The studies accumulated to inform the development of a structurally valid, reliable, yet brief measure of patient experience of hospital quality of care. The key implications for policy, and research arising from this body of work, are as follows:

- National surveys of patient experience are not sufficiently sensitive, nor timely enough, to measure of quality of care at the micro level of the healthcare system, hence necessitating other measures, i.e. CEFIT.

- Choosing an instrument to measure the patient experience of hospital quality of care requires a balanced consideration of all aspects of instrument utility (validity, reliability, cost efficiency, acceptability and educational impact). Using the Beattie and Murphy Utility Index will aid selection.
- Quality of care domains must be contemporary and therefore regular re-evaluation of what constitutes quality is necessary to inform revisions of key domains.

While the individual studies and reported Papers have limitations, the collection of works still offer a robustly and transparently developed instrument to measure the patient experience of hospital quality of care. Lessons learnt from all of the limitations will inform and improve the future research and development of CEFIT. Limitations of the work can be addressed by conducting a generalisability study to determine the number of completed CEFIT questionnaires needed for a reliable sample, with further testing of validity in an inpatient setting. Doing so would provide a much needed measure of patient experience for use at the micro level of the healthcare system

The final Chapter will detail brief conclusions and the dissemination of the findings. The Chapter will detail the contribution of authorship of the included Papers, as well as explaining the standing of the journals in which the papers were submitted.



# Chapter 7

## Conclusion and Dissemination

### 7.1 Final Thoughts

Improving the quality of hospital care remains a practice and policy imperative in Scotland and beyond (DoH 2008, IOM 2001, QIPP 2011, Scottish Government 2010). Measurement is fundamental to this aspiration (Scottish Government 2015). The Care Experience Feedback Improvement Tool (CEFIT) and its associated Papers offer a timely contribution to filling the gap of measuring the patient experience of hospital quality of care at the clinical microsystem. Nurses have a legal and moral duty to continue to improve the quality of hospital care, but require the right tools to do so. The timely and relevant measure of CEFIT has the potential to help frontline staff measure the quality of hospital care from the patient perspective. Doing so could provide an alternative insight and assist the patient voice to be heard in efforts to improve their experience of hospital quality of care. I am looking forward to building on this foundational work to continue to contribute to the field of improving the quality of hospital quality of care. The final section in this thesis considers the appropriateness of the target journal, author contributions of the included Papers, and details of their impact.

### 7.2 Standing of the Journals and Contribution to Published Works

To demonstrate the individual contribution of the author to each Paper and the wider contribution in the field of measuring the patient experience of quality of care, the following paragraphs detail the standing of the journals and a statement of authorship. Details of the impact of each Paper are also considered.

#### 7.2.1 Paper One: An integrative review of dimensions of quality (Beattie et al 2012)

This Paper was published in the *Journal of Research in Nursing (JRN)*. The *JRN* is a peer-reviewed journal in nursing with a specific focus around policy and practice. The target audience is nurses in practice, policy and research. Each issue of the journal contains a collection of papers with a specific focus. The integrative review was initially published online in 2012 before being published in the paper version in June 2013 when the topic of focus was 'Quality and Safety'. As the integrative review aimed to conceptualise a contemporary understanding of quality of healthcare with subsequent

implications for policy and practice, the *JRN* was an appropriate target journal. The findings of the integrative review have implications for the readership of *JRN*; for example, frontline nurses considering what constitutes quality of care and implications of their practice. Also, these findings have implications for policy-makers and managers to consider what domains of quality are important to include in hospital measurement plans. The journal does not report an impact factor, but reports a similar Scimago Institutions Rankings (SIR) of 0.242. However, the journal does score 7/10 for research and theory within Scopus (Sage 2016). Whilst there are higher ranking journals, such as the *Journal of Advanced Nursing (JAN)*, they would be unlikely to publish the study given the methodological limitations of an integrative review. Given all of these factors, the *JRN* was the most appropriate journal for this study.

The *JRN* provides an editorial commentary for each journal addition. In this commentary, the integrative review was described as ‘arguing convincingly’ for the necessity to include the additional dimensions of ‘caring’ and ‘system navigation’ for a modern conception of quality of healthcare (McMahon 2013). The Paper was also selected for a review piece written by Professor Carr, Professor of Nursing in Canada. Professor Carr described the study as a coherent and well written paper which makes an important contribution in the field of healthcare quality. Limitations were also acknowledged in terms of the ambitious attempt to represent the plurality of perspectives relevant to defining quality of healthcare, as well as the limitation of applying the inclusion criteria to titles and abstracts only. Prof Carr concluded by stating that the real challenge is to translate the domains of quality into measureable criteria (Carr 2013). Up to the end February 2016, the Paper has had ten citations in other peer-reviewed journals. Permission was granted from Sage publications to use the PDF version of Paper One within the thesis (see Appendix 13: Approval to use Paper One in Thesis).

### **Author Contributions for Paper One**

MB designed and conceived the study. MB refined the search strategy and retrieved and input papers to RefWorks. MB devised inclusion criteria and applied criteria to all titles and abstracts. AS conducted the duplicate check of the inclusion criteria for 10% of the included papers. BH provided direction for the study methods in his role as PhD supervisor. MB drafted the Paper and amendments were suggested by BH and AS. Estimated percentage contribution to the Paper is: MB 85%, AS 10%, and BH 5%.

### **7.2.2 Paper Two: A cross-sectional study measuring empathy (Beattie et al 2012)**

This study aimed to establish whether the elusive concept of empathy could be measured as an indicator of healthcare quality in the Emergency Department. The Paper was originally submitted to the *Journal of Evaluation in Clinical Practice (JECP)* but rejected before peer review. However, the editor recommended submitting the Paper to the *International Journal of Person Centered Medicine (IJPCM)*, where it was accepted. The *IJPCM* has a multi-disciplinary audience and focuses on the development of theory and practice of Person-Centered Medicine. One of their areas of interest is methods for the evaluation of person-centered care, which suited the cross sectional study enquiry of whether or not empathy was an indicator of healthcare quality. The journal publishes quarterly and was launched in 2011. The journal does not, as yet, report an impact factor. However, the fact that the journal was in its infancy enabled rapid publication of the Paper (within 12 weeks of submission) within an international, subject-specific journal. Disappointingly, as of February 2016, the Paper has only been cited once. This is in part likely to be due to the fact that the journal is not open access and its narrow focus reduces the likelihood of institutions paying for access. Authorisation to use the Paper within the thesis is included in the final paragraph of the Licence to Publish (see Appendix 14: Approval to use Paper Two in Thesis).

#### **Author Contributions for Paper Two**

MB conceived and designed the study. IA and WL assisted with statistical analyses. BM assisted with data collection. MB completed data collection over a nine-day period. MB wrote the Paper and all others contributed to drafts before agreeing the final version. Estimated percentage contribution to the Paper is: MB 85%, IA 5%, BM 5%, and WL 5%.

### **7.2.3 Paper Three: A protocol for systematic review and utility critique (Beattie et al 2014)**

The protocol aimed to develop the methods to conduct a systematic review with a utility critique and was published in *Systematic Reviews*. *Systematic Reviews* publishes high quality systematic reviews within healthcare. This includes rapid reviews, methods papers, protocols, as well as full systematic reviews. The journal does not yet report an impact factor, but is expected to have one within the next 18 months. However, the journal is highly regarded with an editor who is renowned in the field of systematic reviews. Professor Moher is one of the authors of the PRISMA (Preferred Reporting of Items for Systematic Reviews and Meta-Analyses) Statement and its associated

checklists which are used internationally. Authors are required to submit a PRISMA statement and register the review with PROPSERO (Prospectively Registered Systematic Reviews) prior to the Paper being considered for publication. The journal offers a transparent publication process, inclusive of open peer review and publication of all draft manuscript versions alongside the publication. *Systematic Reviews* was the first choice journal for the protocol.

The journal is open access but requires payment for article processing (£1,565 per article in 2015). Funding was obtained from the University of Stirling's Article Processing Charges (APC) Fund. The journal also provides rapid publication; publishing the protocol within 10 weeks of submission despite necessary revisions from the peer review process. The article has been accessed online 15,588 times over a two-year period (January 2014 until January 2016). The Paper also has an Altmetric score of five, which is an average score for articles published for the same length of time and scored by Altmetric. Altmetric reports the number of times a scholarly article is mentioned across the Web, including newspapers and social media, such as Twitter. The article has been cited by 19 authors since publication. Permission was granted from BioMed Central to use the PDF version of Paper Three within the thesis (see Appendix 15: Approval to use Papers Three and Four in Thesis).

### **Author Contributions for Paper Three**

MB conceived and designed the study, devised search strategies, drafted the inclusion selection form and drafted the manuscript. WL participated in study design, statistical advice, piloting of inclusion selection form and revision of manuscript. IA participated in study design, piloting of inclusion selection form and revision of the manuscript. DM provided direction for the study idea and design, provided statistical advice and helped revise the manuscript. All authors have read and approved the final manuscript. Estimated percentage contribution to the Paper is: MB 85%, WL 5%, IA 5%, and DM 5%.

### **7.2.4 Paper Four: A systematic review and instrument utility critique (Beattie et al 2015)**

The systematic review of instruments measuring the patient experience of hospital quality of care was also published in *Systematic Reviews*. This helped link the protocol to the study and be explicit about any deviations from the methods within the protocol. The Paper was reviewed by two international experts in healthcare psychometrics.

Professors Terwee and Mokkink are both authors of the COSMIN checklists, which are used internationally to critique the quality of psychometric instruments in health (Mokkink et al 2010). The Paper required major revision around the development of the utility matrix. The reviewers' feedback helped refine and improve the synthesis of the quality of the methods and results of psychometric studies. Reviewers commented that the Paper makes an important contribution in the field of healthcare psychometrics.

The rapid review and publication processing enabled the Paper to be published online within 4 months of the original submission. The Paper has been accessed 3,223 times within 6 months of publication. The Altmetric score is 27, which is in the top 5% of all research output scored by Altmetric. There has been National and International interest in the Paper, with Twitter demographics noting 50% interest from the UK; whilst the other countries include Canada and Poland. E-mails have also been received from researchers in Spain and Amsterdam and policy-makers from the Australian Commission on Safety and Quality in Health Care expressing interest in the review. Permission was granted from BioMed Central to use the PDF version of Paper Four within the thesis (see Appendix 15: Approval to use Papers Three and Four in Thesis).

#### **Author Contributions for Paper Four**

MB conceived and designed the study, devised search strategies, applied inclusion criteria, applied quality scoring, developed the matrix and drafted the manuscript. DM provided direction for the study idea and design, provided statistical advice, applied quality scoring and helped devise matrix and the manuscript. IA participated in the study design, piloting of inclusion selection form and revision of the manuscript. WL participated in the study design, provided statistical advice, applied inclusion criteria, applied quality scoring and revision of the manuscript. All authors read and approved the final manuscript. Estimated percentage contribution to the Paper is: MB 70%, WL 5%, IA 5%, and DM 20%.

#### **7.2.5 Paper Five: Development and preliminary testing of CEFIT (Beattie et al 2016)**

The Paper described the development and preliminary psychometric testing of an instrument to measure patient experience of hospital quality of care, namely, CEFIT (Care Experience Feedback Improvement Tool). The Paper was submitted to *BMJ (British Medical Journal) Open*. The journal publishes medical research from all disciplines, inclusive of psychometrics. The journal encourages submissions from

research which directly addresses patient outcomes or the practice and delivery of healthcare, which fits the aim of the CEFIT instrument. The journal has an impact factor of 2.271 and is prestigiously associated with the *BMJ*. Publishing CEFIT in this journal will be likely to influence the credibility of CEFIT amongst clinicians, policy-makers and researchers.

*BMJ Open* also operates an open peer review process ensuring fair and transparent decision-making. There is an open access fee of £1,620 which was covered by a successful application to the University of Stirling's APC Fund. The journal rejects 43% of submitted papers.

The Paper was submitted to the *BMJ Open* in November 2015 and has subsequently been peer reviewed. Necessary revisions have been made and the Paper was resubmitted in January 2016. The Paper was accepted for publication in April 2016.

#### **Author Contributions for Paper Five**

MB and DM conceived and designed the CEFIT instrument. MB designed the theoretical model of healthcare quality. WL and IA contributed to the thinking and development of the work in their role as MB's PhD supervisors. MB and DM designed the study. AS facilitated acquisition of data via the Queensland survey. MB and DM designed and collected data for the CVI. WL and IA conducted statistical analysis and interpretation. JC helped in result interpretation and statistical revision. MB drafted the manuscript which was critically revised by all authors before agreeing the final version of the manuscript. Estimated percentage contribution to the Paper is: MB 65%, AS 5%, WL 5%, IA 5%, JC 5%, and DM 15%.

## References

*Adults with Incapacity (Scotland) Act 2000*. Available:

<http://www.legislation.gov.uk/asp/2000/4/contents> [Accessed: 15<sup>th</sup> September 2015].

Ahmed, F., Burt, J. and Roland, M. (2014) Measuring patient experience: concepts and methods, *The Patient – Patient-Centered Outcomes Research*, 7 (3), pp. 235-241.

Aiken, L., Clarke, S., Sloane, D., Lake, E. and Cheney, T. (2008) Effects of hospital care environment on Patient Mortality and Nurse Outcomes, *Journal of Nursing Administration*, 38, pp. 223-229.

Allcock, C., Dormon, F., Taunt, R., & Dixon, J. (2015). Constructive comfort: accelerating change in the NHS. Policy analysis.

Allen, J., Hutchinson, A., Brown, R. and Livingston, P. (2014) Quality care outcomes following transitional care interventions for older people from hospital to home: a systematic review. *BMC Health Services Research*, 14 (1), pp. 346.

Al-Zaru, I., Oweis, A. and Gharaibeh, H. (2013) Cultural definitions of quality of care: perspectives of Jordanian patients. *Journal of Research in Nursing*, 18 (4), pp. 307-317.

Attree, M. (2001) Patients' relatives' experiences and perspectives of 'good' and 'not so good' quality care. *Journal of Advanced Nursing*, 33 (4), pp. 456-460.

Bannigan, K. and Watson, R. (2009) Reliability and validity in a nutshell. *Journal of Clinical Nursing*, 18 (23), pp. 3237-3243.

Barelds, A., Van de Goor, I. and Van Heck, G. (2009a) Quality of care and service trajectories for people with intellectual disabilities: defining the aspects of quality from the client's perspective. *Scandinavian Journal of Caring Sciences*, 24 (1), pp. 164-174.

Barelds, A., Van de Goor, I., Bos, M., Van Heck, G. and Schols, J. (2009b) Care and some trajectories for people with intellectual disabilities: defining course and quality determinants from the client's perspective. *Journal of Policy and Practice Intellectual Disabilities*, 6 (3), pp. 163-172.

Barjenbruch, T., O'Connor, L. and Reichwage, D. (2002) Defining quality of care. *Registered Dental Hygienist*, 22 (6) pp. 41-42.

- Bate, S., Mendel, P. and Robert, G. (2008) *Organising for Quality: The Quality Improvement Journeys of Leading Hospitals in Europe and the United States*. Oxford: Radcliffe Publishing.
- Berg, G.M., Ekengren, F., Lee, F.A., Acuna, D. and Burdsal, C. (2012) Patient satisfaction with surgeons in a trauma population: Testing a structural equation model using perceptions of interpersonal and technical care. *Journal of Trauma & Acute Care Surgery*, 72 (5), pp. 1316-1322.
- Berry, N., Gardner, T. and Anderson, I. (2015) *On Targets: How Targets Can Be Most Effective in the English NHS*. London: Health Foundation.
- Berwick, D. (2009) What 'patient-centered' should mean: confessions of an extremist. *Health Affairs*, 28 (4), pp. 555-565.
- Berwick, D.M. (2016) Era 3 for medicine and health care. *JAMA*, 315, pp. 1329-1330.
- Bevan H (2016) *Helen Bevan on Cost and Quality*. NHS Institute for Innovation and Improvement. Available: [http://www.institute.nhs.uk/quality\\_and\\_value/introduction/article\\_19a.html](http://www.institute.nhs.uk/quality_and_value/introduction/article_19a.html) [Accessed: 23<sup>rd</sup> January 2016].
- Black, N. and Jenkinson, C. (2009) Measuring patients' experiences and outcomes. *British Medical Journal*, 339, pp. b2495.
- Bland, M. (2001) *An Introduction to Medical Statistics*. 3rd Ed. Oxford: Oxford University Press.
- Bleich, S.N., Özaltın, E. and Murray, C.J. (2009) How does satisfaction with the health-care system relate to patient experience? *Bulletin of the World Health Organization*, 87 (4), pp. 271-278.
- Blumenthal, D. (1996) Quality of health care Part 4: The origins of the quality of care debate. *New England Journal of Medicine*, 335 (15), 1146-1149.
- Booth, A.J., Harrison, C.J. and Gardener, G.J. (1992) Waiting times and patient satisfaction in the accident and emergency department. *Archives of Emergency Medicine*, 9 (2), pp. 162-168.



- Boudreaux, E., Ary, R., Mandry, C. and McCabe, B. (2000) Determinants of patient satisfaction in a large, municipal ED: the role of demographic variables, visit characteristics, and patient perceptions. *American Journal of Emergency Medicine*, 18 (4), pp. 394-400.
- Boudreaux, E. and O'Hea, E. (2004) Patient satisfaction in the emergency department: a review of the literature and implications for practice. *The Journal of Emergency Medicine*, 26 (1), pp. 13-26.
- Boulding, W., Kalra, A., Staelin, R. and Zeithaml, V. (1993) A dynamic process model of service quality: from expectations to behavioural intentions. *Journal of Marketing Research*, Feb, pp. 7-27.
- Boyd J (2007) *The 2006 inpatients importance survey*. Oxford: The Acute Coordination Centre for the NHS Acute Patient Survey Programme, Picker Institute Europe.
- Bradshaw, S. A., Playford, E. D. and Riazi, A. (2012) Living well in care homes: a systematic review of qualitative studies. *Age and Ageing*, 41 (4), pp. 429-440.
- Braunstein, J. (2003) Defining quality diabetes care in the new health system. *Current Diabetes Report*, 3 (4), pp. 269-271.
- Brook, R. (1973) Critical issues in the assessment of quality of care and their relationships to HMOs. *Journal of Medical Education*, 48, pp. 114-134.
- Broome, M.E. (1993) Integrating literature reviews for the development of concepts. In: B. Rodgers and K. Knaf, eds. *Concept Development in Nursing*. Second Edition. Philadelphia: WB Saunders.
- Bryman, A. (2006) Integrating quantitative and qualitative research: how is it done? *Qualitative Research*, 6 (1), pp. 97-113.
- Buchanan, D.R. (1992) An Uneasy Alliance: Combining Qualitative and Quantitative Research. *Health Education Quarterly*, 19 (1), pp. 117-135.
- Cameron, P., Schull, M.J. and Cooke, M.W. (2011) A framework for measuring quality in the emergency department. *Emergency Medicine Journal*, 28, pp. 735-740.

Care Act 2014. Available:

<http://www.legislation.gov.uk/ukpga/2014/23/contents/enacted> [Accessed: 17<sup>th</sup> November 2015].

Care Quality Commission (2015) *The Independent Regulator of Health and Social Care in England: About us*. Available: <http://www.cqc.org.uk/content/about-us> [Accessed: 3rd February 2015].

Carifio, J. and Perla, R.J. (2008) Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences*, 3, pp. 106-116.

Carr, E. (2013) Review: Do the Institute of Medicine's (IOM) dimensions of quality capture the current meaning of quality in health care? – An integrative review. *Journal of Research in Nursing*, 18 (4), pp. 305-306.

Carroll, S. (2005) Is patient safety synonymous with quality nursing care? Should it be? A brief discourse. *Quality Management in Health Care*, 14 (4), pp. 229-233.

Cassidy-Smith, T. Baumann, B. and Boudreaux, E. (2007) The disconfirmation paradigm: throughput times and emergency department patient satisfaction. *Journal of Emergency Medicine*, 32 (1), pp. 7-13.

Castilla, J., Hernandez, J., Cabello, Y., Lafuente, A., Pajuelo, N., Marqueta, J. and Coroleu, B. (2008) Defining poor and optimum performance in an IVF programme. *Human Reproduction*, 23 (1), pp. 85-90.

Castle, N.G., Brown, J., Hepner, K.A., & Hays, R.D. (2005) Review of the literature on survey instruments used to collect data on hospital patients' perceptions of care. *Health Services Research*, 40 (6p2), pp. 1996-2017.

Cattell, J.M. (1921). *American Men of Science: A Biographical Directory*. 3rd ed. Garrison, New York: The Science Press.

Centers for Medicare and Medicaid (CMS) (2013) *CAHPS Hospital Survey (HCAHPS) quality assurance guidelines*. Baltimore: CMS.

Centre for Reviews and Dissemination (2009) *Systematic Reviews: CRD Guidance for Undertaking Reviews in Health Care*. University of York. Available: <https://www.york.ac.uk/crd/guidance/> [Accessed: 29<sup>th</sup> September 2015].

Chou, S., Chen, T., Woodard, B. and Yen, M. (2005) Using SERVQUAL to evaluate quality disconfirmation of nursing service in Taiwan. *Journal of Nursing Research* (Taiwan Nurses Association), 13 (2), pp. 75-84.

Chowdhury, S. (2003) *Power of Design for Six Sigma*. Chicago: Kaplan Professional Publishers.

Cleary, P.D., Edgman-Levitan, S., Roberts, M., Moloney, T.W., McMullen, W., Walker, J.D. and Delbanco, T.L. (1991) Patients evaluate their hospital care: a national survey. *Health Affairs*, 10 (4), pp. 254-267.

Cleary, P., Edgman-Levitan, S., McMullen, W. and Delbanco, T.L. (1992) The relationship between reported problems and patient summary evaluations of hospital care. *Quality Review Bulletin*, Feb, pp. 53-59.

Cleary, P.D., Edgman-Levitan, S., Walker, J.D. Gerteis, M. and Delbanco, T.L. (1993) Using patient reports to improve medical care: a preliminary report from 10 hospitals. *Quality Management in Health Care*, 2 (1), pp. 31-38.

Coaley, K. (2014) *An introduction to psychological assessment and psychometrics*. London: Sage Publications.

Collignon, P., Looke, D., Ferguson, J., McLaws, M. and Olesen, D. (2002) Surveillance definitions for multi-resistant organisms (MROs). *Australian Infection Control*, 7 (3), pp. i-iv.

Cornwell, J. and Goodrich, J. (2009) Exploring how to measure patients' experience of care in hospital to improve services. *Nursing Times*, 105 (29), p. 12-15.

Coulter, A., Fitzpatrick, R., Cornwell, J. (2009) *The Point of Care Measures of Patients' Experience in Hospital: Purpose, Methods and Uses*. London: The Kings Fund.

Coyle, J. and Williams, B. (1999) Seeing the wood for the trees: defining the forgotten concept of patient dissatisfaction in the light of patient satisfaction research. *Leadership in Health Services*, 12 (4), pp. 1-9.

Coyle, J. and Williams, B. (2001). Valuing people as individuals: Development of an instrument through a survey of person-centredness in secondary care. *Journal of Advanced Nursing*, 36 (3), pp. 450-459.

Critical Appraisal Skills Programme (2013) *Critical Appraisal Skills Programme: Making Sense of Evidence*. CASP Checklists. Available: <http://www.casp-uk.net/#!/casp-tools-checklists/c18f8> [Accessed: 20<sup>th</sup> November 2015].

Creswell, J. (2007) *Qualitative Inquiry & Research Design: Choosing Among Five Approaches*. Thousand Oaks, CA: Sage Publications.

Crosby, P. (1979) *Quality is Free*. New York: McGraw-Hill.

Crow, R., Gage, H., Hampson, S., Hart, J., Kimber, A., Storey, L. and Thomas, H. (2002) The measurement of satisfaction with healthcare: implications for practice from a systematic review of the literature. *Health Technology Assessment*, 6 (32), pp. 1-244.

Crown Office (2008) *Death and the Procurator Fiscal: Information and Guidance for Medical Practitioners*. Available: [http://www.sudiscotland.org.uk/wmslib/PDFs/Death\\_and\\_the\\_Procurator\\_Fiscal.pdf](http://www.sudiscotland.org.uk/wmslib/PDFs/Death_and_the_Procurator_Fiscal.pdf) [Accessed 17<sup>th</sup> February 2015].

Curry, L., Nembhard, I. and Bradley, E. (2009) Qualitative and mixed methods provide unique contributions to outcomes research. *Circulation*, 119, pp. 1442-1452.

Darwin, C. (1872) *The Origin of Species by Means of Natural Selection; or the preservation of favoured races in the struggle for life*. *Nature*, 5, pp. 318-319.

Davies, H. (2006) *Measuring and reporting the quality of health care: issues and evidence from the international research literature*. Discussion Paper. Edinburgh: NHS Quality Improvement Scotland.

Davies, H., Nutley, S. and Powell, A. (2002) *Regulating for improvement in public services: a review of activity and issues*. St Andrews: Centre for Public Policy and Management, University of St Andrews.

Deming, E. (1986) *Out of the Crisis*. Cambridge, MA: MIT Press.

Department of Health (2008) *High Quality Care for All*. Gateway reference: 10106.

Department of Health (2010a) *The Health and Social Care Act 2008: Code of Practice on the Prevention and Control of Infections and Related Guidance*, London: Stationery Office.

Department of Health (2010b) *Equity and excellence: Liberating the NHS*. Available: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/213823/dh\\_117794.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213823/dh_117794.pdf) [Accessed 11<sup>th</sup> April 2012].

Department of Health. (2011) *A&E Clinical Quality Indicators: Best Practice Guidance for local publication*. Available: [http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_127776](http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_127776) [Accessed: 11<sup>th</sup> April 2012].

Department of Health (2013a) *A Promise to Learn – A Commitment to Act, Improving the Safety of Patients in England*. National Advisory Group on the Safety of Patients in England.

Department of Health (2013b) *Patients First and Foremost: The Initial Government response to the report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. Available: [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/170701/Patients\\_First\\_and\\_Foremost.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/170701/Patients_First_and_Foremost.pdf) [Accessed: 5<sup>th</sup> November 2014].

Dewar, B., Mackay, R., Smith, S., Pullin, S., and Tocher, R. (2010) Use of emotional touchpoints as a method of tapping into the experience of receiving compassionate care in a hospital setting. *Journal of Research in Nursing*, 15 (1), pp. 29-41.

Dewar, B. and Nolan, M. (2013) Caring about caring: developing a model to implement compassionate relationship centred care in an older people care setting. *International Journal of Nursing Studies*, 50 (9), pp. 1247-1258.

De Vet, H., Terwee, C., Mokkink, L. and Knol, D. (2011) *Measurement in Medicine*. New York: Cambridge University Press.

Dingwall, R., Rafferty, A. and Webster, C. (1988) *An Introduction to the Social History of Nursing*. London: Routledge.

Donabedian, A. (1980) *Explorations in Quality Assessment and Monitoring, Vol.1, The Definition of Quality and Approaches to its Assessment*. Ann Arbor, MI: Health Admin Press.

Donabedian, A. (1998) *History of Health Services Research Project Interview with Avedis Donabedian*. National Information Centre on Health Services Research and

Health Care Technology. Available:

<http://www.nlm.nih.gov/hmd/nichsr/donabedian.html> [Accessed: 6<sup>th</sup> November 2014].

Doyle, C., Lennox, L., and Bell, D. (2013) A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. *British Medical Journal Open*, 3: e001570.

Dr Foster Limited (2010) *Intelligent Board 2010: Patient Experience*. Available:

<http://www.healthcaregovernance.org.au/docs/the-intelligent-board-patient-experience-2010.pdf> [Accessed: 15<sup>th</sup> December 2015].

Edwards, J.R. and Bagozzi, R.P. (2000) On the nature and direction of relationships between constructs and measures. *Psychological Reviews*, 5 (2), pp. 155-174.

Egbert, J. and Sanden, S. (2014) Theoretical frameworks: you can't have a how without a why. In: J. Egbert and S. Sanden. *Foundations of Education Research: Understanding Theoretical Components*. New York: Routledge, pp. 57-72.

Elliott, M.N., Lehrman, W.G., Goldstein, E., Hambarsoomian, K., Beckett, M.K. and Giordano, L.A. (2010) Do hospitals rank differently on HCAHPS for different patient subgroups? *Medical Care Research & Review*, 67(1), pp. 56-73.

Fenton, J., Jerant, A., Bertakis, K. and Franks, P. (2012) The cost of satisfaction: a national study of patient satisfaction, health care utilization, expenditures, and mortality. *JAMA Internal Medicine*, 172 (5), pp. 405-411.

Firth-Cozens, J. and Cornwell, J. (2009). *Enabling compassionate care in acute hospital settings*. London: The King's Fund.

Ford, S. (2015) Trial of new Scottish nursing care quality system underway. *Nursing Times*, 11<sup>th</sup> May. Available: <http://www.nursingtimes.net/roles/nurse-managers/trial-of-new-scottish-nursing-care-quality-system-underway/5084887.fullarticle> [Accessed: 15<sup>th</sup> December 2015].

Francis, R. (2013) *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. London: The Stationery Office.

Gadermann, A.M., Guhn, M. and Zumbo, B.D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*, 17 (3), pp. 1-13.

Gailey, F. and Cachia, P. (2010) *Patient Safety: Cost Implications of Adverse Health Events*. NES Patient Safety Multidisciplinary Steering Group, NHS Education for Scotland. Available:

<http://www.nes.scot.nhs.uk/media/6472/PS%20Cost%20Briefing%20Paper.pdf>

[Accessed: 3<sup>rd</sup> November 2014].

Galton, F. (1874) On men of science, their nature and their nurture. *Proceedings of the Royal Institution of Great Britain*, 7, pp. 227–236.

Gebreyesus, S.H., Lunde, T., Mariam, D., Woldehanna, T. and Lindtjørn, B. (2015) Is the adapted Household Food Insecurity Access Scale (HFIAS) developed internationally to measure food insecurity valid in urban and rural households of Ethiopia? *BMC Nutrition*, 1:2.

Gerteis, M., Edgman-Levitan, S. and Delbanco, T.I. (1993) *Through The Patient's Eyes: Understanding and Promoting patient centred care*. San Francisco: Jossey-Bass.

Gillam, S. and Siriwardena, A. (2013) Frameworks for improvement: clinical audit, the plan-do-study-act cycle and significant event audit. *Quality in Primary Care*, 21 (2), pp. 123-130.

Giordano, L., Ellito, M., Goldstein, E., Lehrman, E. and Spencer, P. (2010) Development, Implementation and Public Reporting of the HCAHPS Survey. *Medical Care Research Review*, 67, pp. 27-37.

Giraud, A. (2001) Accreditation and the quality movement in France. *Quality in Health Care*, 10 (2), pp. 111-116.

Go, A.S., Hylek, E.M., Phillips, K., Borowsky, L.H., Henault, L.E., Chang, Y., Selby, J. and Singer, D. (2000) Implications of stroke risk criteria on the anticoagulation decision in nonvalvular atrial fibrillation: the Anticoagulation and Risk Factors in Atrial Fibrillation (ATRIA) study. *Circulation*, 102 (1), pp. 1-13.

Gordon, J., Sheppard, L.A. and Anaf, S. (2010). The patient experience in the emergency department: A systematic synthesis of qualitative research. *International Emergency Nursing*, 18 (2), pp. 80-88.

Grassley, J., Nelms, T. (2009) Tales of resistance and other emancipatory functions of storytelling. *Journal of Advanced Nursing*, 65 (11), pp. 2447-2453.

- Gray, D.E. (2013) Theoretical perspectives and research methodologies. In: D.E. Gray, *Doing Research in the Real World*. London: Sage Publications, pp. 15-38.
- Greaves, F., Pape, U., Lee, H., Smith, D., Darzi, A., Majeed, A. and Millet, C. (2012) Patients' Ratings of Family Physician Practices on the Internet: Usage and Associations With Conventional Measures of Quality in the English National Health Service. *Journal of Medical Internet Research*, 14 (5), e46.
- Greenhalgh, T., Robert, G., Bate, P., Kyriakidou, O., Macfarlane, F. and Peacock, R. (2004) *How to spread good ideas: A systematic review of the literature on diffusion, dissemination and sustainability of innovations in health service delivery and organisation*. London: National Coordinating Centre for the Service Delivery and Organisation (NCCSDO). Available: [http://www.netscc.ac.uk/hsdr/files/project/SDO\\_FR\\_08-1201-038\\_V01.pdf](http://www.netscc.ac.uk/hsdr/files/project/SDO_FR_08-1201-038_V01.pdf) [Accessed: 24<sup>th</sup> August 2015].
- Griffiths, R. (1983) *Griffiths Report on NHS October 1983*. Available: [www.sochealth.co.uk/national-health-service/griffiths-report-ocotber-1983/](http://www.sochealth.co.uk/national-health-service/griffiths-report-ocotber-1983/) [Accessed: 26<sup>th</sup> November 2015]
- Groene, O., Botje, D., Suñol, R., Lopez, M. A. and Wagner, C. (2013) A systematic review of instruments that assess the implementation of hospital quality management systems. *International Journal for Quality in Health Care*, 25 (5), pp. 525-541.
- Haggerty, J. (2010) Are measures of patient satisfaction hopelessly flawed? *British Medical Journal*, 341, p. c4783.
- Haggerty, J., Burge, F., Levesque, J., Gass, D., Pineault, R., Beaulieu, M. and Santor, D. (2007) Operational definitions of attributes of primary health care: consensus among Canadian experts. *Annals of Family Medicine*, 5 (4), pp. 336-344.
- Hall, J.A., Irish, J.T., Roter, D.L., Ehrlich, C.M. and Miller, L.H. (1994) Satisfaction, gender and communication in medical visits. *Medical Care*. 32, pp. 1216-1231.
- Harkness, J. (2010) Adaptation. Cross-Cultural Survey Guidelines. Available: <http://ccsq.isr.umich.edu/pdf/07AdaptationNov2010.pdf> [Accessed: 21<sup>st</sup> October 2015].
- Hart, R.G., Pearce, L.A. and Aguilar, M.I. (2007) Meta-analysis: antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. *Annals of Internal Medicine*, 146 (12), pp. 857-867.



Hass, J., Cook, E.F., Puopolo, A., Burstin, H., Cleary, P. and Brennan, T. (2000) Is the professional satisfaction of general internists associated with patient satisfaction? *Journal of General Internal Medicine*, 15, pp. 122-128.

*Health and Social Care Act 2008*. Code of Practice on the Prevention and Control of Infections and Related Guidance. Available: <http://www.legislation.gov.uk/ukpga/2008/14/contents> [Accessed: 17<sup>th</sup> November 2015].

The Health Foundation (2011) *Evidence Scan: Levels of harm*. Available: <http://www.health.org.uk/public/cms/75/76/313/2593/Levels%20of%20harm.pdf?realName=PYiXMz.pdf> [Accessed: 5<sup>th</sup> November 2014].

The Health Foundation (2013) *Evidence Scan 18: Measuring patient experience*. London: The Health Foundation.

Health and Safety Executive (2013) *RIDDOR – Reporting of Injuries, Diseases and Dangerous Occurrences Regulations*. Available: <http://www.hse.gov.uk/riddor/> [Accessed: 17<sup>th</sup> February 2015].

Healthcare Improvement Scotland (2014) *Highlights of the Scottish Patient Safety Programme National Conference – Driving Improvements in Patient Safety*. Available: <http://www.scottishpatientsafetyprogramme.scot.nhs.uk/Media/Docs/Conference%202014/Scottish%20Patient%20Safety%20Programme%20Conference%20Highlights.pdf> [Accessed: 13<sup>th</sup> October 2015].

Healthcare Improvement Scotland (2016) *Reporting a Suicide*. Available: [http://www.healthcareimprovementscotland.org/our\\_work/mental\\_health/suicide\\_reviews/reporting\\_a\\_suicide.aspx](http://www.healthcareimprovementscotland.org/our_work/mental_health/suicide_reviews/reporting_a_suicide.aspx) [Accessed: 14<sup>th</sup> March 2016].

Heenan, M., Khan, H. and Binkley, D. (2010) From boardroom to bedside: how to define and measure hospital quality. *Healthcare Quarterly*, 13 (1), pp. 55-60.

Hendriks, A., Vrieling, M.R., Van Es, S., De Haes, H. and Smets, E. (2004) Assessing inpatients' satisfaction with hospital care: should we prefer evaluation or satisfaction ratings? *Patient Education & Counseling*, 55 (1), pp. 142-146.

Hesselink, G., Kuis, E., Pijnenburg, M. and Wollersheim, H. (2013) Measuring caring culture in hospitals: a systematic review of instruments. *BMJ Open*, 3:e003416.

Hewitson, P., Skew, A., Graham, C., Jenkinson, C. and Coulter, A. (2014) People with limiting long-term conditions report poorer experiences and more problems with hospital care. *BMC Health Services Research*, 14 (33), pp. 1.

Holbrook, M.B. and Corfman, K.P. (1985) Quality and Value in the Consumption Experience: Phaedrus Rides Again. In: J. Jacoby and J. Olson (eds.), *Perceived Quality: How consumers view stores and merchandise*. Lexington, MA: Lexington Books, pp. 31-57.

Hospital Authority (2010) *Patient Satisfaction Survey in Hong Kong Hospital Authority: Final Report*. Hong Kong.

Howard, J A. and Sheth, J.N. (1969) *The Theory of Buyer Behaviour*. London: John Wiley and Sons, Inc.

Howie, J., Heaney, D. and Maxwell, M. (2004) Quality, core values and the general practice consultation: issues of definition, measurement and delivery. *Family Practice*, 21 (4), pp. 458-468.

Hu, L. and Bentler, P. (1999) Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6 (1), pp. 1-55.

Hughes, R.G. (2008) Tools and strategies for quality improvement and patient safety. In: R.G. Hughes (ed.), *Patient Safety and Quality: An Evidence-Based Handbook for Nurses*. Rockville, MD: Agency for Health care Research and Quality.

Hyder, A. (2002) Structural adjustment in health in Pakistan: Defining the questions. *International Journal of Epidemiology*, 31 (2), pp. 509.

Iedema, R., Allen, S., Britton, K., Piper, D., Baker, A., Grbich, C., Allan, A., Jones, L., Tuckett, A., Williams, A., Manias, E. and Gallagher, T. (2011) Patients' and family members' views on how clinicians enact and how they should enact incident disclosure: the "100 patient stories" qualitative study. *British Medical Journal*, 343: d4423.

Information Services Division (2010) *Quality Measurement Framework*. NHS Scotland Services, Edinburgh. Available: <http://www.isdscotland.org/Health-Topics/Quality-Measurement-Framework/> [Accessed: 5<sup>th</sup> December 2015].

Information Services Division (2014) *NHS Complaints Statistics Scotland 2013/14. Publication Report*. Edinburgh: NHS National Services Scotland.

Information Services Division (2015) *Inpatient and Day Case Activity, Hospital Care*. NHS Scotland Services, Edinburgh. Available: <http://www.isdscotland.org/Health-Topics/Hospital-Care/Inpatient-and-Day-Case-Activity/> [Accessed: 17th November 2015].

Institute of Medicine (1990) *Medicare: A Strategy for Quality Assurance*. Washington DC: National Academy Press.

Institute of Medicine (1999) *To Err is Human: Building a safer healthcare system*. Washington DC: National Academy Press.

Institute of Medicine IOM (2001) *Crossing the Quality Chasm: A New Health System for the 21<sup>st</sup> Century*. Washington DC: National Academy Press.

James, B., Beattie, M., Shepherd, A., Armstrong, L. (2016) Time, fear and transformation: student nurses' experiences of doing a Practicum (Quality Improvement Project) in Practice, *Nurse Education in Practice*. doi: 10.1016/j.nepr.2016.05.004.

Jenkinson, C., Coulter, A., Bruster, S., Richards, N. and Chandola, T. (2002a) Patients' experiences and satisfaction with health care: results of a questionnaire study of specific aspects of care. *Quality and Safety in Healthcare*, 11, pp. 225-339.

Jenkinson, C., Coulter A. and Bruster, S. (2002b) The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. *International Journal of Quality Health Care*, 14, pp. 353-358.

Jha, A., Li, Z., Orav, J. and Epstein, A. (2005) Care in US hospitals – the Hospital Quality Alliance Program. *The New England Journal of Medicine*, 353 (3), pp. 265-274.

Johnston, G., Crombie, K., Davies, H., Alder, E. and Millard, A. (2000) Reviewing audit: barriers and facilitating factors for effective clinical audit. *Quality in Health Care*, 9, pp. 23-36.

Jolly, E. and Clancy, M. (2009) Waiting times are they that important? A patient survey, *Emergency Medicine Journal*, October, 26, pp. 726.

- Jones, P. and Schimanski, K. (2010) The four hour target to reduce emergency department 'waiting time': a systematic review of clinical outcomes, *Emergency Medicine Australia*, 22, pp. 391-298.
- Juran, J.M. (1967) *Management of Quality Control*. New York: McGraw-Hill Book Co. Inc.
- Kano, N. (1984) Attractive quality and must-be quality. *Journal of the Japanese Society for Quality Control*, 14 (2), pp. 39-48.
- Kaplan, S. and Ware, J. (1995) The patient's role in health care and quality assessment. In: N. Goldfield and D. Nash, eds. *Providing Quality of Care: Future Challenges*. 2nd ed. Ann Arbor, MI: Health Administration Press, pp. 25-52.
- Keller, S., O'Malley, A., Hays, R.D., Mathew, R.A., Zaslavsky, A.M., Hepner, A.M., Hepner, K.A. and Cleary, P.D. (2005) Methods used to streamline the CAHPS Hospital Survey. *Health Services Research*, 40, pp. 2057-2077.
- Khan, K., Kunz, R., Kleijnen, J. and Antes, G. (2003) Five steps to conducting a systematic review. *Journal of the Royal Society of Medicine*, 96 (3), pp. 118-121.
- Lagu, T., Goff, S.L., Hannon, N.S., Shatz, A. and Lindenauer, P.K. (2013) A mixed-methods analysis of patient reviews of hospital care in England: implications for public reporting of health care quality data in the United States. *Joint Commission Journal on Quality and Patient Safety*, 39 (1), pp. 7-15.
- Lang, E. (2004) The why and how of evidence-based medicine. *Medical Journal of Malaysia*, 8, pp. 90-94.
- Larrabee, J. and Bolden, L. (2001) Defining patient-perceived quality of nursing care. *Journal of Nursing Care Quality*, 16 (1), pp. 34-60.
- Larsen, R. and Warne, R.T. (2010). Estimating confidence intervals for eigenvalues in exploratory factor analysis. *Behavior Research Methods*, 42, pp. 871-876.
- Larsson, B.W. and Larsson, G. (2002) Development of a short form of the Quality from the Patient's Perspective (QPP) questionnaire. *Journal of Clinical Nursing*, 11 (5), pp. 681-687.
- Laszlo, E. (1991) *The Systems View of the World: A Holistic Vision for Our Time*. Creskill, NJ: Hampton Press.

- LaVela, S. and Gallan, A. (2014) Evaluation and measurement of patient experience. *Patient Experience Journal*, 1 (1), pp. 28-36.
- Leatherman, S., Ferris, T.G., Berwick, D., Omaswa, F., and Crisp, N. (2010) The role of quality improvement in strengthening health systems in developing countries. *International Journal for Quality in Health Care*, 22 (4), pp. 237-243.
- Leonard, K. (2008) Is patient satisfaction sensitive to changes in the quality of care? An exploitation of the Hawthorne effect. *Journal of Health Economics*, 27 (2), pp. 444-459.
- Lepège, A. and Verdier, A. (1995) The adaptation of health status measures: methodological aspects of the translation procedure. *International use and performance of health-related quality of life instruments. Oxford: Oxford Rapid Communication*, 93, pp. 101.
- Levine, R.E., Fowler, F.J. and Brown, J.A. (2005) Role of cognitive testing in the development of the CAHPS Hospital Survey. *Health Services Research*, 40 (6p2), pp. 2037-2056.
- Linn, L.S., Cope, D.W. and Leake, B. (1984) The effect of gender and training of residents on satisfaction ratings by patients. *Medical Education*, 59, pp. 964-966.
- Lopez, L., Weissman, J., Schneider, E., Weingart, S., Cohen, A. and Epstein, A. (2009) Disclosure of hospital adverse events and its association with patients' ratings of the quality of care. *Archives of Internal Medicine*, 169, pp. 1888-1894.
- Luxford, K. (2012) What does the patient know about quality? *International Journal of Quality in Health Care*, 24, pp. 439-440.
- Lynn, M.R. (1986) Determination and quantification of content validity. *Nursing Research*, 35 (6), pp. 382-385.
- McColl, E., Jacoby, A., Thomas, L., Soutter, J., Bamford, C., Steen, N., Thomas, R., Harvey, E., Garratt, A. and Bond, J. (2001) Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. *Health Technology Assessment*, 5 (31).
- McDonagh, M., Peterson, K. and Raina, P. (2013) Avoiding bias in selecting studies in *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Agency for

Healthcare Research and Quality. Available:

<http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318> [Accessed 23<sup>rd</sup> August 2015].

McDowell, I. and Newell, C. (1996) *Measuring Health: A Guide to Rating Scales and Questionnaires*, second ed. Oxford: Oxford University Press.

MacLean, R.N.M. (2014) *The Vale of Leven Hospital Inquiry Report*, Published on behalf of The Vale of Level Hospital Inquiry by APS Group. Available: <http://www.valeoflevenhospitalinquiry.org/Report/j156505.pdf> [Accessed: 8<sup>th</sup> August 2015].

McMahon, A. (2013) Editorial: Making a difference. *Journal of Research in Nursing*, 18 (4), pp. 283-287.

Maxwell, R.J. (1984) Quality assessment in health. *British Medical Journal (Clinical Research Education)*, 288 (6428), pp. 1470-1472.

Maxwell, R.J. (1992) Dimensions of quality revisited: from thought to action. *Quality in Healthcare*, 1 (3), pp. 171-177.

Medicines and Healthcare Products Regulatory Agency (2014) *Yellow Card Scheme*. Available: <https://yellowcard.mhra.gov.uk/> [Accessed 5<sup>th</sup> February 2015].

*Mental Health Act 2007*. Available:

<http://www.legislation.gov.uk/ukpga/2007/12/contents> [Accessed: 15<sup>th</sup> September 2015].

Mental Welfare Commission for Scotland (2011) *Starved of Care: Investigation into the care and Treatment of "Mrs V."* Edinburgh.

Mercer, S.W. and Murphy, D.J. (2008) Validity and reliability of the CARE Measure in secondary care. *Clinical Governance: An International Journal*, 13 (4), pp. 269-283.

Mercer, S. Watt, G. Maxwell, M. and Heaney, D. (2004) The development and preliminary validation of the Consultation and Relational Empathy (CARE) Measure: an empathy-based consultation process measure. *Family Practice*, 21 (6), pp. 699-705.

- Mercer, S.W., McConnachie, A., Maxwell, M., Heaney, D. and Watt, G.C. (2005) Relevance and practical use of the Consultation and Relational Empathy (CARE) Measure in general practice. *Family Practice*, 22 (3), pp. 328-334.
- Mercer, S. Hatch, D. Murray, A. Murphy, D. and Eva, H. (2008) Capturing patients' views on communication with anaesthetists: the CARE Measure. *Clinical Governance: An International Journal*, 13, pp. 128-137.
- Meyer, B. and Bishop, D. (2007) Florence Nightingale: nineteenth century apostle of quality. *Journal of Management History*, 13 (3), pp. 240-254.
- Miller, G. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, 63, pp. 81-97.
- Moher, D., Liberati, A. and Tetzlaff, J. (2009) Preferred Reporting Items for Systematic Reviews and Meta-analysis: The PRISMA Statement. *Plos Med.* 6 (7), e10000097.
- Mokkink, L.B., Terwee, C.B., Knol, D.L., Alonso J., Patrick, D.L., Bouter, L.M. and de Vet, H. (2006) Protocol for the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Medical Research Methodology*, 6:2.
- Mokkink, L., Terwee, C., Stratford, P., Alonso, J., Patrick, D., Riphagen, I., Knol, D., Bouter, L. and de Vet, H. (2009) Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Quality of Life Research*, 18: pp. 313-333.
- Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Stratford, P., Knol, D., Bouter, L. and de Vet, H. (2010) The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, 19:pp. 539-549.
- Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M. and de Vet, H. (2012) *COSMIN checklist manual*. Amsterdam: Institute for Health and Care Research.
- Moret, L., Nguyen, J., Pillet, N., Falissard, B., Lombrail, P. and Gasquet, I. (2007) Improvement of psychometric properties of a scale measuring inpatient satisfaction with care: a better response rate and a reduction of the ceiling effect. *BMC Health Services Research*, 7, pp. 197.

Moumtzoglou, A., Dafogianni, C., Karra, V., Michailidou, D., Lazarou, P. and Bartsocas, C. (2000) Development and application of a questionnaire for assessing parent satisfaction with care. *International Journal for Quality in Health Care*, 12 (4), pp. 331-337.

Mullen, F. (2001) A founder of quality assessment encounters a troubled system firsthand. *Health Affairs*, 20(1), pp. 137-141. Available: <http://content.healthaffairs.org/content/20/1/137.short> [Accessed: 3<sup>rd</sup> August 2015].

*National Health Service (Scotland) Act 1978*. Available: <http://www.legislation.gov.uk/ukpga/1978/29/contents> [Accessed: 17<sup>th</sup> November 2015].

National Institute for Health and Clinical Excellence (2006) Reviewing Evidence. In: *Methods for Development of NICE Public Health Guidance*. Chapter 4, p. 25.

Nelson, J., Puntillo, K., Pronovost, P., Walker, A., McAdam, J., Ilaoa, D. and Penrod, J. (2010) In their own words: patients and families define high quality palliative care in the intensive care unit. *Critical Care Medicine*, 38 (3), pp. 808-818.

Nerney, M. Chin, M. Jin, L. Karrison, T. Walter, J. Mulliken, R. Miller, A. Hayley, D. and Friedman, P. (2001) Factors associated with older patients' satisfaction with care in an inner city emergency department. *Annals of Emergency Medicine*, 38 (2), pp. 140-145.

NHS Confederation (2013) *Changing care, improving quality: reframing the debate on reconfiguration*. NHS Confederation, London. Available: [http://www.aomrc.org.uk/doc\\_view/9692-changing-care-improving-quality](http://www.aomrc.org.uk/doc_view/9692-changing-care-improving-quality) [Accessed: 20<sup>th</sup> October 2015].

Noblit, G.W. and Hare, R.D. (1988) *Meta-ethnography: synthesizing qualitative studies*, Newbury Park, CA: Sage Publications.

Norman, G. (2010) Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, 15 (5), pp. 625-632.

Nursing and Midwifery Council (2015) *The Code: Professional Standards of Practice and Behaviour for Nurses and Midwives*. London: NMC.



- Nystrom, M., Dahlberg, K. and Carlsson, G. (2003). Non-caring encounters at an emergency care unit – a life-world hermeneutic analysis of an efficiency-driven organization. *International Journal of Nursing Studies*, 40, pp. 761-769.
- Oltedal, S., Garratt, A., Bjertnæs, Ø., Bjørnsdottir, M., Freil, M. and Sachs, M. (2007) The NORPEQ patient experiences questionnaire: data quality, internal consistency and validity following a Norwegian inpatient survey. *Scandinavian Journal of Public Health*, 35 (5), pp. 540-547.
- Onwuegbuzie, A.J., Johnson, R.B. and Collins, K.M. (2009) A call for mixed analysis: a philosophical framework for combining qualitative and quantitative. *International Journal of Multiple Research Approaches*, 3, pp. 114-139.
- O'Reilly, P. (2007) Involving service users in defining and evaluating the service quality of a disability service. *International Journal of Health Care Quality*, 20 (2-3), pp. 808-818.
- Orb, M.P. (2009) Phenomenology. In: S. Littlejohn and K. Foss (editors). *Encyclopaedia of Communication Theory*. Thousand Oaks, CA: Sage Publications.
- Ovretveit, J. (1992) *Health Service Quality: An Introduction to Quality Methods for Health Services*. Oxford: Blackwell Scientific.
- Oxman, A.D. and Guyatt, G.H. (1993) The science of reviewing research. In: K.S. Warren and F. Mosteller, eds. *Doing More Good than Harm: The evaluation of health care interventions*. New York: New York Academy of Sciences, pp. 156-163.
- Parasuraman, A., Zeithaml, V. and Berry, L. (1988) SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64 (1), pp. 12-40.
- Parker, J., Nester C.J., Long, A.F. and Barrie, J. (2003) The problem with measuring patient perceptions of outcome with existing outcome measures in foot and ankle surgery. *Foot Ankle International*, 24, pp. 56-60.
- Parry, G. (2014) A brief history of quality improvement. *Journal of Oncology Practice*, 10 (3), pp. 196-198.
- Pawson, R. and Tilley, N. (1997) *Realistic Evaluation*. London: Sage.

Perez-Carceles, M., Gironda, J., Osuna, E., Falcon, M. and Luna, A. (2010) Is the right to information fulfilled in an emergency department? Patients' perceptions of the care provided. *Journal of Evaluation in Clinical Practice*, 16, pp. 456-463.

Personal Communication (2012) *E-mail correspondence from Professor Kevin Rooney following discussion with Carol Haraden, PhD, Vice President, Institute for Healthcare Improvement.* 21<sup>st</sup> March 2012.

Personal Communication (2015) *Discussion with Diane Campbell, Associate Director for Patient Safety, NHS Tayside.* 5<sup>th</sup> November 2015.

Picker Institute Europe (2012) *Guidance Manual for the NHS Adult Inpatient Survey 2012.* Oxford: Picker Institute Europe.

Pirsig, R. (1974) *Zen and the Art of Motorcycle Maintenance.* New York: Bantam Books.

Pitrou, I., Lecourt, A., Bailly, L., Brousse, B., Dauchet, L. and Ladner, J. (2009) Waiting time and assessment of patient satisfaction in a large reference emergency department: prospective cohort study, France. *European Journal of Emergency Medicine*, 16 (4), pp. 177-182.

Plesk, P. and Greenhalgh, T. (2001) The challenge of complexity in health care. *British Medical Journal*, 323 (7313), pp. 625-628.

Polit, D.F. and Yang, F.M. (2016) Content validity and face validity. In: *Measurement and the Measurement of Change.* London: Lippincott Williams & Wilkins, pp. 168-181.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M. and Britten, N. (2006) Guidance on the Conduct of Narrative Synthesis in Systematic Reviews. *ESRC Methods Programme*, 15 (1), pp. 047-71.

Preston, C.C. and Coleman, A.M. (2000) Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104 (1), pp. 1-15.

Quality, Innovation, Productivity and Prevention (QIPP) Right Care. (2011) *The NHS Atlas of Variation in Healthcare: reducing unwarranted variation to increase value and improve quality.* Available: <http://www.rightcare.nhs.uk/index.php/nhs-atlas/> [Accessed: 11<sup>th</sup> April 2012].

- Raleigh, V. and Foot, C. (2010) *Getting the measure of quality. Opportunities and challenges*. London: The Kings Fund.
- Rao, K.D., Peters, D.H. and Bandeen-Roche, K., (2006) Towards patient-centered health services in India—a scale to measure patient perceptions of quality. *International Journal for Quality in Health Care*, 18 (6), pp. 414-421.
- Rathert, C., Huddleston, N. and Pak, Y. (2011) Acute care patients discuss the patient role in patient safety. *Health Care Management Review*, 36, pp. 134-144.
- Reason, J. (2000) Human error: models and management. *British Medical Journal*, 320 (7237), pp. 768-770.
- Reeves, S., Kuper, A. and Hodges, B.D. (2008) Qualitative research methodologies: ethnography. *British Medical Journal*, 337, a1020.
- Resar, R., Pronovost, P., Haraden, C., Simmonds, T., Rainey, T., and Nolan, T. (2005) Using a bundle approach to improve ventilator care processes and reduce ventilator-associated pneumonia. *The Joint Commission Journal on Quality and Patient Safety*, 31(5), pp. 243-248.
- Reynolds, W. and Scott, P.A. (2000) Do nurses and other professional helpers normally display much empathy? *Journal of Advanced Nursing Studies*, 31 (1), pp. 226- 234.
- Right Care (2011) *The NHS Atlas of Variation in Healthcare: reducing unwarranted variation to increase value and improve quality*. Available: <http://www.rightcare.nhs.uk/index.php/nhs-atlas/> [Accessed: 11th April 2015].
- Roach, K. (2006) Measurement of Health Outcomes: Reliability, Validity and Responsiveness. *Journal of Professions and Organization*, 18, pp. 8-12.
- Robert, G., Anderson, J., Burnett, S., Aase, K., Anderson-Gate, B., Bal, R., Calltorp, J., Nunes, F., Weggelaar, A., Vincent, C. and Fulop, N. (2011) A longitudinal, multi-level comparative study of quality and safety in European hospitals: the QUASER study protocol. *BMC Health Services Research*, 11, pp. 285.
- Rolstad, S., Adler, J. and Rydén, A. (2011) Response burden and questionnaire length: is shorter better? A review and meta-analysis. *Value in Health*, 14(8), pp. 1101-1108.

Royal College of Nursing (2009) *Comments on Public Service Reform (Scotland) Bill*, RCN, PSR19, Available:

<http://archive.scottish.parliament.uk/s3/committees/hs/inquiries/PSRBill/documents/PSR19RCNScotland.pdf> [accessed 11th February 2015]

Royal College of Nursing (2014) *Defining Nursing*. London: RCN.

Royal College of Nursing (2016) *If Scotland's current framework of NHS targets has had its day, what next?* Available: <https://www.rcn.org.uk/about-us/policy-briefings/sco-pol-sustainability-opinion-articles> [Accessed 3<sup>rd</sup> March 2016].

Rubin, H., Ware, J., Nelson, E. and Meterko, M. (1990) The Patient judgements of hospital quality (PJHQ) questionnaire. *Medical Care*, 28: S17-S18.

Russell, T. (2007) From my perspective: The fragmented and labyrinthine schematics that define our current health care system regularly frustrate surgeons' efforts to deliver, and patients' attempts to receive, optimal care. *Bulletin of the American College of Surgeons*, 92 (11), pp. 4-5

SAGE (2016) *Journal of Research in Nursing*.

Available: <http://jrn.sagepub.com/> [Accessed: 3<sup>rd</sup> March 2016].

Salisbury, C., Wallace M. and Montgomery, A.A. (2010) Patients' experience and satisfaction in primary care: secondary analysis using multilevel modelling. *British Medical Journal*, 341:c5004.

Sally, G. and Donaldson, L.J. (1998) Clinical governance and the drive for quality improvement in the new NHS in England. *British Medical Journal*, 317 (7150), pp. 61-65.

Scales, C. Jnr. and Schulman, K. (2014) Triggering Management for Quality Improvement. *Health Services Research*, 49 (5), pp. 1401-1406.

Scottish Government (2007a) *The Crerar Review: The report of the independent review of regulation, audit, inspection and complaints handling of public services in Scotland*. Edinburgh: Scottish Government.

Scottish Government (2010) *The Healthcare Quality Strategy for NHS Scotland*. Edinburgh: Scottish Government.

Scottish Government (2011a) *Achieving Sustainable Quality in Scotland's Healthcare: A 20:20 Vision*. Edinburgh: Scottish Government.

Scottish Government (2011b) *Information Services Division, Accident & Emergency Waiting Times*, Available: <http://isd.scot.nhs.uk/isd/4024.html> [Accessed 9<sup>th</sup> May 2012].

Scottish Government (2012) *Health Service Provision – Waiting Times*. Available: <http://www.gov.scot/Topics/Statistics/Browse/Health/TrendWaitingTimes> [Accessed: 21<sup>st</sup> October 2015].

Scottish Government (2014a) *Scottish Inpatient Patient Experience Survey, Volume 1: National Results*. Available: <http://www.gov.scot/Resource/0045/00458511.pdf> [Accessed 21<sup>st</sup> October 2015].

Scottish Government (2014b) *Consultation on Proposals to Introduce a Statutory Duty of Candour for Health and Social Care Services: October*. <http://www.scotland.gov.uk/Publications/2014/10/9897/0> [Accessed: 5<sup>th</sup> November 2014].

Scottish Government (2015) *Excellence in Care Deliverables: September*. Available: <http://www.gov.scot/Publications/2015/09/8281/3> [Accessed: 5<sup>th</sup> February 2015].

Scottish Government (2016) *Local Delivery Plan Guidance: January*. Available: [http://www.sehd.scot.nhs.uk/dl/DL\(2016\)01.pdf](http://www.sehd.scot.nhs.uk/dl/DL(2016)01.pdf) [Accessed: 4<sup>th</sup> February 2016].

Shaikh, B.T., Mobeen, N., Azam, S.I. and Rabbani, F., (2008) Using SERVQUAL for assessing and improving patient satisfaction at a rural health facility in Pakistan. *Eastern Mediterranean Health Journal*, 14 (2), pp. 447-456.

Sheng, Y. and Sheng, Z. (2012) Is Coefficient Alpha Robust to Non-Normal Data? *Frontiers in Psychology*, 3, p.34.

Sipkoff, M. (2004) The new consensus favouring IOM's definition of quality. *Managed Care*, 13 (6), pp. 18-21.

Sirota, R. (2006) Defining error in anatomic pathology. *Archives of Pathology and Laboratory Medicine*, 103 (5), pp. 604-606.

Sizmur, S. and Redding, D. (2012) *Core Domains for Measuring Inpatients' Experience of Care*, Oxford: Picker Institute Europe.

- Slavin, R. (1995) Best evidence synthesis: an intelligent alternative to meta-analysis. *Journal of Clinical Epidemiology*, 48 (1), pp. 9-18.
- Smith, R. (1998) All changed, changed utterly: British medicine will be transformed by the Bristol case. *British Medical Journal*, 316 (7149), pp. 1917-1918.
- Smith, P.C., Mossialos, E., and Papanicolas, I. (2008) Performance measurement for health system improvement: experiences, challenges and prospects. Copenhagen: World Health Organization Regional Office for Europe. Available: <https://stir.box.com/s/uwodn4p7hfqoa3h87opd26xjcx6j415s> [Accessed: 12th June 2016].
- Sofaer, S. and Firminger, K. (2005) Patient perceptions of the quality of health services. *Annual Review Public Health*, 26, pp. 513-559.
- Sofaer, S., Crofton, C., Goldstein, E., Hoy, E., and Crabb, J. (2005) What do consumers want to know about the quality of care in hospitals? *Health Services Research*, 40 (6p2), pp. 2018-2036.
- Stelfox, H.T., Palmisani, S., Scurlock, C., Orav, E.J., and Bates, D.W. (2006). The “*To Err is Human*” report and the patient safety literature. *Quality & Safety in Health Care*, 15 (3), pp. 174-178.
- Storm-Versloot, M.N., Vermeulen, H., van Lammeren, N., Luitse, J.S. and Goslings, J.C. (2014) Influence of the Manchester triage system on waiting time, treatment time, length of stay and patient satisfaction: a before and after study. *Emergency Medicine Journal*, 31 (1), pp. 13-18.
- Strauss, M.E. and Smith, G.T. (2009) Construct validity: advances in theory and methodology. *Annual Review of Clinical Psychology*, 5, pp. 1-25.
- Streiner, D.L. and Norman, G.R. (2003) *Health Measurement Scales*, 3<sup>rd</sup> ed., Oxford: Medical Publications.
- Streiner, D., Norman, G. and Cairney, J. (2015) *Health measurement scales: a practical guide to their development and use*, 5<sup>th</sup> edition. Oxford: Oxford University Press.
- Sullivan, G.M. and Artino, A.R. (2013) Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education*, 5 (4), pp. 541-542.

Swanson, K. (1999) What is known about caring in nursing science. In: A.S. Hinshaw, S. Fleetham and J. Shaver, (eds). *Handbook of Clinical Nursing Research*. Thousand Oaks, CA: Sage Publications, pp. 31-60.

Taguchi, G. (1992) *Taguchi on Robust Technology Development*. ASME Press.

Terwee, C.B., Bot, S.D., de Boer, M., van der Windt, D.A., Knol, D.L., Dekker, J., Bouter, L. and de Vet, H.C.W. (2007) Quality criteria were proposed for measurement properties of health status instruments. *Journal of Clinical Epidemiology*, 60, pp. 34-42.

Thisted, R. (2006) *The Cross-Sectional Study: Investigating Prevalence and Association*. University of Chicago. Available: <http://health.bsd.uchicago.edu/thisted/epor/Lectures/061011-EPOR-CrossSection.pdf> [Accessed: 10<sup>th</sup> September 2015].

Toma, G., Triner, W. and McNutt, L. (2009) Patient satisfaction as a function of emergency department revisit expectations. *Annals of Emergency Medicine*, 54 (3), pp360-367e.

Trivers, R. (1971) The evolution of reciprocal altruism. *The Quarterly Review of Biology*, 46 (1), pp. 35-57.

Van der Vleuten, C. (1996) The assessment of professional competence: developments, research and practical implications. *Advances in Health Science Education*, 1m, pp. 41-67.

Watson, J. (2009) *Assessing and Measuring Caring in Nursing and Health Sciences*. New York: Springer.

Webster, T.R., Mantopoulos, J., Jackson, E., Cole-Lewis, H., Kidane, L., Kebede, A., Abebe, Y., Lawson, R. and Bradley, E.H. (2011) A brief questionnaire for assessing patient healthcare experiences in low-income settings. *International Journal of Quality Health Care*, 23, pp. 258-268.

Whittemore, R. and Knafl, K. (2005) The integrative review: an updated methodology. *Journal of Advanced Nursing*, 52 (5), pp. 546-533.

Wiig, A., Aase, K., von Plessen, Burnett, S., Nunes, F., Weggelaar, A., Anderson-Gare, B., Calltorp, J. and Fulop, N. (2014a) Talking about quality: exploring how 'quality' is

conceptualized in European hospitals and healthcare systems. *BMC Health Services Research*, 14, pp. 478.

Wiig, A., Robert, G., Anderson, J., Pietikainen, E., Reiman, T., Macchi, L. and Aase, K. (2014b) Applying different quality and safety models in healthcare improvement work: Boundary objects and systems thinking. *Reliability Engineering and System Safety*, 125, pp. 134-144.

Wilde B, Starrin B, Larsson G and Larsson M. (1993) Quality of care from a patient perspective: a grounded theory study. *Scandinavian Journal of Caring Sciences*, 7, pp. 113-120.

Wilde, B., Larsson, G., Larsson, M. and Starrin, B. (1994) Quality of care: development of a patient-centred questionnaire based on a grounded theory model. *Scandinavian Journal of Caring Sciences*, 8 (1), pp. 39-48.

Williams, B. (1994) Patient satisfaction: a valid concept? *Social Science and Medicine*, 38 (4), pp. 509-516.

Williams, B., Coyle, J. and Healy, D. (1998) The meaning of patient satisfaction: An explanation of high reported levels. *Social Science and Medicine*, 47 (9), pp. 1351-1359.

Wiman, E. and Wikblad, K. (2004) Caring and uncaring encounters in nursing in the emergency department. *Journal of Clinical Nursing*, 13, pp. 422-429.

Wong, E., Coulter, A., Cheung, A.W., Yam, C.H., Yeoh, E. and Griffiths, S.M. (2013) Validation of Inpatient Experience Questionnaire. *International journal of Quality Health Care*, 25, pp. 443-451.

Woodcock, T., Poots, A.J. and Bell, D. (2013). The impact of changing the 4<sup>th</sup> emergency access standard on patient waiting times in emergency departments in England. *Emergency Medicine Journal*, 30 (3), pp. e22-e22.

World Health Organisation (2006) *Quality of Care: a Process for Making Strategic Choices in Health Systems*. Geneva: WHO.

World Medical Association (1964) *WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects*, Principle 30. Available: <http://www.wma.net/en/30publications/10policies/b3/> [Accessed: 15<sup>th</sup> September 2015].



# Appendices

## Appendix 1: Local Delivery Plan (LDP) Standards

<p><b>NHS LDP Standards</b></p> <p><b>People diagnosed and treated in 1<sup>st</sup> stage of breast, colorectal and lung cancer (25% increase)</b> 31 days from decision to treat (95%) 62 days from urgent referral with suspicion of cancer (95%) <i>Early diagnosis and treatment improves outcomes.</i></p> <p><b>People newly diagnosed with dementia will have a minimum of 1 years post-diagnostic support</b> <i>Enable people to understand and adjust to a diagnosis, connect better and plan for future care</i></p> <p><b>12 weeks Treatment Time Guarantee (TTG 100%)</b> <b>18 weeks Referral to Treatment (RTT 90%)</b> <b>12 weeks for first outpatient appointment (95% with stretch 100%)</b> <i>Shorter waits can lead to earlier diagnosis and better outcomes for many patients as well as reducing unnecessary worry and uncertainty for patients and their relatives.</i></p> <p><b>At least 80% of pregnant women in each SIMD quintile will have booked for antenatal care by the 12th week of gestation</b> <i>Antenatal access supports improvements in breast feeding rates and other important health behaviours.</i></p> <p><b>Eligible patients commence IVF treatment within 12 months (90%)</b> <i>Shorter waiting times across Scotland will lead to improved outcomes for patients.</i></p> <p><b>18 weeks referral to treatment for specialist Child and Adolescent Mental Health Services (90%)</b> <i>Early action is more likely to result in full recovery and improve wider social development outcomes.</i></p> <p><b>18 weeks referral to treatment for Psychological Therapies (90%)</b> <i>Timely access to healthcare is a key measure of quality and that applies equally to mental health services.</i></p> <p><b>Clostridium difficile infections per 1000 occupied bed days (0.32)</b> <b>SAB infections per 1000 acute occupied bed days (0.24)</b> <i>NHS Boards area expected to improve SAB infection rates during 2015/16. Research is underway to develop a new SAB standard for inclusion in LDP for 2016/17.</i></p> <p><b>Clients will wait no longer than 3 weeks from referral received to appropriate drug or alcohol treatment that supports their recovery (90%)</b> <i>Services for people are recovery focused, good quality and can be accessed when and where they are needed.</i></p> <p><b>Sustain and embed alcohol brief interventions in 3 priority settings (primary care, A&amp;E, antenatal) and broaden delivery in wider settings</b> <b>Sustain and embed successful smoking quits, at 12 weeks post quit, in the 40% SIMD areas</b> <i>Enabling people at risk of health inequalities to make better choices and positive steps toward better health.</i></p> <p><b>48 hour access or advance booking to an appropriate member of the GP team (90%)</b> <i>Often a patient's first contact with the NHS is through their GP practice. It is vital, therefore, that every member of the public has fast and convenient access to their local primary medical services to ensure better outcomes and experiences for patients.</i></p> <p><b>Sickness absence (4%)</b> <i>A refreshed Promoting Attendance Partnership Information Network Policy will be published in 2015.</i></p> <p><b>4 hours from arrival to admission, discharge or transfer for A&amp;E treatment (95% with stretch 98%)</b> <i>High correlation between emergency departments with 4 hour wait performance between 95 and 98% and elimination of long waits in A&amp;E which result in poorer outcomes for patients</i></p> <p><b>Operate within agreed revenue resource limit; capital resource limit; and meet cash requirement</b> <i>Sound financial planning and management are fundamental to effective delivery of services.</i></p>
---

## Appendix 2: Patient Information Leaflet

# Patient Information Leaflet



UNIVERSITY OF  
STIRLING

SCHOOL OF  
NURSING, MIDWIFERY  
AND HEALTH

### **A Study Comparing Waiting Time & Perceived Empathy with Satisfaction of Care**

The Accident and Emergency (A&E) Department are always looking for ways to improve the service they deliver to you, the patient. This study invites you to help us decide the best way to measure how satisfied you are with the service. Before you decide whether or not to participate it is important for you to understand why the research is being done and what it will involve.

#### **What is the purpose of the study?**

Currently, the quality of care within A&E is measured by the speed in which you have been seen and treated (waiting time). We think that another important aspect of how satisfied you are with your care may be related to how you feel the staff understand and respond to your needs. This study aims to determine whether or not you think that staff understanding is a good measure of how satisfied you are with your care.

All Accident and Emergency (A&E) Departments in the UK must record the duration of your visit in the Department. This information is used to determine the quality of care you receive. This study seeks to identify if how you feel you were treated in A&E is a better measure of the quality of care you have received in the Department.

#### **Why have I been chosen?**

The study aims to ask at least 70 patients (over 18 years) to complete a questionnaire following primary treatment, and before discharge from A&E. Anyone willing to participate and consent is asked to complete the questionnaire.

#### **Do I have to take part?**

No. It is up to you to decide whether to take part. If you do decide to take part, you will be asked to sign a consent form. Even if you decide to take part you are still free to withdraw at any time and without giving a reason. Your decision will not affect the standard of care you receive. You will be asked to complete the questionnaire after you have received the necessary care you require.

#### **What will happen to me if I take part?**

You will be asked to complete a questionnaire at the end of your A&E visit – either before you go home or before you go to another area for further care. The questionnaire takes about five minutes to complete. Your answers will only be seen by the person conducting the study.

Highland Campus:  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JH  
Tel: +44 (0) 1463 255655  
Fax: +44 (0) 1463 255654

Stirling Campus:  
Stirling  
FK9 4LA  
Tel: +44 (0) 1786 466340  
Fax: +44 (0) 1786 466333

Western Isles Campus:  
Western Isles Hospital  
MacAuley Road  
Stornoway Isle of Lewis HS1 2AF  
Tel: +44 (0) 1851 708243  
Fax: +44 (0) 1851 708070

Version 4: 27 October 2011

## Appendix 3: University of Stirling Ethics Approval and Submission

### Tracking Projects



**UNIVERSITY OF  
STIRLING**

DEPARTMENT OF  
NURSING AND MIDWIFERY

<b>Project Title</b>	A & E Study: Is there a correlation between and waiting time and empathy?
<b>Project Number/Cost Centre</b>	
<b>Project Description</b>	<p>To determine the relationship between the CARE measure (not currently measured) and waiting times (current Health Improvement; Efficiency; Access and Treatment, commonly known as HEAT target) within a local Accident and Emergency Department.</p> <p>Patients included in the study (those not meeting the exclusion criteria) will be given patient information by A&amp;E staff. Those willing to participate will be given a consent form and CARE questionnaire. Patients will be asked to complete the questionnaire following their A&amp;E consultation. Completed questionnaires will be collected by the PI. CHI numbers will be used to match the questionnaire to the patients wait time in A&amp;E (data already collected). Participant will be given a unique identifier and CHI will be destroyed to ensure patient anonymity. Data will be entered into SPSP for analysis.</p>
<b>Principal Investigator</b>	Michelle Beattie
<b>Principal's Department and University</b>	School of Nursing, Midwifery and Health

<b>Will Ethical Review be Sought?</b>	Yes	No <input type="checkbox"/>
<b>Reasons if No</b>		
<b>Authorised by</b>		

<b>All Staff Employed in Project</b>	Michelle Beattie			
<b>Other Investigators (University of Stirling)</b>	None			
<b>Other Investigators (External)</b>	None			
<b>Financial Year Funds Awarded</b>	<b>Start Date</b>		<b>End Date</b>	

<b>Funding Body (In Full)</b>	None		
<b>Total Amount Requested</b>	£	<b>Total Amount Awarded</b>	£
<b>Eligible for inclusion in RAE?</b>	Yes		

<b>Extension Date</b>		<b>Additional Funds</b>	£
-----------------------	--	-------------------------	---

<b>Programme</b>	PhD
------------------	-----

<b>Links to Other Programmes</b>	
----------------------------------	--

<b>Research Group or Centre</b> (if applicable)	Educational Research and Practice Development
--	---

<b>Contact Person</b> (Research Assistant)	Michelle Beattie
---	------------------

<b>Key Words</b>	Quality, Research, Accident and Emergency
------------------	---

	<b>submitted</b>	<b>accepted</b>	<b>unacceptable</b>	<b>Date</b>
<b>Interim Report</b>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
<b>Final Report</b>	<input type="checkbox"/>			

**Departmental Research Ethics  
Committee**



**UNIVERSITY OF  
STIRLING**

SCHOOL OF  
NURSING, MIDWIFERY  
AND HEALTH

**Application Form**

**This form must be submitted as part of your application.**

NAME	Michelle Beattie
DESIGNATION	Lecturer
INSTITUTION	School of Nursing, Midwifery and Health
ADDRESS	Highland Campus, Centre for Health Science, Old Perth Road, Inverness, IV3 3JH
EMAIL ADDRESS	<a href="mailto:Michelle.beattie@stir.ac.uk">Michelle.beattie@stir.ac.uk</a>
TELEPHONE NUMBER	01463 255622
STUDENT NO. <i>(where relevant)</i>	1929599
SUPERVISOR	Professor William Lauder
FUNDING BODY <i>(where relevant)</i>	None
PROJECT TITLE	A & E Study: Is there a correlation between and waiting time and relational empathy?
PROJECT PROPOSED START DATE	August 2011
PROJECT PROPOSED COMPLETION DATE	December 2011
APPLICANT'S SIGNATURE	<input type="text"/>
<b>SUPERVISOR'S SIGNATURE</b> <i>(where relevant)</i>	<hr/> <input type="text"/> <hr/>
<b>ADDITIONAL COMMENTS FROM SUPERVISOR INCLUDED</b>	<input type="checkbox"/>
DATE	13 June 2011 <input type="text"/>



Application Checklist Form

---

**This form must be submitted as part of your application.**

**1 copy** of the DREC application form and an electronic version with all supporting documentation are required.

Applications will **only** be processed on receipt of the appropriate documents.

Please  
mark

- This form: The completed DREC application form
- Study proposal  
*(that follows the DREC guidelines for the content of study proposal for ethical approval)*
- Patient information sheet
- Written consent forms
- Interview schedules/questionnaires
- Research tracking form
- Others (please specify)

Please also ensure that patient information sheets, GP letters and other documents are on headed paper and have version numbers and version dates recorded on them.

Please return this checklist together with your application form to the address below:

Dr Fiona Harris (Acting Chair)  
c/o Sarahjane Gilvear  
R.G Bomont Building  
School of Nursing, Midwifery and Health  
University of Stirling  
STIRLING FK9 4LA

Telephone: 01786 466404  
Email: [nm.research@stir.ac.uk](mailto:nm.research@stir.ac.uk)

### **Study 3 Proposal**

Is there a relationship between the CARE measure and waiting times within a local Accident and Emergency Department? – A correlation study

#### **Rationale**

Measurement of processes and systems have become commonplace in health care in an attempt to improve quality. The measures are often derived from the Institute of Medicines' (IOM) dimensions of quality – safety, timeliness, effectiveness, efficiency, equity and patient-centred (IOM, 2001). The measures provide an indicator of the level of quality in a particular area or speciality, hence the name quality indicator.

Much of quality improvement thinking aligns well with systems theory – acknowledging the significant influence of the environment, or complex system, in which health care practitioners operate and therefore targeting system change, rather than the agents of the system (Deming 1986). Systems thinking would suggest that where an organisation performs well in compulsory quality measures; they would likely perform well in other unmeasured dimensions of quality. The assumptions are that measurements of quality within the same organisation will be influenced by similar characteristics of the organisation or system. This study seeks to test this theory by conducting a correlation study to determine the relationship between the CARE measure (not currently measured) and waiting times (current HEAT target) within a local Accident and Emergency (A&E) Department.

Within the UK, waiting times have been a significant element of health care policy aiming to improve the quality of health care services. The Scottish Government determined the benchmark that all patients must be seen, treated and moved on to

Version 1: 7June11



an appropriate area or care, or discharged within 4 hours of arrival to the A&E Department (Scottish Government, 2011)

### **Aim**

To determine the relationship between the CARE measure (not currently measured) and waiting times (current Health Improvement; Efficiency; Access and Treatment, commonly known as HEAT target) within a local Accident and Emergency Department.

### **Study Design**

A correlation study will be used to examine what relationship, if any, exists between the scores of the CARE measure and waiting time in A&E. A positive correlation would indicate that as the number of minutes waited decreases, the relational empathy (CARE) measure score would increase.

### **Population Sample**

Questionnaires (inclusive of CARE measure and demographic details), a study explanation sheet and consent form would be distributed to all adult patients attending the local A&E Department during set time frames of data collection. Patients requiring immediate resuscitation, those unable to give informed consent due to incapacity or altered levels of consciousness will be excluded from the study (this would include those whose capacity has been temporarily affected by opioid analgesia). Patients returning for a second planned visit will also be excluded as they are unrepresentative of the population attending A&E, as their wait time and CARE score may be influenced by their primary visit.

Version 1: 7June11

Patients eligible to participate in the study would be given a study explanation sheet by A&E staff to determine whether or not they wished to participate in the study. Consent forms would be explained and administered by the principal investigator. Participants will be asked to complete the questionnaire at the end of their A&E consultation. The questionnaire would contain the patients CHI (Community Health Index) number. The length of time the patient spent in A&E would be obtained from the database in the A&E department using the patients CHI number. Each participant would be allotted a study number and their CHI number would be kept separately from their responses to ensure anonymity.

### Data Collection

The study aims to collect a minimum of 50 questionnaires to enable the results to be of statistical significance. There are recognised time periods when A&E Departments are busier and quieter. To ensure the sample of questionnaires are representative of these differences data collection will be within specific time frames (see table 1).

A&E Departments are busiest on Mondays between 10am and lunchtime. There is a similar pattern across the UK. Attendances peak between midday and 4pm.

To ensure sufficient numbers are achieved, particularly over quieter episodes of time, the data collection would likely need to be repeated in over a second week.

Day	Time Frame	Number of questionnaire to be completed
Monday	08.00 – 12.00 (quiet time)	12
Monday	12.00 – 16.00 (busy time)	12
Wednesday	08.00 – 12.00 (quiet time)	12
Wednesday	12.00 – 16.00 (busy time)	12
Saturday	08.00 – 12.00 (quiet time)	12
Saturday	12.00 – 16.00 (busy time)	12

Relational empathy will be determined using the CARE measure questionnaire (Mercer et al 2005). The CARE measure tool has been validated for use in primary

Version 1: 7June11

care (Mercer et al 2004) and some studies have utilised this within a secondary care setting (Mercer & Murphy 2008; Mercer et al 2008). The tool has not been previously used within A&E.

### **Analysis**

Data would be coded and entered into SPSS for analysis. The values would be standardised and Pearson's correlation coefficient test would be used to determine whether a positive, negative, or indeed, no relationship existed between time spent in A&E and CARE measure scores.

### **Potential risks**

Patients may perceive a risk that their care in A&E would be influenced depending on their responses to the CARE measure, or, whether or not they choose to participate. Staff may also have reservations about the study as the tool may highlight poor levels of perceived empathy by patients. There is a risk that insufficient patients are recruited during data collection.

### **Safeguarding against risk**

Patients will be reassured that their questionnaire responses will only be seen by the principal investigator and that clinical staff providing their care will not see their responses. Clinical Staff will ask whether or not the patient would like to participate and be reminded that their decision to participate will not affect their care in any way. This information will also be reinforced on the patient advice leaflet, which aims to help the patient to decide whether or not to participate.

Staff concerns will be reduced or eliminated by ensuring the purpose of the study is known to staff, meetings with the Charge Nurse and emphasising that individual staff

Version 1: 7June11

members will not be 'scored' in relation to the CARE measure. The A&E Department will receive a CARE measure score, not individuals working within the area.

#### **Informed consent**

The information sheet will provide the patients with enough information to decide whether or not they want to participate. Staff of the PI (who will be on site during all phases of data collection) would be able to answer any further questions the patient may have about the study. Patients willing to participate must sign a consent form stating that they understand the study and are willing to proceed. Those volunteering to participate would be given a questionnaire to complete at the end of their consultation (prior to discharge or before moving to an appropriate area of care).

#### **Researcher's role**

The PI is a nurse, but not known to the area, therefore the PI will adopt the role of researchers within the study. No questions are required to be directly asked by the PI, therefore researcher bias is limited. The patient will only meet the PI to return the questionnaire or if they have any questions.

#### **Sensitive topics**

No other areas of sensitivity are predicted.

#### **Confidentiality, anonymity and data protection**

The CARE questionnaire will include the patients CHI number to enable matching of their waiting time with the Departments electronic database. Once the wait time had been documented on the questionnaire the CHI number will be removed from the

Version 1: 7June11

---

questionnaire. Each participant will then be given a unique identifier as the information is inputted to SPSS. The database will be password protected and only accessible to the PI and research supervisor. The patient would not be identifiable from the information held within SPSS. Data will be stored in accordance with the Data Protection Act

#### **Other research ethics committees**

The application will also be submitted to the North of Scotland Research Ethics Committee.

#### **Main Ethical Issues**

The research will include the use of patient identifiable information. The patient identifier (CHI) would be destroyed once the waiting time was recorded to ensure anonymity following data collection. Completed questionnaires would be secured in a locked filing cabinet. Subjects would be given a unique identifier.

#### **Outputs**

A paper would be written for publication and sent to an appropriate target journal. A report would be prepared for NHS Highland Board.

#### **Costs/Resources**

PI time as determined by PhD professional development.

Version 1: 7June11

## **Timeline**

### **June 11**

Check study proposal with supervisor and A&E Department.

### **June 11**

Submit to university ethics

### **July 11**

Submit to IRAS

### **May – June 11**

Literature Review

Target Journal Identified

Detailed Plan of Study

### **July 11**

Draft background and methodology written

### **August 11**

Pilot data collection

### **August and September 11**

Data collection

### **September – October 11**

Data input and analysis

### **November – December 11**

Write paper

Submit to target journal.

Version 1: 7June11

## References

Deming E (1986) *Out of a Crisis*, Massachusetts Institute of Technology, Blackwell Publisher.

Institute of Medicine (2001) *Crossing the Quality Chasm: A New Health System for 21<sup>st</sup> Century*, Washington DC, National Academy Press. Accessed 10<sup>th</sup> November 2010.

[http://books.nap.edu/openbook.php?record\\_id=10027&page=R1](http://books.nap.edu/openbook.php?record_id=10027&page=R1)

Mercer S, Watt G, Maxwell M, Heaney D and (2004) The development and preliminary validation of the Consultation and Relational Empathy (CARE) Measure: an empathy-based consultation process measure, *Family Practice*, 21 (6), 699-705.

Mercer S, McConnachie A, Maxwell M, Heaney D and Watt G (2005) Relevance and performance of the Consultation and Relational Empathy (CARE) Measure in general practice. *Family Practice*, 22 (3), 328-334.

Mercer S and Murphy D. (2008) Validity and reliability of the CARE Measure in secondary care. *Clinical Governance: An International Journal*, 13; 261-283.

Mercer S, Hatch D, Murray A, Murphy D and Eva H (2008) Capturing patients' views on communication with anaesthetists: the CARE Measure. *Clinical Governance: An International Journal*, 13; 128-137.

Scottish Government (2011) *Accident and Emergency Waiting Times*, ISD Scotland, 31<sup>st</sup> May 2011 available online, accessed 13 June 2011.

<http://www.isdscotland.org/Health-Topics/Waiting-Times/Emergency-Departments/>

Version 1: 7June11

RJ/SG

25 July 2011

Michelle Beattie  
Lecturer  
School of Nursing, Midwifery and Health  
Highland Campus  
Centre for Health Science  
Old Perth Road  
Inverness  
IV3 3JH



UNIVERSITY OF  
STIRLING

SCHOOL OF  
NURSING, MIDWIFERY  
AND HEALTH

Email: nursingmidwifery@stir.ac.uk  
Web: www.nm.stir.ac.uk

Ruth Jepson  
Acting Chair (on behalf of Fiona Harris)  
School Research Ethics Committee

School of Nursing, Midwifery and Health  
University of Stirling  
Stirling FK9 4LA

Tel: +44 (0) 1786 466402  
Fax: +44 (0) 1786 466333  
Email: ruth.jepson@stir.ac.uk

Dear Michelle

### **A & E Study: Is there a correlation between and waiting time and relational empathy?**

The Committee discussed your application for research ethics approval at the meeting on 13 July and after careful consideration we decided to pass this proposal subject to chair's action on the points below.

#### **Rationale for the study hypothesis**

We would like to see further rationale as why waiting times may impact on empathy. We would also like you to provide a clear definition (for us and the participants) as to what empathy means in this context. We felt that defining this for participants might address the risk that not all patients would be familiar with this term.

#### **Data measures**

As you point out, the CARE measure was developed for primary care consultations which are almost exclusively one-to-one consultations. In the A & E setting, the patient may be seen by a number of health professionals or may not actually be seen by a doctor at all. We would like more clarification as to whether a) the patient will fill out a separate survey for each health professional, or whether b) they would choose a particular health professional. Both of these approaches have limitations which need to be discussed. For example, the CARE measure has been developed and validated for doctors but not other health professionals. How will you address this if the patient is only seen by a nurse? The CARE questionnaire as it stands asks the patient to think about the 'doctor'.

We did not think it was appropriate or necessary to access the patient records, and ask you to remove this from the proposal, the consent sheet and the information sheet.

Highland Campus:  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JH

Tel: +44 (0) 1463 255655  
Fax: +44 (0) 1463 255654

Stirling Campus:  
Stirling  
FK9 4LA

Tel: +44 (0) 1786 466340  
Fax: +44 (0) 1786 466333

Western Isles Campus:  
Western Isles Hospital  
MacAulay Road  
Stornoway Isle of Lewis HS1 2AF

Tel: +44 (0) 1851 708243  
Fax: +44 (0) 1851 706070

The University of Stirling is recognised as a Scottish Charity with number SC 011159



### **Sample size and recruitment issues**

You need to further justify your sampling strategy and sample size. Also you need to consider the potential for bias that may occur between, for example patients recruited during busy times and quiet times, and how you might address these.

It would be important to add in further detail about how you will manage potentially vulnerable patients (other than those on opioid medication), for example those who have experienced trauma related to their use of A&E.

### **Analysis**

The section on analysis may need to be re-written for the NHS Ethics Committee. We suggest that you look at the terminology being used for example rather than 'determining the relationship' framing this as rejecting/not rejecting the null hypothesis etc.

### **Patient Information Leaflet**

As it stands, the language and terminology needs to be made simpler and clearer for the participants. For example, under point 3 you introduce the CARE describing detail at a methods level. We also suggest that point 9 should be removed, but that details of an 'external contact' be added, who participants can contact if they have a query or concern about the research.

As noted above, please take out the reference to accessing patients' medical records.

You have also left in the descriptions of what is needed for each section and these need to be removed.

### **Other issues**

The Committee would like further information on:

1. Who are the staff of the PI and what role will they have? For example are they researchers or health professionals or both?
2. Will the PI be on site during data collection?

Please respond to each of the points in an email to me. You cannot submit your proposal to IRAS until we have signed off your application, but you do not have to wait until the next Ethics Committee meeting for this.

We wish you well with your research.

Yours sincerely



Ruth Jepson  
Chair (Acting) on behalf of Fiona Harris  
School of Nursing, Midwifery and Health Research Ethics Committee

3<sup>rd</sup> August 2011SCHOOL OF  
NURSING, MIDWIFERY  
AND HEALTHEmail: nursingmidwifery@stir.ac.uk  
Web: www.stir.ac.uk/nmhealthMichelle Beattie  
Lecturer  
School of Nursing, Midwifery and Health  
University of Stirling  
Highland Campus  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JHTel: +44 (0) 1463 255622  
Fax: +44 (0) 1463 255654  
Email: michelle.beattie@stir.ac.uk

Dear Ruth

**A Study Comparing Waiting Time and Perceived Empathy with Overall Satisfaction**

Many thanks for your response on behalf of the Ethics Committee on the above study. The queries posed have helped me formulate a clearer study proposal. I have addressed the queries and questions below.

**Rationale for the study hypothesis**

The study purpose is to compare whether patient's scores of relational empathy capture their perceptions of health care quality more accurately than waiting time in Accident and Emergency (A&E). Although waiting times have improved across the UK there are concerns that time is an inaccurate reflection of the patients' health care experience within (Jolly and Clancy, 2009). There is a risk that as we focus on measurable outcomes, such as time, important aspects of caring, namely empathy, become marginalised. For the purpose of the study empathy will be defined as an ability to understand, communicate and act on the patient's situation, perspective and feelings (Mercer and Reynolds, 2002).

**Data measures**

Although the CARE measure was developed and validated for primary care consultations with doctors, latterly the tool has been used in several studies on nurses and other health care professionals, as well as in the hospital setting (see attached list of research conducted using the CARE measure). Participant will be asked to complete the questionnaire on the health care professional who has conducted the majority of their consultation; this will be either a doctor or a nurse. The wording on the CARE measure tool will be amended to include doctor/nurse. The patient information leaflet will be amended to ensure clarity (enclosed).

**Sample size and recruitment issues**

Although Mercer et al (2004) recommend the use of 50 questionnaires; the statistical test requires 64 questionnaires to establish an adequate sample size. Using a t-test, this number will provide a sample sufficient to give a 0.5% effect size with a power of 80% (Machin et al 2008). A 0.5% effect size has been estimated to provide a moderate effect size (Cohen, 1988). The study therefore aims to collect a minimum of 70 self completed questionnaires.

Equal numbers of questionnaires will be administered during specific time frames to enable an accurate distribution at recognised busy and quiet times. Data collection will likely be required for two weeks to enable an equal distribution of questionnaires between busy and quiet times. Recognised vulnerable

Highland Campus:  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JH  
Tel: +44 (0) 1463 255655  
Fax: +44 (0) 1463 255654Stirling Campus:  
Stirling  
FK9 4LA  
Tel: +44 (0) 1796 466340  
Fax: +44 (0) 1796 466333Western Isles Campus:  
Western Isles Hospital  
MacAulay Road  
Stornoway Isle of Lewis H81 2AF  
Tel: +44 (0) 1851 708243  
Fax: +44 (0) 1851 706070

adults will not be approached to participate and, where, vulnerability is unknown, the patient has the option to not participate.

#### **Analysis**

The study hypothesises that as patients' empathy scores increase, their overall rating of health care satisfaction will increase. The study also hypothesises that patients' waiting time will not influence their rating of health care satisfaction. These hypothesis will be answered by comparing mean scores of patients overall ratings of satisfaction with their care measure score using the t-test in SPSS.

#### **Patient information leaflet**

The language and terminology used in the patient information leaflet have been revised to ensure appropriateness for the lay public (see attached). The reference to accessing patients' medical notes was taken following attendance at a North of Scotland Research Ethics Committee (NoSREC) drop in session, where the ethical advisor recommended requesting access to the patient notes in case the study was selected for audit purposes. This would ensure NoSREC could access the medical notes for audit. On reflection, it would seem that NoSREC would obtain little information on this specific study from the medical notes and therefore this has been removed.

#### **Other issues**

The Principal Investigator will be an academic, who will be on site during data collection. The other member of staff specifically involved will be a Clinical Educator, who is based in A&E and works for NHS Highland.

I hope this information helps to clarify the points raised. Please get in touch if you require any further information.

Yours sincerely

Michelle Beattie  
Lecturer/PhD Student

Enc 3

## **The CARE Measure – summary of research and current use (2009)**

### **Background**

The Consultation and Relational Empathy (CARE) Measure is a consultation process measure developed by Professor Stewart Mercer and colleagues in the Departments of General Practice at Glasgow University and Edinburgh University. It is based on a broad definition of empathy in context of a therapeutic relationship within the consultation (clinical encounter). The wording reflects a desire to produce a holistic, patient-centred measure that is meaningful to patients irrespective of their socioeconomic position. The measure has been extensively validated (see below) and is widely used in the UK and internationally. A summary of the research is shown below, followed by a summary of the ways in which the CARE Measure is currently being used in practice.

### **Published work on the CARE Measure (2002-2009)**

#### *Theoretical underpinning*

- Mercer SW and Reynolds W. Empathy and quality of care. BJGP 2002, 52 (Supplement), S9-S12
- Neumann M, Bensing J, Mercer S, Ernstmann N, Ommen O, Pfaff H. Analyzing the “nature” and “specific effectiveness” of clinical empathy: A theoretical overview and contribution towards a theory-based research agenda. Patient Education and Counseling 2009, 74:339-46

#### *Validity and reliability of the CARE Measure*

- Mercer SW, Watt, GCM, Maxwell M, and Heaney DH. The development and preliminary validation of the Consultation and Relational Empathy (CARE) Measure: an empathy-based consultation process measure. Family Practice 2004, 21 (6), 699-705
  - Mercer SW, McConnachie A, Maxwell M, Heaney DH, and Watt GCM. Relevance and performance of the Consultation and Relational Empathy (CARE) Measure in general practice. Family Practice 2005, 22 (3), 328-334
  - Mercer SW and Howie JGR. CQI-2, a new measure of holistic, interpersonal care in primary care consultations. BJGP 2006, 56 (525), 262-268
  - Mercer SW, Murphy DJ. Validity and reliability of the CARE Measure in secondary care. Clinical Governance: An International Journal 2008, 13; 261-283
-

- Mercer SW, Hatch DJ, Murray A, Murphy DJ, Eva HW. Capturing patients' views on communication with anaesthetists: the CARE Measure. *Clinical Governance: an international journal* 2008, 13 (2) : 128-137
- Murphy DJ, Bruce DA, Mercer SW, Eva KW . The reliability of workplace-based assessment in postgraduate medical education and training: a national evaluation in general practice in the United Kingdom. *Advances in Health Sciences Education* 2009, 13: 219-232
- Fung C, Mercer SW. A qualitative study of patients' views on quality of primary care consultations in Hong Kong and comparison with the UK CARE Measure. *BMC Family Practice* 2009, 10:10
- Fung C, Hua A, Tam L, Mercer SW. Reliability and validity of the Chinese version of the CARE Measure in a primary care setting in Hong Kong. *Family Practice* 2009 26(5):398-406

#### *Effect of CARE on outcomes*

- McPherson H, Mercer SW, Scullion T, Thomas KJ. Empathy, enablement, and outcome: an exploratory study of acupuncture patients' perceptions. *J Alt. Comp. Med.* 2003, 9(6), 869-876.
- Mercer SW. Practitioner empathy, patient enablement, and health outcomes of patients attending the Glasgow Homoeopathic Hospital : a retrospective and prospective comparison. *Wien. Med. Wochenschr.* 2005, 155, 21-22, 498-501
- Bikker AP, Mercer SW, Reilly D. A pilot prospective study on the consultation and relational empathy, patient enablement, and health changes over 12 months, in patients going to the Glasgow Homoeopathic Hospital. *J Alt. Comp. Med.* 2005, 11 (4), 591-600
- Price S, Mercer SW, McPherson H. Practitioner empathy, patient enablement, and health outcomes: a prospective study of acupuncture patients. *Patient Education and Counseling* 2006, 63 (1-2), 239-245
- Neumann M, Wirtz W, Bollschweiler E, Mercer SW, Warm M, Wolf J, Pfaff H. Determinants and patient-reported long-term outcomes of physician empathy in oncology: A structural equation modelling approach. *Patient Education and Counseling* 2007, 69: 63-75
- Mercer SW, Neumann M, Wirtz W, Fitzpatrick B, Vojt G. Effect of General Practitioner empathy on patient enablement, and patient-reported outcomes in primary care in an area of high socio-economic deprivation in Scotland - A pilot

prospective study using structural equation modelling. *Patient Education and Counseling* 2008, 73; 240-245

#### *Other published studies that have used the CARE Measure*

- Mercer SW and Watt GMC. The inverse care law: clinical primary care encounters in deprived and affluent areas of Scotland. *Annals of Family Medicine* 2007, 5: 503-510
- Mercer SW, Fitzpatrick B, Gourlay G, Vojt G, McConnachie A, Watt GCM. More time for complex consultations in a high deprivation practice is associated with increased patient enablement. *BJGP* 2007, 57: 960-966

#### *Studies that informed or support the use of the CARE Measure*

- Mercer SW, Cawston PG, Bikker AP. Patients' views on consultation quality in primary care in an area of high deprivation; a qualitative study. *BMC Family Medicine* 2007, 8:22
- Mercer SW and Reilly D. A qualitative study of patients views on the consultation at the Glasgow Homoeopathic Hospital. *Patient Education and Counselling* 2004, 53 (1), 13-18
- Mercer SW, Reilly D and Watt GCM. The importance of empathy in the enablement of patients attending the Glasgow Homoeopathic Hospital. *BJGP* 2002, 52 (484), 901-905

#### **Current use of the CARE Measure**

- The CARE Measure has been named in the draft Quality Strategy of the Scottish Government as the measure of choice for healthcare staff in the NHS in Scotland for patient feedback in appraisal and revalidation.
- The CARE Measure is accredited and is routinely used in the appraisal of GPs in Scotland by NHS Education Scotland and the RCGP Scotland (since 2003) and has been available as a web-based feedback system run by RCGP Scotland since 2006 ([www.caremeasure.org](http://www.caremeasure.org))
- Also accredited by RCGP (UK) for use in membership by assessment and international membership by assessment since 2006
- The CARE Measure is a compulsory component of the assessment of all GPs in training in the UK ( since 2007), as part of workplace-based assessment for the new MRCGP exam

- The CARE Measure is being used (with the CQI\_2) in the 'Year of Care'. The year of care initiative is a partnership between the Department of Health, Diabetes UK, NHS Diabetes and The Health Foundation which aims to deliver a personalised approach to care for people with long term conditions, including support for self management
- The CARE Measure has been piloted as a patient-feedback tool by the Royal College of Anaesthetists
- The CARE Measure has been included in the National Quality Measures Clearing House (NQMCH) in the USA, which is sponsored by the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services, and is a public repository for evidence-based quality measures and measure sets. See <http://www.qualitymeasures.ahrq.gov/>
- The CARE Measure has been translated and validated into a number of languages and is being used by independently by researchers in Europe, USA, Canada, China, and Japan
- Current collaborative research with Stewart Mercer:
  - Japan: Daiwa Foundation funding in 2006 to translate, validate, and pilot the utility of the CARE Measure in primary care (ongoing) with Nagoya University.
  - Hong Kong: Translation, validity and performance of the CARE Measure with Dr Colman Fung, Chinese University of Hong Kong (ongoing)
  - Germany: German version of the CARE Measure translated and validated by Dr Melanie Neumann and colleagues at the Universities of Cologne and Frieburg and used in studies in oncology and primary care (ongoing)

#### **Independent research / use of the CARE Measure**

- UK:
  - Dr Susan Kennedy, Lecturer, Department of Nursing Studies at Glasgow University is using the CARE Measure as an assessment of nurses communication in 'Keep Well' the Scottish Governments' flagship health inequalities intervention
  - Dr Madeline Murtagh, Senior Lecturer in Social Science and Public Health at the Institute for Health and Society at University of Newcastle is using the CARE Measure in a relational approach to decision making support in consultations

- Dr Sarah Flyer, Clinical Neuropsychologist, is using the CARE measure to evaluate a training course for nurses and unqualified care staff for the Acquired Brain Injury Service in Cumbria
  - Dr. Johnson D'souza, a General Practitioner based in Castleford in West Yorkshire, undertaking a MSc in Diabetes Care, who is using the CARE Measure to assess patient satisfaction with the care plans in diabetes care
  - Dr Selena Ledwidge, a Specialist Registrar in General Surgery in Oxford currently studying for a Masters in Surgical Education at Imperial College, London who is using the CARE measure to obtain feedback from patients
  - Michelle McArthur, PhD student in Clinical Psychology, UK, is using an adopted version of the CARE Measure to assess veterinarians' communication skills and client (owner) satisfaction
- International:
- Professor Bruce Barrett, Department of Family Medicine, University of Wisconsin, USA, in a National Institute of Health-funded major RCT (with over 700 patients) has found that the CARE Measure (as a measure of the doctor-patient interaction) predicts severity and duration of illness and immune response in the common cold. (see Rakel DP, Hoelt TJ, Barrett BP, Chewing BA, Craig BM, Niu M. Practitioner empathy and duration of the common cold. *Fam Med* 2009, 41(7); 494-501)
  - Dr Robin Muller, Behavioral Health, Consultant in Behavioural Health training, The Center for Family Practice of The Greenville Hospital System University Medical Group, South Carolina, USA, is using the CARE Measure to evaluate the effectiveness of Balint groups on consultation quality in primary care
  - Emily Bower, PhD student from Department of Family medicine, University of West Virginia, USA, is assessing medical students communication and empathy in OSCE using CARE Measure
  - Dr Lee McKinley has included the CARE Measure in the curriculum for second year medical students at Indiana University School of Medicine, USA
  - Dr Chris Dietz, Instructor of Anaesthetics, is using the CARE measure in an ethical training intervention with anaesthetists, at the Mayo College of Medicine. Mayo Clinic, Rochester, Minnesota, USA
  - Pieter Greef, a Master's student in Research Psychology, enrolled at the North-West University, South Africa, using the CARE Measure in his research on the development of a short-term training programme for forensic



interviewing.

- Robin Takeshita, a PhD student from Argosy University in Hawaii, is using the CARE Measure to assess oriental medicine practitioner empathy in the patient population in Hawaii
- Dr Jean-Pierre Jacquet is translating the CARE Measure into French and using it in the teaching of GPs at the National College of GP Teachers (College National des Generalistes Enseignants)
- Dr Pedro Laja and Professor Francisco Cardoso, Universidade de Trás-os-Montes e Alto Douro (UTAD), Portugal have translated the CARE Measure into Portugese and are using it in a study on emotional intelligence and therapeutic alliance in clinical psychology
- Dr Julia Strupp is using the German version of the CARE Measure in a study with people suffering from MS at the Cologne Centre for Palliative Care
- Greek translation of the CARE Measure used at the 3rd International Geriatric Symposium, September, 2007, Rhodes, Greece, in an interactive clinical seminar with Greek general practitioners concerning their attitudes towards the frail geriatric patient in primary health care
- Dr Helen Richards, Lead Clinical Psychologist, is using the CARE measure with patients who have been assessed by medical students, at Mercy University Hospital, Cork

**The CARE Measure**

**1. Please rate the following statements about today's consultation. Please tick one box for each statement and answer every statement.**

<i>How was the doctor at ...</i>	Poor	Fair	Good	Very Good	Excellent	Does Not Apply
<b>1. Making you feel at ease.....</b> <i>(being friendly and warm towards you, treating you with respect; not cold or abrupt)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>2. Letting you tell your " story".....</b> <i>(giving you time to fully describe your illness in your own words; not interrupting or diverting you)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>3. Really listening .....</b> <i>(paying close attention to what you were saying; not looking at the notes or computer as you were talking)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>4. Being interested in you as a whole person ...</b> <i>(asking/knowing relevant details about your life, your situation; not treating you as "just a number")</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>5. Fully understanding your concerns.....</b> <i>(communicating that he/she had accurately understood your concerns; not overlooking or dismissing anything)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>6. Showing care and compassion...</b> <i>(seeming genuinely concerned, connecting with you on a human level; not being indifferent or "detached")</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>7. Being Positive.....</b> <i>(having a positive approach and a positive attitude; being honest but not negative about your problems)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

<b>8. Explaining things clearly.....</b> <i>(fully answering your questions, explaining clearly, giving you adequate information; not being vague)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>9. Helping you to take control.....</b> <i>(exploring with you what you can do to improve your health yourself; encouraging rather than "lecturing" you)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>10. Making a plan of action with you ...</b> <i>(discussing the options, involving you in decisions as much as you want to be involved; not ignoring your views)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

The scoring system for each item is 'poor'=1, 'fair' = 2, 'good' = 3, 'very good' = 4, and 'excellent'= 5. All ten items are then added, giving a maximum possible score of 50, and a minimum of 10. Up to two 'Not Applicable' responses or missing values are allowable, and are replaced with the average score for the remaining items. Questionnaires with more than two missing values or 'Not Applicable' responses are removed from the analysis.

© Stewart W Mercer 2004

*The CARE measure can be used free of charge. The Intellectual Property rights rest with Professor Stewart Mercer on behalf of the Scottish Government. The measure may not be used on a commercial basis. If you would like more information, please contact;*

For further information, and to download the measure please visit:

[www.gla.ac.uk/departments/generalpractice/caremeasure.htm](http://www.gla.ac.uk/departments/generalpractice/caremeasure.htm)

# Patient Information Leaflet



UNIVERSITY OF  
STIRLING

SCHOOL OF  
NURSING, MIDWIFERY  
AND HEALTH

## 1 A Study Comparing Waiting Time & Perceived Empathy with Overall Satisfaction

The health service is constantly looking for ways to improve their services. This study aims to determine whether or not patients' perceptions of how understanding staff in A&E is a good measure of satisfaction of care, in comparison to time spent in the department.

## 2 Invitation

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether you wish to take part. Thank you for reading this.

## 3 What is the purpose of the study?

All Accident and Emergency (A&E) Departments in the UK must record the duration of your visit in the Department. This information is used to determine the quality of care you receive. This study seeks to identify if how you feel you were treated in A&E is a better measure of the quality of care you have received in the department. The self-completed questionnaire has been developed to enable you to assess health care staffs' communication and understanding (empathy) during your consultation. The questionnaire consists of items relating to your perception of health care staffs' understanding of, and response to any concerns and fears you may have in relation to your reason for your visit.

## 4 Why have I been chosen?

The study aims to get 70 adult patients (over 16 years) to complete the questionnaire during specific time periods in A&E. Anyone willing to participate and consent is asked to complete the questionnaire.

## 5 Do I have to take part?

No. It is up to you to decide whether to take part. If you do decide to take part, you will be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

Highland Campus:  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JH  
Tel: +44 (0) 1463 255655  
Fax: +44 (0) 1463 255654

Stirling Campus:  
Stirling  
FK9 4LA  
Tel: +44 (0) 1798 466340  
Fax: +44 (0) 1798 466333

Western Isles Campus:  
Western Isles Hospital  
MacAulay Road  
Stornoway Isle of Lewis HE1 2AF  
Tel: +44 (0) 1851 708243  
Fax: +44 (0) 1851 708070

Version 9 - August 2014

## 6 What will happen to me if I take part?

You will be asked to complete questionnaire at the end of your A&E visit – either before you go home or before you go to another area for further care. The questionnaire takes a few minutes to complete. Your answers will only be seen by the person conducting the study.

## 7 What are the possible benefits of taking part?

The study will help us to understand and possibly develop measures of quality that are meaningful to you, the patient.

## 8 Will my taking part in this study be kept confidential?

All information, which is collected, about you during the course of the research will be kept strictly confidential. Any information about you which leaves the hospital/surgery will have your name and address removed so that you cannot be recognised from it.

## 9 What will happen to the results of the research study?

The responses will be held in a database and analysed to determine the results. The results will be published in an academic health care journal and shared with the Health Board. You will not be identified in any report/publication.

## 13 Who is organising and funding the research?

The organisation and sponsor of the research is the University of Stirling.

## 14 Who has reviewed the study?

The study has been reviewed by the Research and Ethics Committee, University of Stirling.

## 15 Contacts for Further Information

Many thanks for participating in this study. If you require any further information on the study you can contact Michelle Beattie, Lecturer, School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Centre for Health Science, Old Perth Road, Inverness, IV2 3JH. Telephone 01463 255622 or e-mail [michelle.beattie@stir.ac.uk](mailto:michelle.beattie@stir.ac.uk)

You can also choose to contact William Lauder, Head of Department, School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Centre for Health Science, Old Perth Road, Inverness, IV2 3JH. Telephone 01463 255619 or e-mail [william.lauder@stir.ac.uk](mailto:william.lauder@stir.ac.uk)

Highland Campus:  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JH  
Tel: +44 (0) 1463 255655  
Fax: +44 (0) 1463 255654

Stirling Campus:  
Stirling  
FK9 4LA  
Tel: +44 (0) 1796 466340  
Fax: +44 (0) 1796 466333

Western Isles Campus:  
Western Isles Hospital  
MacAulay Road  
Stomoway Isle of Lewis HB1 2AF  
Tel: +44 (0) 1851 708243  
Fax: +44 (0) 1851 708070

FILE MESSAGE

Ignore X Delete Reply Reply All Forward Meeting More

Admin To Manager Done Create New

Team Email Reply & Delete

Rules OneNote Actions

Delete Respond Quick Steps Move



Sun 14/08/2011 19:42

Ruth Jepson

RE: A&E Ethics Study

To Michelle Beattie

Cc Fiona Harris; SarahJane Gilvear

You replied to this message on 18/08/2011 15:04.

Hi Michelle

Thanks for your reply and addressing our comments satisfactorily. Your study has now been approved by Chair's action.

Best wishes

Ruth

Dr Ruth Jepson  
Senior Research Fellow  
Co-Director of Centre for Public Health and Population Health Research  
School of Nursing, Midwifery and Health  
University of Stirling  
Stirling  
Scotland, UK  
FK9 4LA  
Email: [ruth.jepson@stir.ac.uk](mailto:ruth.jepson@stir.ac.uk)  
Tel: 01786 466402

Room 4T 21

## Appendix 4: National Health Services Research and Ethics Committee (NHS REC) North of Scotland, Ethics Application and Approval

NHS REC Form

Reference:  
11-NS/0025

IRAS Version 3.3

Welcome to the Integrated Research Application System

IRAS Project Filter

The integrated dataset required for your project will be created from the answers you give to the following questions. The system will generate only those questions and sections which (a) apply to your study type and (b) are required by the bodies reviewing your study. Please ensure you answer all the questions before proceeding with your applications.

Please enter a short title for this project (maximum 70 characters)  
Time and Perceived Empathy with Overall Patient Satisfaction in A&E

1. Is your project research?

Yes  No

2. Select one category from the list below:

- Clinical trial of an investigational medicinal product
- Clinical investigation or other study of a medical device
- Combined trial of an investigational medicinal product and an investigational medical device
- Other clinical trial to study a novel intervention or randomised clinical trial to compare interventions in clinical practice
- Basic science study involving procedures with human participants
- Study administering questionnaires/interviews for quantitative analysis, or using mixed quantitative/qualitative methodology
- Study involving qualitative methods only
- Study limited to working with human tissue samples (or other human biological samples) and data (specific project only)
- Study limited to working with data (specific project only)
- Research tissue bank
- Research database

If your work does not fit any of these categories, select the option below:

Other study

2a. Please answer the following question(s):

- a) Does the study involve the use of any ionising radiation?  Yes  No
- b) Will you be taking new human tissue samples (or other human biological samples)?  Yes  No
- c) Will you be using existing human tissue samples (or other human biological samples)?  Yes  No

3. In which countries of the UK will the research sites be located?(Tick all that apply)

- England
- Scotland
- Wales
- Northern Ireland

3a. In which country of the UK will the lead NHS R&D office be located:

Date: 08/09/2011

1

88069/244424/1/539

- England  
 Scotland  
 Wales  
 Northern Ireland  
 This study does not involve the NHS

**4. Which review bodies are you applying to?**

- NHS/HSC Research and Development offices  
 Research Ethics Committee  
 National Information Governance Board for Health and Social Care (NIGB)  
 Ministry of Justice (MoJ)  
 National Offender Management Service (NOMS) (Prisons & Probation)

*For NHS/HSC R&D offices, the CI must create Site-Specific Information Forms for each site, in addition to the study-wide forms, and transfer them to the PIs or local collaborators.*

**5. Will any research sites in this study be NHS organisations?**

- Yes    No

**6. Do you plan to include any participants who are children?**

- Yes    No

**7. Do you plan at any stage of the project to undertake intrusive research involving adults lacking capacity to consent for themselves?**

- Yes    No

*Answer Yes if you plan to recruit living participants aged 16 or over who lack capacity, or to retain them in the study following loss of capacity. Intrusive research means any research with the living requiring consent in law. This includes use of identifiable tissue samples or personal information, except where application is being made to the NIGB Ethics and Confidentiality Committee to set aside the common law duty of confidentiality in England and Wales. Please consult the guidance notes for further information on the legal frameworks for research involving adults lacking capacity in the UK.*

**8. Do you plan to include any participants who are prisoners or young offenders in the custody of HM Prison Service or who are offenders supervised by the probation service in England or Wales?**

- Yes    No

**9. Is the study or any part of it being undertaken as an educational project?**

- Yes    No

Please describe briefly the involvement of the student(s):  
The only student involved would be the applicant, Michelle Beattie, undertaking the study as part of a PhD

**9a. Is the project being undertaken in part fulfilment of a PhD or other doctorate?**

- Yes    No

**10. Will this research be financially supported by the United States Department of Health and Human Services or any of its divisions, agencies or programs?**



Yes  No

**Integrated Research Application System**  
**Application Form for Research administering questionnaires/interviews for quantitative analysis or mixed methodology study**



**Application to NHS/HSC Research Ethics Committee**

The Chief Investigator should complete this form. Guidance on the questions is available wherever you see this symbol displayed. We recommend reading the guidance first. The complete guidance and a glossary are available by selecting [Help](#).

Please define any terms or acronyms that might not be familiar to lay reviewers of the application.

**Short title and version number:** (maximum 70 characters - this will be inserted as header on all forms)  
 Time and Perceived Empathy with Overall Patient Satisfaction in A&E

Please complete these details after you have booked the REC application for review.

**REC Name:**  
 NRES Committee North of Scotland

**REC Reference Number:**  
 11-NS/0025

**Submission date:**  
 08/09/2011

**PART A: Core study information**

**1. ADMINISTRATIVE DETAILS**

**A1. Full title of the research:**

Are We Measuring the Correct Dimensions of Quality in A&E? A Study Comparing Time Waited and Perceived Empathy with Overall Patient Satisfaction

**A2-1. Educational projects**

Name and contact details of student(s):

Name and contact details of academic supervisor(s):

**Academic supervisor 1**

	Title	Forename/Initials	Surname
	Professor	William	Lauder
Address	University of Stirling School of Nursing, Midwifery and Health Stirling		
Post Code	FK9 4LA		

E-mail	william.lauder@stir.ac.uk
Telephone	01786486344
Fax	01786486344

Please state which academic supervisor(s) has responsibility for which student(s):  
Please click "Save now" before completing this table. This will ensure that all of the student and academic supervisor details are shown correctly.

Student(s)	Academic supervisor(s)
------------	------------------------

A copy of a current CV for the student and the academic supervisor (maximum 2 pages of A4) must be submitted with the application.

#### A2-2. Who will act as Chief Investigator for this study?

- Student  
 Academic supervisor  
 Other

#### A3-1. Chief Investigator:

	Title Forename/Initials Surname
	Mrs Michelle Beattie
Post	Lecturer
Qualifications	MSc Advanced Practice Nursing BSc Hons Health and Social Care RN (Adult)
Employer	University of Stirling
Work Address	Centre for Health Science Old Perth Road Inverness
Post Code	IV2 3JH
Work E-mail	michelle.beattie@stir.ac.uk
* Personal E-mail	
Work Telephone	01463255622
* Personal Telephone/Mobile	
Fax	01463255654

\* This information is optional. It will not be placed in the public domain or disclosed to any other third party without prior consent.

A copy of a current CV (maximum 2 pages of A4) for the Chief Investigator must be submitted with the application.

#### A4. Who is the contact on behalf of the sponsor for all correspondence relating to applications for this project?

This contact will receive copies of all correspondence from REC and R&D reviewers that is sent to the CI.

	Title Forename/Initials Surname
	Mrs Carol Johnstone
Address	Research and Enterprise Office University of Stirling Stirling
Post Code	FK9 4LA

E-mail	carol.johnstone@stir.ac.uk
Telephone	01786486890
Fax	

**A5-1. Research reference numbers.** *Please give any relevant references for your study:*

Applicant's/organisation's own reference number, e.g. R & D (if available): Highland 779

Sponsor's/protocol number:

Protocol Version:

Protocol Date:

Funder's reference number:

Project website:

**Additional reference number(s):**

Ref.Number	Description	Reference Number

*Registration of research studies is encouraged wherever possible. You may be able to register your study through your NHS organisation or a register run by a medical research charity, or publish your protocol through an open access publisher. If you have registered your study please give details in the "Additional reference number(s)" section.*

**A5-2. Is this application linked to a previous study or another current application?**

Yes  No

*Please give brief details and reference numbers.*

**2. OVERVIEW OF THE RESEARCH**

*To provide all the information required by review bodies and research information systems, we ask a number of specific questions. This section invites you to give an overview using language comprehensible to lay reviewers and members of the public. Please read the guidance notes for advice on this section.*

**A6-1. Summary of the study.** *Please provide a brief summary of the research (maximum 300 words) using language easily understood by lay reviewers and members of the public. This summary will be published on the website of the National Research Ethics Service following the ethical review.*

Much emphasis has been placed on the 4 hour wait target in accident and emergency (A&E) from the UK Government since it was implemented in 2007. There are concerns that time is an inaccurate reflection of the patients' health care experience within A&E departments. There is a risk that as we focus on measurable outcomes, such as time, important aspects of caring, namely empathy become marginalised.

This study aims to compare whether patient's scores of relational empathy capture their perception of health care quality more accurately than waiting time in A&E.

Do patients rate their overall quality of care in A&E higher when they have been seen quicker, or been treated in an empathetic manner, or both? This study hypothesises that there will be a positive relationship between patients' empathy scores and their overall rating of quality of care. There will be no relationship between the time patients have spent in A&E and their overall rating of quality of care.

Seventy patients will be required to complete the CARE measure questionnaire (validated measure of perceived empathy) with additional questions on demographics and their overall rating of satisfaction of care. CARE measure scores will be compared with time waited in A&E to determine whether empathy scores or time waited have an influence of the patients overall satisfaction score.

The analysis of the data will be carried out using a range of statistical tests (e.g. t-test) and will aim to prove that empathy is either more important or at least as important as time waited. The conclusion will aim to suggest that a

new approach to measuring patient satisfaction be developed for hospital HEAT targets in Scotland.

**A6-2. Summary of main issues.** Please summarise the main ethical, legal, or management issues arising from your study and say how you have addressed them.

*Not all studies raise significant issues. Some studies may have straightforward ethical or other issues that can be identified and managed routinely. Others may present significant issues requiring further consideration by a REC, R&D office or other review body (as appropriate to the issue). Studies that present a minimal risk to participants may raise complex organisational or legal issues. You should try to consider all the types of issues that the different reviewers may need to consider.*

#### Ethical Issues

Consent - patients will be given an information leaflet on arrival by A&E staff. Those willing to participate will be required to sign a consent form. No patients unable to consent for themselves will be included in the study.

Use of identifiable patient data - The principal investigator (PI) will insert the patients CHI (community health index) number onto the questionnaire before patient completion. CHI numbers are required to match the patients questionnaire to the database which holds information on the length of time each patient has spent in the department. Once the time has been recorded on the questionnaire the CHI number will be removed and study numbers will be used to identify participants. Patients will consent to the use of their CHI number.

Patients may perceive a risk that their care may be influenced depending on their responses to the CARE measure, or, whether or not they choose to participate. Care measure questionnaires will only be distributed after the patient has had their consultation and primary treatment. Patients will be reassured that their questionnaire responses will only be seen by those conducting the research and that responses should be put in the envelope provided. Information on the voluntary nature of participation will be reinforced on the patient advice leaflet. Feedback to the participating Board and preparation of a paper for publication will not identify any patients.

Staff may also have reservations about the study as the tool may highlight poor levels of perceived empathy. Information to staff in written and presentation format will be given and staff will be informed that the CARE measure score will be calculated as a departmental score, rather than an individual score. The Senior Charge Nurse, Area Manager and Clinical Educator of the A&E Department are aware of the proposed study.

**A6-3. Proportionate review of REC application** The initial project filter has identified that your study may be suitable for proportionate review by a REC sub-committee. Please consult the current guidance notes from NRES and indicate whether you wish to apply through the proportionate review service or, taking into account your answer to A6-2, you consider there are ethical issues that require consideration at a full REC meeting.

Yes - proportionate review  No - review by full REC meeting

Further comments (optional):

Note: This question only applies to the REC application.

### 3. PURPOSE AND DESIGN OF THE RESEARCH

**A7. Select the appropriate methodology description for this research. Please tick all that apply.**

- Case series/ case note review
- Case control
- Cohort observation
- Controlled trial without randomisation
- Cross-sectional study
- Database analysis
- Epidemiology
- Feasibility/ pilot study
- Laboratory study

- Metanalysis
- Qualitative research
- Questionnaire, interview or observation study
- Randomised controlled trial
- Other (please specify)

**A10. What is the principal research question/objective? Please put this in language comprehensible to a lay person.**

To determine how and if staff display and communicate understanding of a patient's needs and whether this influences patient's overall satisfaction of the service they receive in A&E in comparison to the total time they have spent in A&E

**A11. What are the secondary research questions/objectives if applicable? Please put this in language comprehensible to a lay person.**

Do patients rate their overall quality of care in A&E higher when they have been seen quicker, or been treated in an empathetic manner, or both?

Is there a relationship between the relational empathy score of patients and overall satisfaction score of those patients?

Is there a relationship between total time spent in A&E measured in minutes and the overall satisfaction score of those patients?

Is there a positive relationship between patients' empathy scores and their overall rating of quality of care.

Is there a null relationship between the time patients have spent in A&E and their overall rating of quality of care.

**A12. What is the scientific justification for the research? Please put this in language comprehensible to a lay person.**

The NHS in Scotland are always looking for ways to improve their services. Currently, the only quality measure of A&E services is the 4 hour wait target (Audit Scotland, 2010). Although waiting times have improved across the UK there are concerns that time is an inaccurate reflection of the patients' health care experience within accident and emergency (A&E) departments (Jolly and Clancy, 2009). There is a risk that as we focus on measurable outcomes, such as time, important aspects of caring, namely empathy become marginalised. As the Scottish Government have three quality ambitions which include supporting person-centredness there is a need to develop measures of quality which are more reflective of the patients experience (Scottish Government, 2010). The Scottish Government are currently working on a Care Governance Measurement Framework (Scottish Government, 2011) to underpin the delivery of the three quality ambitions. The framework is flexible enough to enable the identification of local and national variables to measure for quality improvement purposes. Boards are required to address significant gaps to provide the data that will enable healthcare staff to focus on activities that will underpin the delivery of the Quality Strategy (Scottish Government, 2010).

This study aims to determine whether or not the CARE measure (validated to capture relational empathy) captures an important dimension of health care quality from the patients perspective. The quality of direct clinical care has long been associated with caring theories. While these theories have derived mainly from nursing, they are relevant to all healthcare staff. The attributes of the caring behaviours identified from these theories map well to the 7Cs - care and compassion, communication, collaboration, clean and safe environment, continuity and clinical excellence - which have been defined by the people of Scotland as high quality healthcare (Scottish Government, 2010). Like concepts of quality, caring remains difficult to precisely determine. However, many definitions or conceptions of caring capture the notion of empathy, or the ability to communicate an understanding of the patients world (Reynolds, 2000). There remains much debate over whether or not caring can indeed be measured. Although, there is a convergence in understanding that although the wholeness of caring is likely unmeasurable, there are important elements of caring that can indeed be captured, and subsequently, measured (Watson, 2009). The Consultation and Relational Empathy (CARE) measure has demonstrated validity in measuring the elusive notion of empathy (Mercer et al 2004, Mercer et al 2005). Although a multitude of tools exist to measure aspects of caring (Larsen & Ferketich 1983, Nyberg 1990, Duffy 1992, Watson and Lea 1997). Many other tools have limited transferability into practice, due to lengthy and complex questionnaires (Wolf, 1988), the need to measure staff and patients perceptions (Larsen & Ferketich 1983) or the inappropriateness to fit in alternative settings (Reynolds, 2000). The CARE measure has been chosen as it captures a modern conception of caring, 'collaboration', rather than 'doing for'; it has a simplistic understanding of empathy, has

face validity from a Scottish population, is quick to complete and is currently the tool of choice in many other national (for example Bikker et al 2005) and international research projects (for example Fung et al 2009).

#### References

Audit Scotland (2010) Emergency Departments, Prepared for the Auditor General Scotland, August. [http://www.audit-scotland.gov.uk/docs/health/2010/nr\\_100812\\_emergency\\_departments.pdf](http://www.audit-scotland.gov.uk/docs/health/2010/nr_100812_emergency_departments.pdf)

Bikker AP, Mercer SW, Reilly D. (2005) A pilot prospective study on the consultation and relational empathy, patient enablement, and health changes over 12 months, in patients going to the Glasgow Homoeopathic Hospital. *J Alt. Comp. Med.* 2005, 11 (4), 591-600

Fung C, Hua A, Tam L, Mercer SW. (2009) Reliability and validity of the Chinese version of the CARE Measure in a primary care setting in Hong Kong. *Family Practice* 2009 28(5):398-406

Jolly E and Clancy M (2009) Waiting times are they that important? A patient survey, *Emergency Medicine Journal*, October, Vol. 26, No. 10726.

Larsen P & Ferketich S (1983) Patients' satisfaction with nurses' caring during hospitalisation. *Western Journal for Nursing Research*, 15 (6), 187-193.

Mercer S, Watt G, Maxwell M, Heaney D and (2004) The development and preliminary validation of the Consultation and Relational Empathy (CARE) Measure: an empathy-based consultation process measure, *Family Practice*, 21 (6), 699-705.

Mercer S, McConnachie A, Maxwell M, Heaney D and Watt G (2005) Relevance and performance of the Consultation and Relational Empathy (CARE) Measure in general practice. *Family Practice*, 22 (3), 328-334.

Nyberg J (1990) cited by Watson J (2009) *Assessing and Measuring Caring in Nursing and Health Sciences*, Springer publishers, Chapter 9, Nyberg Caring Assessment Scale, 113-116.

Reynolds W (2000) *The Measurement and Development of Empathy in Nursing*. Ashgate publishers.

Scottish Government (2010) *The Healthcare Quality Strategy for NHS Scotland*, May. <http://www.scotland.gov.uk/Resource/Doc/311867/0098354.pdf>  
Accessed 15th October 2010

Scottish Government (2011) *The NMAHP Contribution to Quality: Towards a Care Governance Measurement Framework, Draft Version Three 30/08/11*. <http://www.knowledge.scot.nhs.uk/media/CLT/ResourceUploads/1012046/The%20Care%20Governance%20Measurement%20Framework300811.pdf>

Watson J and Lea A (1997) *The Caring Dimensions Inventory (CDI) content validity, reliability and scaling*, *Journal of Advanced Nursing*, 25, 87-94.

Watson J (2002) *Assessing and Measuring Caring in Nursing and Health Science*, New York, Springer publisher.

Wolf, 1986

**A13. Please summarise your design and methodology.** *It should be clear exactly what will happen to the research participant, how many times and in what order. Please complete this section in language comprehensible to the lay person. Do not simply reproduce or refer to the protocol. Further guidance is available in the guidance notes.*

#### Sampling

Empathy will be measured using the CARE measurement tool (Mercer et al 2004). Mercer et al (2005) suggest that at least 50 people be included in any study using this tool to ensure validity and reliability. Using a t-test to determine difference in scores between patients who are satisfied and dissatisfied with their overall care will require a minimum of 64 participants. This number will provide a sample sufficient to detect an effect size of 0.5 at a power of 80% (Machin et al 2008). The study therefore aims to collect a minimum of 70 self completed questionnaires.

There are recognised time periods when A&E Departments are busier and quieter. To ensure the sample of questionnaires are representative of differences in time, data will be collected during recognised busy and quiet

times.

Patients requiring immediate resuscitation, those unable to give informed consent due to incapacity or altered levels of consciousness and children, less than 18 years, will be excluded from the study (this would include those whose capacity has been temporarily affected by opioid analgesia). Patients returning for a second planned visit will also be excluded as they are unrepresentative of the population attending A&E, as their wait time and CARE score may be influenced by their primary visit. All of these excluded patients will be identified by A&E staff and the researcher will only approach patients fitting the inclusion criteria once identified by A&E staff and following their treatment.

#### Data Collection

A member of A&E staff will provide an information leaflet to all eligible patients and determine whether or not they wish to participate. Either the principal applicant or clinical educator (based in A&E) would be on site during data collection times to answer any patient queries, ensure consent forms were signed, insert the patients CHI (community health index) number onto the questionnaire and distribute and collect the questionnaires. Those who have consented will be given a self-completion questionnaire following their consultation and prior to their departure. Completed questionnaires will be returned in sealed envelopes to the research team or in a questionnaire collection box within A&E reception. Results of relational empathy scores and overall satisfaction scores will be linked to an existing database to establish the length of time the patient has been in A&E using the patients CHI number. The following variables will be collected:-

**Empathy Score** – this will be measured using the existing Consultation and Relational Empathy (CARE) measure questionnaire (Mercer et al 2004). The tool is composed of 10 questions, each requiring the patient to select the most appropriate response based on a five point Likert scale, and should only take about 5 minutes to complete. Each response totals to provide an overall score of relational empathy ranging between 10 and 50. The CARE measure tool has been validated for use in primary care (Mercer et al 2004; Mercer et al 2005) and some studies have utilised this within a secondary care setting (Mercer & Murphy 2008; Mercer et al 2008).

**Patient Satisfaction** – patients will be asked to determine their overall level of satisfaction using a five point Likert scale. This will assess whether the patient perceives their care to have been very good, good, fair, poor, or very poor.

**Time** – the time waited in A&E is recorded from the time the patient arrives in the department until they leave the department. This is currently recorded within an existing database for Information Services Division at the Scottish Government. Time will be recorded as a continuous variable as number of minutes.

**Demographics** – Age and gender will also be recorded and used as covariants in analysis in order to assess if these variables account for differences in quality of care scores.

**Day and time of arrival** – to identify the busier and quieter times in A&E, the day and time of arrival will be recorded to enable analysis to assess if they affect perceptions of empathy of quality of care.

**Health care professional** – patients will be asked to select whether their primary consultation and care had been from either a doctor or a nurse.

#### A14-1. In which aspects of the research process have you actively involved, or will you involve, patients, service users, and/or their carers, or members of the public?

- Design of the research
- Management of the research
- Undertaking the research
- Analysis of results
- Dissemination of findings
- None of the above

*Give details of involvement, or if none please justify the absence of involvement.*

Patients will ONLY be involved as consented participants in the research.

#### 4. RISKS AND ETHICAL ISSUES

##### RESEARCH PARTICIPANTS



**A17-1. Please list the principal inclusion criteria (list the most important, max 5000 characters).**

All patients entering A&E for a primary visit within data collection time frames.  
All patients 18 years or older.  
All patients with capacity to consent and complete the questionnaire.

**A17-2. Please list the principal exclusion criteria (list the most important, max 5000 characters).**

Patients under 18 years.  
Patients requiring immediate resuscitation.  
Those unable to give informed consent due to incapacity as determined by A&E staff members i.e. dementia, confusion  
Those who are temporarily incapacitated as determined by A&E staff members i.e. altered levels of consciousness, opioid analgesia.  
Patients returning for a second planned visit will also be excluded as they are unrepresentative of the population attending A&E, as their wait time and CARE score may be influenced by their primary visit. For example those being asked to return the next day for a wound dressing review.

**RESEARCH PROCEDURES, RISKS AND BENEFITS**

**A18. Give details of all non-clinical intervention(s) or procedure(s) that will be received by participants as part of the research protocol. These include seeking consent, interviews, non-clinical observations and use of questionnaires.**

Please complete the columns for each intervention/procedure as follows:

1. Total number of interventions/procedures to be received by each participant as part of the research protocol.
2. If this intervention/procedure would be routinely given to participants as part of their care outside the research, how many of the total would be routine?
3. Average time taken per intervention/procedure (minutes, hours or days)
4. Details of who will conduct the intervention/procedure, and where it will take place.

Intervention or procedure	1	2	3	4
Patient Information Sheet	1	0	5mins	Administered by A&E staff
Informed Consent	1	0	1mins	Taken by PI or Clinical Educator
Patient Complete Questionnaire	1	0	5mins	Administered by PI or Clinical Educator. Collected in envelope by PI or returned to completion box in A&E reception

**A21. How long do you expect each participant to be in the study in total?**

Each participant will take approximately 11 mins to read information sheet, sign consent forms and complete questionnaire. As the patient will complete the questionnaire following consultation i.e. prior to discharge or transfer to another area of care, their length of time in the department will be dictated by their clinical condition, not the study. Following completion of the questionnaire the patient's involvement in the study will end.

**A22. What are the potential risks and burdens for research participants and how will you minimise them?**

*For all studies, describe any potential adverse effects, pain, discomfort, distress, intrusion, inconvenience or changes to lifestyle. Only describe risks or burdens that could occur as a result of participation in the research. Say what steps would be taken to minimise risks and burdens as far as possible.*

Breach of confidentiality - As the patients CHI (community health index) number will be added to the questionnaire there is a risk that patient information could be disclosed. This risk will be removed by removing the patients CHI number from the questionnaire once their time spent in the Department is obtained from the database. This step will only be completed by the PI or A&E Clinical Educator. The matching of questionnaire to time spent in the department will happen immediately where possible or, during busy periods, at the end of the data collection period (4 hour slots).

All data collected will then be anonymised and patients data will be given a study number. Questionnaires will be stored in a locked filing cabinet at the University of Stirling, Highland Campus. All data added to Statistical Software Package for Social Scientists (SPSS) will be held on the University of Stirling's secure IT network. No laptops or removable storage devices will be used in line with University policy and the Data Protection Act.

Patients may perceive a risk that their care may be influenced depending on their responses to the CARE measure, or, whether or not they choose to participate. Care measure questionnaires will only be distributed after the patient has had their consultation and primary treatment. Patients will be reassured that their questionnaire responses will only be seen by the researcher and that responses should be put in the envelope provided. Information on the voluntary nature of participation will be reinforced on the patient advice leaflet. Staff may also have reservations about the study as the tool may highlight poor levels of perceived empathy. Information to staff in written and presentation format will be given and staff will be informed that the CARE measure score will be calculated as a departmental score, rather than an individual score. The Senior Charge Nurse, Clinical Educator and Manager have been consulted and involved in the proposed study.

**A23. Will interviews/ questionnaires or group discussions include topics that might be sensitive, embarrassing or upsetting, or is it possible that criminal or other disclosures requiring action could occur during the study?**

Yes  No

**A24. What is the potential for benefit to research participants?**

The study will give patients the opportunity to express how satisfied they were with their care, as well as, influence health care quality measures for the future in A&E.

**A26. What are the potential risks for the researchers themselves? (if any)**

There is a risk that insufficient patients are recruited during data collection. Time required for subsequent data collection will be calculated into the estimated time line.

#### RECRUITMENT AND INFORMED CONSENT

*In this section we ask you to describe the recruitment procedures for the study. Please give separate details for different study groups where appropriate.*

**A27-1. How will potential participants, records or samples be identified? Who will carry this out and what resources will be used? For example, identification may involve a disease register, computerised search of GP records, or review of medical records. Indicate whether this will be done by the direct healthcare team or by researchers acting under arrangements with the responsible care organisation(s).**

All patients entering A&E who do not meet any of the exclusion criteria are eligible to be recruited. A&E staff will determine whether or not the patient has the capacity to participate and then give out the study information sheets. Only those who consent will be given a questionnaire to complete after their consultation.

**A27-2. Will the identification of potential participants involve reviewing or screening the identifiable personal information of patients, service users or any other person?**

Yes  No

Please give details below:

**A28. Will any participants be recruited by publicity through posters, leaflets, adverts or websites?**

Yes  No

**A29. How and by whom will potential participants first be approached?**

The first member of the A&E clinical team will determine whether the patient has capacity to complete the questionnaire and administer the patient information leaflet.

**A30-1. Will you obtain informed consent from or on behalf of research participants?**

Yes  No

*If you will be obtaining consent from adult participants, please give details of who will take consent and how it will be done, with details of any steps to provide information (a written information sheet, videos, or interactive material). Arrangements for adults unable to consent for themselves should be described separately in Part B Section 6, and for children in Part B Section 7.*

*If you plan to seek informed consent from vulnerable groups, say how you will ensure that consent is voluntary and fully informed.*

Consent forms will be explained and administered by the PI or Clinical Educator in A&E. No children or vulnerable groups will be included.

*If you are not obtaining consent, please explain why not.*

Please enclose a copy of the information sheet(s) and consent form(s).

**A30-2. Will you record informed consent (or advice from consultees) in writing?**

Yes  No

**A31. How long will you allow potential participants to decide whether or not to take part?**

From time entering the Department until discharge or transfer to another area of care. Most patients (95%) are seen and transferred or discharged within 4 hours. Patients will be provided with the PIS once deemed suitable for the study by the A&E staff, and will have the opportunity to think about the study until the point where they are approaching discharge or removal to another department. The CI is aware that a potentially large number of patients may be subject to exclusion given that A&E patients are likely to be suffering from pain, but is confident that the provision of the PIS to eligible patients over the period of their stay in A&E will not compromise their care or their levels of distress.

**A33-1. What arrangements have been made for persons who might not adequately understand verbal explanations or written information given in English, or who have special communication needs?(e.g. translation, use of interpreters)**

Only those with access to a translator or interpreter being used for clinical purposes will be included.

**A35. What steps would you take if a participant, who has given informed consent, loses capacity to consent during the study? Tick one option only.**

- The participant and all identifiable data or tissue collected would be withdrawn from the study. Data or tissue which is not identifiable to the research team may be retained.
- The participant would be withdrawn from the study. Identifiable data or tissue already collected with consent would be retained and used in the study. No further data or tissue would be collected or any other research procedures carried out on or in relation to the participant.
- The participant would continue to be included in the study.
- Not applicable – informed consent will not be sought from any participants in this research.
- Not applicable – it is not practicable for the research team to monitor capacity and continued capacity will be assumed.

**Further details:**

If a patient's condition deteriorates during their stay, they would then be subject to the exclusion criteria, for example, incapacity due to altered levels of consciousness and any data collected would be destroyed and not included in the study.

**CONFIDENTIALITY**

In this section, personal data means any data relating to a participant who could potentially be identified. It includes pseudonymised data capable of being linked to a participant through a unique code number.

**Storage and use of personal data during the study**

**A36. Will you be undertaking any of the following activities at any stage (including in the identification of potential participants)? (Tick as appropriate)**

- Access to medical records by those outside the direct healthcare team
- Electronic transfer by magnetic or optical media, email or computer networks
- Sharing of personal data with other organisations
- Export of personal data outside the EEA
- Use of personal addresses, postcodes, faxes, emails or telephone numbers
- Publication of direct quotations from respondents
- Publication of data that might allow identification of individuals
- Use of audio/visual recording devices
- Storage of personal data on any of the following:
  - Manual files including X-rays
  - NHS computers
  - Home or other personal computers
  - University computers
  - Private company computers
  - Laptop computers

**Further details:**

An existing NHS database will require to be accessed to document the time waited in A&E for each participant.

All data from questionnaires will be entered into the Statistical Package for Social Scientists (SPSS) on a university computer. This information will contain no identifiable patient data and subjects will be given a study number.

**A38. How will you ensure the confidentiality of personal data? Please provide a general statement of the policy and procedures for ensuring confidentiality, e.g. anonymisation or pseudonymisation of data.**

No data is required to be personalised, therefore each participant will be given a study number within SPSS.

**A40. Who will have access to participants' personal data during the study? Where access is by individuals outside the direct care team, please justify and say whether consent will be sought.**

No persons other than the CI and Clinical Educator within A&E will have access to participant's personal data during the study.

**Storage and use of data after the end of the study**

**A43. How long will personal data be stored or accessed after the study has ended?**

- Less than 3 months
- 3 – 6 months

- 6 – 12 months  
 12 months – 3 years  
 Over 3 years

**INCENTIVES AND PAYMENTS**

**A46. Will research participants receive any payments, reimbursement of expenses or any other benefits or incentives for taking part in this research?**

- Yes  No

**A47. Will individual researchers receive any personal payment over and above normal salary, or any other benefits or incentives, for taking part in this research?**

- Yes  No

**A48. Does the Chief Investigator or any other investigator/collaborator have any direct personal involvement (e.g. financial, share holding, personal relationship etc.) in the organisations sponsoring or funding the research that may give rise to a possible conflict of interest?**

- Yes  No

**NOTIFICATION OF OTHER PROFESSIONALS**

**A49-1. Will you inform the participants' General Practitioners (and/or any other health or care professional responsible for their care) that they are taking part in the study?**

- Yes  No

*If Yes, please enclose a copy of the information sheet/letter for the GP/health professional with a version number and date.*

**PUBLICATION AND DISSEMINATION**

**A50. Will the research be registered on a public database?**

- Yes  No

*Please give details, or justify if not registering the research.*

The study will not be registered on a public database as no such database exists. The study will be published in a peer reviewed journal.

*Registration of research studies is encouraged wherever possible.*

*You may be able to register your study through your NHS organisation or a register run by a medical research charity, or publish your protocol through an open access publisher. If you are aware of a suitable register or other method of publication, please give details. If not, you may indicate that no suitable register exists. Please ensure that you have entered registry reference number(s) in question A5-1.*

**A51. How do you intend to report and disseminate the results of the study? Tick as appropriate:**

- Peer reviewed scientific journals  
 Internal report

- Conference presentation  
 Publication on website  
 Other publication  
 Submission to regulatory authorities  
 Access to raw data and right to publish freely by all investigators in study or by Independent Steering Committee on behalf of all investigators  
 No plans to report or disseminate the results  
 Other (please specify)  
 The study will form one or more papers to contribute towards the PhD by publication of the CI.

**A53. Will you inform participants of the results?**

Yes  No

*Please give details of how you will inform participants or justify if not doing so.*  
 Information of the study could be disseminated on the NHS Highland public website.

**5. Scientific and Statistical Review**

**A54. How has the scientific quality of the research been assessed? Tick as appropriate:**

- Independent external review  
 Review within a company  
 Review within a multi-centre research group  
 Review within the Chief Investigator's institution or host organisation  
 Review within the research team  
 Review by educational supervisor  
 Other

*Justify and describe the review process and outcome. If the review has been undertaken but not seen by the researcher, give details of the body which has undertaken the review:*

A review of the study proposal was conducted by Iain Atherton, Lecturer, university of Stirling with statistical expertise. Feedback was received via review form and discussion and the proposal was amended in relation to the planned statistical test and explanation of the power calculation within the proposal.

The proposal was also reviewed by Una Lyons, Lead Nurse, NHS Highland who commented on the applicability and timeliness of the study.

Professor Lauder reviewed the proposal in his role as supervisor of the PhD and required clarity on the variable being measured. Again, the proposal was refined to reflect the feedback.

*For all studies except non-doctoral student research, please enclose a copy of any available scientific critique reports, together with any related correspondence.*

*For non-doctoral student research, please enclose a copy of the assessment from your educational supervisor/ institution.*

**A56. How have the statistical aspects of the research been reviewed? Tick as appropriate:**

- Review by independent statistician commissioned by funder or sponsor  
 Other review by independent statistician  
 Review by company statistician  
 Review by a statistician within the Chief Investigator's institution  
 Review by a statistician within the research team or multi-centre group  
 Review by educational supervisor

Other review by individual with relevant statistical expertise

No review necessary as only frequencies and associations will be assessed – details of statistical input not required

*In all cases please give details below of the individual responsible for reviewing the statistical aspects. If advice has been provided in confidence, give details of the department and institution concerned.*

	Title Forename/Initials Surname
	Dr Iain Atherton
Department	School of Nursing, Midwifery and Health
Institution	University of Stirling
Work Address	Centre for Health Science Old Perth Road Inverness
Post Code	IV2 3JH
Telephone	01483255615
Fax	01483255654
Mobile	
E-mail	iain.atherton@stir.ac.uk

*Please enclose a copy of any available comments or reports from a statistician.*

**A57. What is the primary outcome measure for the study?**

Measure of patient satisfaction with quality of care using a likert scale.

**A58. What are the secondary outcome measures? (if any)**

Measurement of perceived empathy (CARE measure) and time waited.

**A59. What is the sample size for the research? How many participants/samples/data records do you plan to study in total? If there is more than one group, please give further details below.**

Total UK sample size: 70  
Total international sample size (including UK):  
Total in European Economic Area:

Further details:

**A60. How was the sample size decided upon? If a formal sample size calculation was used, indicate how this was done, giving sufficient information to justify and reproduce the calculation.**

Empathy will be measured using the CARE measurement tool (Mercer et al 2004). Mercer et al (2005) suggest that at least 50 people be included in any study using this tool to ensure validity and reliability. Using a t-test to determine difference in scores between patients who are satisfied and dissatisfied with their overall care will require a minimum of 64 participants. This number will provide a sample sufficient to detect an effect size of 0.5 at a power of 80% (Machin et al 2008). The calculation was taken from the power calculation tables available in Machin et al 2008. The study therefore aims to collect a minimum of 70 self completed questionnaires.

**A61. Will participants be allocated to groups at random?**

Yes  No

**A62. Please describe the methods of analysis (statistical or other appropriate methods, e.g. for qualitative research) by which the data will be evaluated to meet the study objectives.**

What is the relationship between the patients CARE measure score, time waited and their overall perception of their quality of care in A&E?

Data will be analysed using the statistical package for social science (SPSS) version 17. The research question would be answered by comparing the mean score of the patients overall ratings of satisfaction with their CARE measure score using the t-test.

#### 6. MANAGEMENT OF THE RESEARCH

**A63. Other key investigators/collaborators. Please include all grant co-applicants, protocol co-authors and other key members of the Chief Investigator's team, including non-doctoral student researchers.**

	Title	Forename/Initials	Surname
	Mrs	Beverley	MacLennon
Post	Clinical Educator A&E		
Qualifications			
Employer	NHS Highland		
Work Address	A&E Department Raigmore Hospital Old Perth Rd, Inverness		
Post Code	IV2 3UJ		
Telephone	01463706350		
Fax			
Mobile			
Work Email	beverley.maclennon@nhs.net		

#### A64. Details of research sponsor(s)

##### A64-1. Sponsor

<b>Lead Sponsor</b>	
Status:	Commercial status:
<input type="radio"/> NHS or HSC care organisation	
<input checked="" type="radio"/> Academic	
<input type="radio"/> Pharmaceutical industry	
<input type="radio"/> Medical device industry	
<input type="radio"/> Local Authority	
<input type="radio"/> Other social care provider (including voluntary sector or private organisation)	
<input type="radio"/> Other	
<i>If Other, please specify:</i>	
<b>Contact person</b>	
Name of organisation	University of Stirling
Given name	Carol
Family name	Johnstone



Address University of Stirling  
 Town/city Stirling  
 Post code FK94LA  
 Country UNITED KINGDOM  
 Telephone 01786466690  
 Fax  
 E-mail carol.johnstone@stir.ac.uk

## Is the sponsor based outside the UK?

Yes  No

*Under the Research Governance Framework for Health and Social Care, a sponsor outside the UK must appoint a legal representative established in the UK. Please consult the guidance notes.*

## A65. Has external funding for the research been secured?

- Funding secured from one or more funders  
 External funding application to one or more funders in progress  
 No application for external funding will be made

Please give details of funding applications.

Organisation NHS Highland  
 Address Research and Development Office  
 Centre for Health Science, Old Perth Rd  
 Inverness  
 Post Code IV2 3JH  
 Telephone 01463255280  
 Fax  
 Mobile  
 Email frances.hines@nhs.net

Funding Application Status:  Secured  In progress

Date Funding decision expected: 14/09/2011

Amount: 3,000

Duration

Years: 0

Months: 6

*If applicable, please specify the programme/ funding stream:*

What is the funding stream/ programme for this research project?

What type of research project is this?

- Standalone project  
 Project that is part of a programme grant  
 Project that is part of a fellowship/ personal award/ research training award  
 Other

Other – please state:  
Part of PhD by publication

**A67. Has this or a similar application been previously rejected by a Research Ethics Committee in the UK or another country?**

Yes  No

*Please provide a copy of the unfavourable opinion letter(s). You should explain in your answer to question A6-2 how the reasons for the unfavourable opinion have been addressed in this application.*

**A68. Give details of the lead NHS R&D contact for this research:**

	Title Forename/Initials Surname
	Mrs Frances Hines
Organisation	NHS Highland
Address	Research and Development Office Centre for Health Science, Old Perth Rd Inverness
Post Code	IV2 3JH
Work Email	frances.hines@nhs.net
Telephone	01463255280
Fax	
Mobile	

*Details can be obtained from the NHS R&D Forum website: <http://www.rdforum.nhs.uk>*

**A69-1. How long do you expect the study to last in the UK?**

Planned start date: 16/09/2011  
Planned end date: 01/02/2012  
Total duration:  
Years: 0 Months: 6 Days:

**A71-2. Where will the research take place? (Tick as appropriate)**

- England  
 Scotland  
 Wales  
 Northern Ireland  
 Other countries in European Economic Area

Total UK sites in study 1

**Does this trial involve countries outside the EU?**

Yes  No

**A72. What host organisations (NHS or other) in the UK will be responsible for the research sites? Please indicate the type of organisation by ticking the box and give approximate numbers of planned research sites:**

<input type="checkbox"/> NHS organisations in England	
<input type="checkbox"/> NHS organisations in Wales	
<input checked="" type="checkbox"/> NHS organisations in Scotland	1
<input type="checkbox"/> HSC organisations in Northern Ireland	
<input type="checkbox"/> GP practices in England	
<input type="checkbox"/> GP practices in Wales	
<input type="checkbox"/> GP practices in Scotland	
<input type="checkbox"/> GP practices in Northern Ireland	
<input type="checkbox"/> Social care organisations	
<input type="checkbox"/> Phase 1 trial units	
<input type="checkbox"/> Prison establishments	
<input type="checkbox"/> Probation areas	
<input type="checkbox"/> Independent hospitals	
<input type="checkbox"/> Educational establishments	
<input type="checkbox"/> Independent research units	
<input type="checkbox"/> Other (give details)	
Total UK sites in study:	1

#### A76. Insurance/ indemnity to meet potential legal liabilities

*Note: in this question to NHS indemnity schemes include equivalent schemes provided by Health and Social Care (HSC) in Northern Ireland*

**A76-1. What arrangements will be made for insurance and/or indemnity to meet the potential legal liability of the sponsor(s) for harm to participants arising from the management of the research? Please tick box(es) as applicable.**

*Note: Where a NHS organisation has agreed to act as sponsor or co-sponsor, indemnity is provided through NHS schemes. Indicate if this applies (there is no need to provide documentary evidence). For all other sponsors, please describe the arrangements and provide evidence.*

- NHS indemnity scheme will apply (NHS sponsors only)  
 Other insurance or indemnity arrangements will apply (give details below)

The University of Stirling has indemnity policies to cover the management of research. A letter from R&D will be provided as documentary evidence.

*Please enclose a copy of relevant documents.*

**A76-2. What arrangements will be made for insurance and/ or indemnity to meet the potential legal liability of the sponsor(s) or employer(s) for harm to participants arising from the design of the research? Please tick box(es) as applicable.**

*Note: Where researchers with substantive NHS employment contracts have designed the research, indemnity is provided through NHS schemes. Indicate if this applies (there is no need to provide documentary evidence). For other protocol authors (e.g. company employees, university members), please describe the arrangements and provide evidence.*

- NHS indemnity scheme will apply (protocol authors with NHS contracts only)  
 Other insurance or indemnity arrangements will apply (give details below)

The University of Stirling has indemnity policies to cover the design of the research. A letter from R&D will be provided as documentary evidence.

Please enclose a copy of relevant documents.

**A76-3. What arrangements will be made for insurance and/ or indemnity to meet the potential legal liability of investigators/collaborators arising from harm to participants in the conduct of the research?**

*Note: Where the participants are NHS patients, indemnity is provided through the NHS schemes or through professional indemnity. Indicate if this applies to the whole study (there is no need to provide documentary evidence). Where non-NHS sites are to be included in the research, including private practices, please describe the arrangements which will be made at these sites and provide evidence.*

- NHS indemnity scheme or professional indemnity will apply (participants recruited at NHS sites only)  
 Research includes non-NHS sites (give details of insurance/ indemnity arrangements for these sites below)

Please enclose a copy of relevant documents.

**PART C: Overview of research sites**

Please enter details of the host organisations (Local Authority, NHS or other) in the UK that will be responsible for the research sites. For NHS sites, the host organisation is the Trust or Health Board. Where the research site is a primary care site, e.g. GP practice, please insert the host organisation (PCT or Health Board) in the Institution row and insert the research site (e.g. GP practice) in the Department row.

Research site		Investigator/ Collaborator/ Contact	
Institution name	University of Stirling, School of Nursing, Midwifery & Health	Title	Mrs
Department name	Centre for Health Science	First name/ Initials	Michelle
Street address	Old Perth Road	Surname	Beattie
Town/city	Inverness		
Post Code	IV2 3JH		

**PART D: Declarations****D1. Declaration by Chief Investigator**

1. The information in this form is accurate to the best of my knowledge and belief and I take full responsibility for it.
2. I undertake to abide by the ethical principles underlying the Declaration of Helsinki and good practice guidelines on the proper conduct of research.
3. If the research is approved I undertake to adhere to the study protocol, the terms of the full application as approved and any conditions set out by review bodies in giving approval.
4. I undertake to notify review bodies of substantial amendments to the protocol or the terms of the approved application, and to seek a favourable opinion from the main REC before implementing the amendment.
5. I undertake to submit annual progress reports setting out the progress of the research, as required by review bodies.
6. I am aware of my responsibility to be up to date and comply with the requirements of the law and relevant guidelines relating to security and confidentiality of patient or other personal data, including the need to register when necessary with the appropriate Data Protection Officer. I understand that I am not permitted to disclose identifiable data to third parties unless the disclosure has the consent of the data subject or, in the case of patient data in England and Wales, the disclosure is covered by the terms of an approval under Section 251 of the NHS Act 2006.
7. I understand that research records/data may be subject to inspection by review bodies for audit purposes if required.
8. I understand that any personal data in this application will be held by review bodies and their operational managers and that this will be managed according to the principles established in the Data Protection Act 1998.
9. I understand that the information contained in this application, any supporting documentation and all correspondence with review bodies or their operational managers relating to the application:
  - Will be held by the main REC or the GTAC (as applicable) until at least 3 years after the end of the study; and by NHS R&D offices (where the research requires NHS management permission) in accordance with the NHS Code of Practice on Records Management.
  - May be disclosed to the operational managers of review bodies, or the appointing authority for the main REC, in order to check that the application has been processed correctly or to investigate any complaint.
  - May be seen by auditors appointed to undertake accreditation of RECs.
  - Will be subject to the provisions of the Freedom of Information Acts and may be disclosed in response to requests made under the Acts except where statutory exemptions apply.
10. I understand that information relating to this research, including the contact details on this application, may be held on national research information systems, and that this will be managed according to the principles established in the Data Protection Act 1998.
11. I understand that the summary of this study will be published on the website of the National Research Ethics Service (NRES), together with the contact point for enquiries named below. Publication will take place no earlier than 3 months after issue of the ethics committee's final opinion or the withdrawal of the application.

**Contact point for publication(Not applicable for R&D Forms)**

*NRES would like to include a contact point with the published summary of the study for those wishing to seek further information. We would be grateful if you would indicate one of the contact points below.*

- Chief Investigator  
 Sponsor  
 Study co-ordinator

- Student
- Other – please give details
- None

**Access to application for training purposes (Not applicable for R&D Forms)**

*Optional – please tick as appropriate:*

I would be content for members of other RECs to have access to the information in the application in confidence for training purposes. All personal identifiers and references to sponsors, funders and research units would be removed.

Signature: .....

Print Name: Michelle Beattie

Date: 28/08/2011 (dd/mm/yyyy)

**D2. Declaration by the sponsor's representative**

*If there is more than one sponsor, this declaration should be signed on behalf of the co-sponsors by a representative of the lead sponsor named at A64-1.*

I confirm that:

1. This research proposal has been discussed with the Chief Investigator and agreement in principle to sponsor the research is in place.
2. An appropriate process of scientific critique has demonstrated that this research proposal is worthwhile and of high scientific quality.
3. Any necessary indemnity or insurance arrangements, as described in question A76, will be in place before this research starts. Insurance or indemnity policies will be renewed for the duration of the study where necessary.
4. Arrangements will be in place before the study starts for the research team to access resources and support to deliver the research as proposed.
5. Arrangements to allocate responsibilities for the management, monitoring and reporting of the research will be in place before the research starts.
6. The duties of sponsors set out in the Research Governance Framework for Health and Social Care will be undertaken in relation to this research.
7. I understand that the summary of this study will be published on the website of the National Research Ethics Service (NRES), together with the contact point for enquiries named in this application. Publication will take place no earlier than 3 months after issue of the ethics committee's final opinion or the withdrawal of the application.

This section was signed electronically by Ms Carol Johnstone on 26/08/2011 16:26.

Job Title/Post: Research Development Manager  
Organisation: University of Stirling  
Email: carol.johnstone@stir.ac.uk



**D3. Declaration for student projects by academic supervisor(s)**

1. I have read and approved both the research proposal and this application. I am satisfied that the scientific content of the research is satisfactory for an educational qualification at this level.
2. I undertake to fulfil the responsibilities of the supervisor for this study as set out in the Research Governance Framework for Health and Social Care.
3. I take responsibility for ensuring that this study is conducted in accordance with the ethical principles underlying the Declaration of Helsinki and good practice guidelines on the proper conduct of research, in conjunction with clinical supervisors as appropriate.
4. I take responsibility for ensuring that the applicant is up to date and complies with the requirements of the law and relevant guidelines relating to security and confidentiality of patient and other personal data, in conjunction with clinical supervisors as appropriate.

**Academic supervisor 1**

Signature: .....

Print Name: William Lauder

Post: Professor and Head of School

Organisation: University of Stirling

Date: (dd/mm/yyyy)

Date: 11<sup>th</sup> November 2011  
Your Ref:  
Our Ref: GK/AH

Enquiries to: Amanda Hume  
Extension: 4358  
Direct Line: 01463 704358  
Email: [amanda.hume2@nhs.net](mailto:amanda.hume2@nhs.net)

Dear Sir/Madam

**A Study Comparing Waiting Time & Perceived Empathy with Overall Satisfaction of Care**

Within the Emergency Department we are always looking for ways to improve the service we deliver to you, the patient. This study invites you to help us decide the best way to measure how satisfied you are with the service. Before you decide whether or not to participate it is important for you to understand why the research is being done and what it will involve.

Currently, we mostly measure quality of care within the Emergency Department by the speed in which you have been seen and treated (waiting time). We think that another important aspect of how satisfied you are with your care may be related to how you feel the staff understand and respond to your needs. This study aims to determine whether or not you think that staff understanding is a good measure of how satisfied you are with your care.

Please take time to read the further information carefully and discuss it with others if you wish. Ask if there is anything that is not clear or if you would like more information. I have included the researchers contact details so you can contact her before or after the study should you have any questions.

Michelle Beattie  
Lecturer  
School of Nursing, Midwifery and Health  
University of Stirling  
Highland Campus  
Centre for Health Science  
Old Perth Road  
Inverness  
IV2 3JH

Tel: 01463 255622  
E-mail: [michelle.beattie@stir.ac.uk](mailto:michelle.beattie@stir.ac.uk)

Continued /.....

Version 1 – 11.11.11



**Headquarters:** Assynt House, Beechwood Park, INVERNESS IV2 3BW

Chair: Garry Coutts  
Chief Executive: Elaine Mead



Many thanks for taking the time to read this and consider whether or not you wish to participate in this study.

Yours sincerely

A handwritten signature in black ink, appearing to read 'Gary Kerr', written over a light blue horizontal line.

**Gary Kerr**  
Head of Service of the Emergency Department  
GMC No: 3317051

# Patient Information Leaflet



UNIVERSITY OF  
STIRLING

SCHOOL OF  
NURSING, MIDWIFERY  
AND HEALTH

## **A Study Comparing Waiting Time & Perceived Empathy with Satisfaction of Care**

The Accident and Emergency (A&E) Department are always looking for ways to improve the service they deliver to you, the patient. This study invites you to help us decide the best way to measure how satisfied you are with the service. Before you decide whether or not to participate it is important for you to understand why the research is being done and what it will involve.

### **What is the purpose of the study?**

Currently, the quality of care within A&E is measured by the speed in which you have been seen and treated (waiting time). We think that another important aspect of how satisfied you are with your care may be related to how you feel the staff understand and respond to your needs. This study aims to determine whether or not you think that staff understanding is a good measure of how satisfied you are with your care.

All Accident and Emergency (A&E) Departments in the UK must record the duration of your visit in the Department. This information is used to determine the quality of care you receive. This study seeks to identify if how you feel you were treated in A&E is a better measure of the quality of care you have received in the Department.

### **Why have I been chosen?**

The study aims to ask at least 70 patients (over 18 years) to complete a questionnaire following primary treatment, and before discharge from A&E. Anyone willing to participate and consent is asked to complete the questionnaire.

### **Do I have to take part?**

No. It is up to you to decide whether to take part. If you do decide to take part, you will be asked to sign a consent form. Even if you decide to take part you are still free to withdraw at any time and without giving a reason. Your decision will not affect the standard of care you receive. You will be asked to complete the questionnaire after you have received the necessary care you require.

### **What will happen to me if I take part?**

You will be asked to complete a questionnaire at the end of your A&E visit – either before you go home or before you go to another area for further care. The questionnaire takes about five minutes to complete. Your answers will only be seen by the person conducting the study.

Highland Campus:  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JH  
Tel: +44 (0) 1463 255655  
Fax: +44 (0) 1463 255654

Stirling Campus:  
Stirling  
FK9 4LA  
Tel: +44 (0) 1786 466340  
Fax: +44 (0) 1786 466333

Western Isles Campus:  
Western Isles Hospital  
MacAulay Road  
Stornoway Isle of Lewis HS1 2AF  
Tel: +44 (0) 1851 703243  
Fax: +44 (0) 1851 706070

Version 4: 27 October 2011

The questionnaire aims to measure how understanding you think the staff are and whether this is a good measure of how satisfied you are with your care. You are likely to be seen by more than one member of staff. Try and have one member of staff in your mind when completing the questionnaire. The staff member could be either a doctor or a nurse.

**What are the possible benefits of taking part?**

The study will help us to decide the best way to measure how satisfied you are with the service in A&E.

**Will my taking part in this study be kept confidential?**

All information, which is collected, about you during the course of the research will be kept strictly confidential. Any information about you which leaves the hospital will have your name and address removed so that you cannot be recognised from it.

**What will happen to the results of the research study?**

The responses will be held in a database. The results will be published in an academic health care journal and shared with the Health Board. You will not be identified in any report/publication.

**Who is organising and funding the research?**

The organisation and sponsor of the research is the University of Stirling.

**Who has reviewed the study?**

The study has been reviewed by the Research and Ethics Committee, University of Stirling and NRES Committees – North of Scotland.

**Contacts for Further Information**

Many thanks for considering/participating in this study. If you require any further information on the study you can contact Michelle Beattie, Lecturer, School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Centre for Health Science, Old Perth Road, Inverness, IV2 3JH. Telephone 01463 255622 or e-mail [michelle.beattie@stir.ac.uk](mailto:michelle.beattie@stir.ac.uk)

You can also choose to contact William Lauder, Head of Department, School of Nursing, Midwifery and Health, University of Stirling, Highland Campus, Centre for Health Science, Old Perth Road, Inverness, IV2 3JH. Telephone 01463 255619 or e-mail [william.lauder@stir.ac.uk](mailto:william.lauder@stir.ac.uk)

Highland Campus:  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JH  
Tel: +44 (0) 1463 255655  
Fax: +44 (0) 1463 255654

Stirling Campus:  
Stirling  
FK9 4LA  
Tel: +44 (0) 1786 466340  
Fax: +44 (0) 1786 466333

Western Isles Campus:  
Western Isles Hospital  
MacAulay Road  
Stornoway Isle of Lewis HB1 2AF  
Tel: +44 (0) 1851 703243  
Fax: +44 (0) 1851 706070

Version 4: 27 October 2011

Page 2

# Consent Form



UNIVERSITY OF  
STIRLING

SCHOOL OF  
NURSING, MIDWIFERY  
AND HEALTH

## A Study Comparing Waiting Time & Perceived Empathy with Overall Satisfaction of Care

Name of Researcher: Michelle Beattie

Please initial box

- 1 I confirm that I have read and understand the information sheet dated October 2011 (Version 4) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.
- 2 I understand that my Registration number will be used to obtain information on the length of time I have spent in the Accident and Emergency Department.
- 3 I understand that paperwork from data collected during the study may be looked at by individuals from the University of Stirling, from regulatory authorities or from the NHS Trust/Health Board, for research governance purposes.
- 3 I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected.

\_\_\_\_\_  
Name of Patient

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Name of Person taking consent  
(if different from researcher)

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Michelle Beattie  
Researcher

\_\_\_\_\_  
Date

\_\_\_\_\_  
Signature

Highland Campus:  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JH  
Tel: +44 (0) 1463 255655  
Fax: +44 (0) 1463 255654

Stirling Campus:  
Stirling  
FK9 4LA  
Tel: +44 (0) 1786 466340  
Fax: +44 (0) 1786 466333

Western Isles Campus:  
Western Isles Hospital  
MacAuley Road  
Stornoway, Isle of Lewis H81 2AF  
Tel: +44 (0) 1851 703243  
Fax: +44 (0) 1851 706070

Version 6: 27 October 2011

## The CARE Measure

© Stewart W Mercer 2004

Registration Number: \_\_\_\_\_

<b>1. Please rate the following statements about today's consultation. Please tick one box for each statement and answer every statement.</b>						
<i>How was the doctor/nurse at ...</i>	Poor	Fair	Good	Very Good	Excellent	Does Not Apply
<b>1. Making you feel at ease.....</b> <i>(being friendly and warm towards you, treating you with respect; not cold or abrupt)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>2. Letting you tell your "story".....</b> <i>(giving you time to fully describe your illness in your own words; not interrupting or diverting you)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>3. Really listening .....</b> <i>(paying close attention to what you were saying; not looking at the notes or computer as you were talking)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>4. Being interested in you as a whole person ...</b> <i>(asking/knowing relevant details about your life, your situation; not treating you as "just a number")</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>5. Fully understanding your concerns.....</b> <i>(communicating that he/she had accurately understood your concerns; not overlooking or dismissing anything)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>6. Showing care and compassion....</b> <i>(seeming genuinely concerned, connecting with you on a human level; not being indifferent or "detached")</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>7. Being Positive.....</b> <i>(having a positive approach and a positive attitude; being honest but not negative about your problems)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>8. Explaining things clearly.....</b> <i>(fully answering your questions, explaining clearly, giving you adequate information; not being vague)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>9. Helping you to take control.....</b> <i>(exploring with you what you can do to improve your health yourself; encouraging rather than "lecturing" you)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<b>10. Making a plan of action with you ...</b> <i>(discussing the options, involving you in decisions as much as you want to be involved; not ignoring your views)</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Version 2: 29 August 2011



**UNIVERSITY OF  
STIRLING**

STIRLING FK9 4LA SCOTLAND

Carol Johnstone  
Business Development Manager  
RESEARCH & ENTERPRISE OFFICE  
Tel: 01786 466690  
Fax: 01786 466688  
E-mail: carol.johnstone@stir.ac.uk

19 August 2011

To Whom It May Concern

**Research Study: Time and Perceived Empathy with Overall patient Satisfaction in A&E**

This study is included in the following cover put in place by Aon Ltd. These policies are renewed annually and the current period of insurance is 1 February 2011 - 31 January 2012.

I confirm that the following cover is in place under the Professional Indemnity policy of the University of Stirling. This policy provides indemnity to University of Stirling for legal liability to third parties arising from breach of professional duty due to neglect, error or omission in the course of the business of the University of Stirling.

The limit of the Professional Indemnity cover is £5,000,000 for any one event and in aggregate in any one period of insurance.

In addition the University carries Public Liability cover in respect of its Legal Liability for accidental loss of or damage to Third Party property or for death, injury, illness or disease arising out of the business of the University of Stirling, including liability arising from goods sold or supplied.

The limit of the Public Liability cover is £10,000,000 any one incident and in the aggregate of Products.

I trust that this is sufficient for your requirements. Please however do not hesitate to get in touch with me should you have any queries.

Yours sincerely

Carol Johnstone  
Business Development Manager

---



Recruitment & Employment Services Section  
Human Resource Services  
Raigmore Hospital, Old Perth Road  
Inverness IV2 3UJ  
Telephone 01463 704000  
Fax 01463 704598  
Textphone users can contact us via  
Typetalk: Tel 0800 959598  
[www.nhshighland.scot.nhs.uk](http://www.nhshighland.scot.nhs.uk)



Mrs Michelle Beattie  
Lecturer  
School of Nursing, Midwifery & Health  
University of Stirling  
Highland Campus  
Centre for Health Science  
Old Perth Road  
INVERNESS  
IV2 3JH

Date 19 February 2016  
Your Ref  
Our Ref  
Enquiries to  
Extension 5914  
Direct Line 01463 705914  
Email [carol.mackay@nhs.net](mailto:carol.mackay@nhs.net)

Dear Mrs Beattie

**Letter of Access for Research**

**Project Title: Are We Measuring the Correct Dimensions of Quality in A&E? A Study Comparing Time Waited and Perceived Empathy with Overall Patient Satisfaction**

This letter confirms your right of access to conduct research through NHS Highland for the purpose and on the terms and conditions set out below. This right of access commences on 18/11/11 and ends on 30/04/12 unless terminated earlier in accordance with the clauses below.

You have a right of access to conduct such research as confirmed in writing in the letter of permission for research from this NHS organisation. Please note that you cannot start the research until the Principal Investigator for the research project has received a letter from us giving permission to conduct the project.

The information supplied about your role in research at NHS Highland has been reviewed and you do not require an honorary research contract with this NHS organisation. We are satisfied that such pre-engagement checks as we consider necessary have been carried out.

You are considered to be a legal visitor to NHS Highland premises. You are not entitled to any form of payment or access to other benefits provided by this NHS organisation to employees and this letter does not give rise to any other relationship between you and this NHS organisation, in particular that of an employee.

While undertaking research through NHS Highland, you will remain accountable to your employer the **University of Stirling** but you are required to follow the reasonable instructions of **Dr Ken Proctor Associate Medical Director** in this NHS organisation or those given on her/his behalf in relation to the terms of this right of access.

Where any third party claim is made, whether or not legal proceedings are issued, arising out of or in connection with your right of access, you are required to co-operate fully with any investigation by this NHS organisation in connection with any such claim and to give all such assistance as may reasonably be required regarding the conduct of any legal proceedings.

You must act in accordance with NHS Highland policies and procedures, which are available to you upon request, and the Research Governance Framework.



Headquarters: ~~Assynt House, Beechwood Park~~, INVERNESS IV2 3BW

Chair: Garry Coutts  
Chief Executive: Elaine Mead

You are required to co-operate with NHS Highland in discharging its duties under the Health and Safety at Work etc Act 1974 and other health and safety legislation and to take reasonable care for the health and safety of yourself and others while on NHS Highland premises. You must observe the same standards of care and propriety in dealing with patients, staff, visitors, equipment and premises as is expected of any other contract holder and you must act appropriately, responsibly and professionally at all times.

You are required to ensure that all information regarding patients or staff remains secure and *strictly confidential* at all times. You must ensure that you understand and comply with the requirements of the NHS Confidentiality Code of Practice (<http://www.dh.gov.uk/assetRoot/04/06/92/54/04069254.pdf>) and the Data Protection Act 1998. Furthermore you should be aware that under the Act, unauthorised disclosure of information is an offence and such disclosures may lead to prosecution.

You should ensure that, where you are issued with an identity or security card, a bleep number, email or library account, keys or protective clothing, these are returned upon termination of this arrangement. Please also ensure that while on the premises you wear your ID badge at all times, or are able to prove your identity if challenged. Please note that this NHS organisation accepts no responsibility for damage to or loss of personal property.

We may terminate your right to attend at any time either by giving seven days' written notice to you or immediately without any notice if you are in breach of any of the terms or conditions described in this letter or if you commit any act that we reasonably consider to amount to serious misconduct or to be disruptive and/or prejudicial to the interests and/or business of this NHS organisation or if you are convicted of any criminal offence. Your substantive employer is responsible for your conduct during this research project and may in the circumstances described above instigate disciplinary action against you.

NHS Highland will not indemnify you against any liability incurred as a result of any breach of confidentiality or breach of the Data Protection Act 1998. Any breach of the Data Protection Act 1998 may result in legal action against you and/or your substantive employer.

If your current role or involvement in research changes, or any of the information provided in your Research Passport changes, you must inform your employer through their normal procedures. You must also inform your nominated manager in this NHS organisation.

Yours sincerely

John Huband  
Head of Recruitment and Employment Services

Cc Frances Hines, R&D Manager, NHS Highland, Room S101, Centre for Health Science, Old Perth Road, Inverness, IV2 3JH  
Director, Human Resources Services, Cottrell Building (Room 4B1), University of Stirling, Stirling FK9 4LA

Please sign and date BOTH copies of this Letter of Access and return ONE copy to:

Frances Hines, R&D Manager, NHS Highland, Room S101, Centre for Health Science, Old Perth Road, Inverness, IV2 3JH

REF: MB3/REC11-NS/0025



UNIVERSITY OF  
STIRLING

7<sup>th</sup> November 2011

SCHOOL OF  
NURSING, MIDWIFERY  
AND HEALTH

Email: [nursingmidwifery@stir.ac.uk](mailto:nursingmidwifery@stir.ac.uk)  
Web: [www.stir.ac.uk/nmhealth](http://www.stir.ac.uk/nmhealth)

**Michelle Beattie**  
Lecturer  
School of Nursing, Midwifery and Health  
University of Stirling  
Highland Campus  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JH

Tel: +44 (0) 1463 255622  
Fax: +44 (0) 1463 255634  
Email: [michelle.beattie@stir.ac.uk](mailto:michelle.beattie@stir.ac.uk)

Dear Dr Venables

Time and Perceived Empathy with Overall Patient Satisfaction in A&E

Many thanks for your timely response and request for clarification from my response sent on 14<sup>th</sup> October 2011. I have detailed the response to your points below.

*The Committee suggest that a separate Letter of Invitation is written and this should come from the Head of Department and contain the researchers contact details and suggest that it could be stapled to the front of the Participant Information Sheet.*

A separate Letter of Invitation has been endorsed by Gary Kerr, Head of Accident and Emergency Department, Raigmore Hospital, Inverness. The letter contains details of my contact details (see Letter Ref: GK/AH, 4<sup>th</sup> November, 2011).

*The Committee ask that you re-review the Information Sheet to ensure that the language is appropriate for your target audience.*

The Information Sheet has been re-written to reflect the target audience (see Patient Information Leaflet, Version 4, 27<sup>th</sup> October 2011). The Consent Form has also been revised to refer to the most recent version of the Patient Information Leaflet (see Consent Form, Version 5, 27<sup>th</sup> October 2011).

I hope these additions will suffice, however please contact me should you require any further information.

Yours sincerely

Michelle Beattie  
Lecturer

Highland Campus:  
Centre for Health Science  
Old Perth Road  
Inverness IV2 3JH  
Tel: +44 (0) 1463 255665  
Fax: +44 (0) 1463 255654

Stirling Campus:  
Stirling  
FK9 4LA  
Tel: +44 (0) 1796 466340  
Fax: +44 (0) 1796 466333

Western Isles Campus:  
Western Isles Hospital  
MacAulay Road  
Stomoway Isle of Lewis HS1 2AF  
Tel: +44 (0) 1851 708243  
Fax: +44 (0) 1851 706070

The University of Stirling is recognised as a Scottish Charity with number SC 011159

**NRES Committees - North of Scotland**

Summerfield House  
2 Eday Road  
Aberdeen  
AB15 6RE

Telephone: 01224 558474  
Facsimile: 01224 558609  
Email: nosres@nhs.net



14 November 2011

Mrs Michelle Beattie  
Lecturer  
University of Stirling  
Centre for Health Science  
Old Perth Road  
Inverness  
IV2 3JH

Dear Mrs Beattie

**Study title:** Are We Measuring the Correct Dimensions of Quality in A&E? A Study Comparing Time Waited and Perceived Empathy with Overall Patient Satisfaction  
**REC reference:** 11/NS/0025

Thank you for your letter of 14 November 2011 responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information was considered by the Ethics Coordinator.

**Confirmation of ethical opinion**

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

**Ethical review of research sites**

**NHS sites**

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

**Non-NHS sites**

**Conditions of the favourable opinion**

The favourable opinion is subject to the following conditions being met prior to the start of the study.

**Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.**

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

**It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).**

**Approved documents**

The final list of documents reviewed and approved by the Committee is as follows:

Document	Version	Date
(None)		14 November 2011
Covering Letter		14 November 2011
Evidence of insurance or indemnity		19 January 2011
Investigator CV		22 August 2011
Letter of invitation to participant	1	11 November 2011
Other: CV - Beverley MacLennan		25 August 2011
Other: CV - Professor William Lauder		29 August 2011
Other: Review from University of Stirling Ethics Committee		25 July 2011
Other: The CARE Measure	2	29 September 2011
Other: Letter of Confirmation of Sponsor		19 August 2011
Other: NHS SSI Form		
Participant Consent Form	5	27 October 2011
Participant Information Sheet	4	27 October 2011
Protocol	2	26 August 2011
REC application	3.3	08 September 2011
Referees or other scientific critique report		18 August 2011
Response to Request for Further Information		

**Statement of compliance**

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

**After ethical review**

Reporting requirements

The attached document "*After ethical review – guidance for researchers*" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

Further information is available at National Research Ethics Service website > After Review

<b>11/NS/0025</b>	<b>Please quote this number on all correspondence</b>
-------------------	---

With the Committee's best wishes for the success of this project

Yours sincerely

pp

*RS Bhattacharya*

**Professor Siladitya Bhattacharya  
Chair**

Enclosures:            After ethical review – guidance for researchers]

Copy to:                Mrs Carol Johnstone  
                              NHS Highland R&D Department

## Appendix 5: Care Experience Feedback Improvement Tool (CEFIT)

Version 1

The five statements below have been identified as important aspects of care by patients. We would welcome your feedback on how often you experienced these aspects of quality whilst receiving care. Examples are given to help describe each of the statements. You may have different examples or interpretations of the statements, which is OK.

Please tick one box like this  with a ball point pen. If you change your mind just cross out your old response and make a new choice. Please answer every statement.

	Never <sup>1</sup>	Occasionally <sup>2</sup>	Sometimes <sup>3</sup>	Often <sup>4</sup>	Always <sup>5</sup>
<b>1. I received safe care...</b> <input type="checkbox"/> Never <sup>1</sup> <input type="checkbox"/> Occasionally <sup>2</sup> <input type="checkbox"/> Sometimes <sup>3</sup> <input type="checkbox"/> Often <sup>4</sup> <input type="checkbox"/> Always <sup>5</sup> <i>Examples might include:</i> Staff washed their hands before and after any direct contact with me I received the right medication dose					
<b>2. I received timely care...</b> <input type="checkbox"/> Never <sup>1</sup> <input type="checkbox"/> Occasionally <sup>2</sup> <input type="checkbox"/> Sometimes <sup>3</sup> <input type="checkbox"/> Often <sup>4</sup> <input type="checkbox"/> Always <sup>5</sup> <i>Examples might include:</i> I waited an acceptable amount of time to be seen I received procedures and treatments within a reasonable time					
<b>3. My care met my personal needs...</b> <input type="checkbox"/> Never <sup>1</sup> <input type="checkbox"/> Occasionally <sup>2</sup> <input type="checkbox"/> Sometimes <sup>3</sup> <input type="checkbox"/> Often <sup>4</sup> <input type="checkbox"/> Always <sup>5</sup> <i>Examples might include:</i> My care was tailored to my individual needs i.e. adjustments made for my other conditions and personal lifestyle I was involved in all decisions about my care					
<b>4. Staff were caring to me ...</b> <input type="checkbox"/> Never <sup>1</sup> <input type="checkbox"/> Occasionally <sup>2</sup> <input type="checkbox"/> Sometimes <sup>3</sup> <input type="checkbox"/> Often <sup>4</sup> <input type="checkbox"/> Always <sup>5</sup> <i>Examples might include:</i> I was treated pleasantly and with courtesy Staff really listened to me and took my concerns seriously					
<b>5. I was able to get the care I needed...</b> <input type="checkbox"/> Never <sup>1</sup> <input type="checkbox"/> Occasionally <sup>2</sup> <input type="checkbox"/> Sometimes <sup>3</sup> <input type="checkbox"/> Often <sup>4</sup> <input type="checkbox"/> Always <sup>5</sup> <i>Examples might include:</i> I knew how to access the care I needed There were no barriers to me getting the treatment/services that I needed					

## Appendix 6: Patient Feedback Tool Ethical Decision Letter

Frances Hines  
Research, Development & Innovation Manager  
NHS Highland R, D & I Department  
The Centre for Health Science  
Old Perth Road  
Inverness  
IV2 3JH



E-mail: [frances.hines@nhs.net](mailto:frances.hines@nhs.net)

Tel: 01463 255822

Fax: 01463 255838

[www.show.scot.nhs.uk/nhshighland/](http://www.show.scot.nhs.uk/nhshighland/)

22 January 2016

Ms Michelle Beattie  
Lecturer  
School of Health Science  
University of Stirling  
Highland Campus  
Centre for Health Science  
Old Perth Road  
Inverness  
IV2 3JH

### RE: Query Regarding Requirement for NHS Ethical Approval of Feedback Tool as Part of Initial Project Development

Dear Ms Beattie,

Following our previous discussions, I thought it would be sensible for me to outline what NHS Highland's view is regarding the Feedback Tool development and whether it required NHS Research Ethics Committee approval.

I did give you advice regarding this position at the time when you were in the process of developing the Feedback Tool, and my view and position have not changed from this time. Essentially, the initial feedback was gained through a standard patient feedback on service improvement process not a research process, and was therefore considered to be of the same status as any audit or service evaluation that patients (and staff) regularly take part in. At this time there was no plan to use this feedback as part of a research project – it was done as general engagement with service users and staff providers. It was on this basis that I decided that you did not require REC approval.

Later on, the information from this process was revisited and it was then decided that it might be useful to develop this aspect of your PhD further. To that end, the data was then employed in a research project. However, from my perspective this was no different from any individual (NHS or University) who having gained information during an audit or service evaluation at some time in the past, then deciding to use this already collected data (which did not require REC approval) to form a component of a new research project (which in itself would be submitted for REC approval).

So, as previously I do not consider that there is a need for NHS REC approval and I believe that this corresponds with the view from the North of Scotland REC, who also felt that it would not require REC approval if it was part of a service evaluation.



Headquarters: Assynt House, Beechwood Park, INVERNESS IV2 3BW

Chair: David Alston

Chief Executive: Elaine Mead





Please do not hesitate to contact me if you need to discuss this further, otherwise I hope that your research continues to progress smoothly.

Yours sincerely,

A handwritten signature in black ink that reads 'Frances Hines'.

Frances Hines  
R, D & I Manager  
NHS Highland

## Appendix 7: Expert Feedback Form: Content Validity Index

### Expert Feedback: Content Validity Index

**Please rate the individual items of the quality of care feedback tool using the rating scale below.**

1=not relevant  
 2=unable to assess relevance without item revision or item is in need of such revision that it would no longer be relevant  
 3=relevant but needs minor alterations  
 4=very relevant and succinct

Please document any suggestions for improvement below. This would include any suggestions for re-wording, or revision, or inclusion or removal of an item.

Item	Descriptor	Rating Score 1-4	Suggestions for improvement
1. Safety	<i>I received safe care ...</i>		
2. Timely	<i>I received timely care ...</i>		
3. Effective	<i>My care met my personal needs ...</i>		
4. Caring	<i>Staff were caring to me ...</i>		
5. System Navigation	<i>I was able to get the care I needed ...</i>		
6. Global Question	<i>Overall rating of experience</i>		

## Appendix 8: CEFIT Results of COSMIN Checklist for Content Validity

COSMIN Questions for Content Validity	Response	Rating
1. Was there an assessment of whether all items refer to relevant aspects of the construct to be measured?	Yes	Excellent
2. Was there an assessment of whether all items are relevant for the study population? (e.g. age, gender, disease characteristics, country, setting)	No	Fair
3. Was there an assessment of whether all items are relevant for the purpose of the measurement instrument? (discriminative, evaluative, and/or predictive)	Partial	Fair
4. Was there <u>an</u> assessment of whether all items together comprehensively reflect the construct to be measured?	Yes	Excellent
5. Were there any important flaws in the design or methods of the study?	No	Fair
6. Was there evidence that items were theoretically informed?	Yes	Excellent
<b>TOTAL RATING</b>	<b>Lowest score counts</b>	<b>FAIR</b>

## Appendix 9: CEFIT Results of COSMIN Checklist for Structural Validity

COSMIN Questions for Structural Validity	Response	Rating
1. Does the scale consist of effect indicators, i.e. is it based on a reflective model?	Yes	Excellent
2. Was the percentage of missing items given?	Yes	Excellent
3. Was there a description of how missing items were handled?	Not necessary as none	Excellent
4. Was the sample size included in the analysis adequate?	Yes	Excellent
5. Were there any important flaws in the design or methods of the study?	No	Excellent
6. for CTT: Was exploratory or confirmatory factor analysis performed?	Yes	Excellent
7. for IRT: Were IRT tests for determining the (uni-) dimensionality of the items performed?	N/A	N/A
<b>TOTAL RATING</b>	<b>Lowest score counts</b>	<b>EXCELLENT</b>

## Appendix 10: CEFIT Results of COSMIN Checklist for Internal Consistency

COSMIN Questions for Internal Consistency	Response	Rating
1. Does the scale consist of effect indicators, i.e. is it based on a reflective model?	Yes	Excellent
2. Was the percentage of missing items given?	Yes	Excellent
3. Was there a description of how missing items were handled?	Not necessary as none	Excellent
4. Was the sample size included in the internal consistency analysis adequate?	Yes	Excellent
5. Was the <u>unidimensionality</u> of the scale checked? i.e. was factor analysis or IRT model applied?	Yes	Excellent
6. Was the sample size included in the <u>unidimensionality</u> analysis adequate?	Yes	Excellent
7. Was an internal consistency statistic calculated for each (unidimensional) (sub)scale separately?	N/A	N/A
8. Were there any important flaws in the design or methods of the study?	No	Excellent
9. for Classical Test Theory (CTT): Was Cronbach's alpha calculated?	Yes	Excellent
10. for dichotomous scores: Was Cronbach's alpha or KR-20 calculated?	N/A	N/A
11. for IRT: Was a goodness of fit statistic at a global level calculated? E.g. $\chi^2$ , reliability coefficient of estimated latent trait value (index of (subject or item)	N/A	N/A
<b>TOTAL RATING</b>	<b>Lowest score counts</b>	<b>GOOD</b>

## Appendix 11: Quality Criteria for Measurement Properties (Terwee et al 2007)

Property	Rating	Quality Criteria
<i>Reliability</i>		
Internal consistency	(+)	(Sub)scale unidimensional AND Cronbach's alpha(s) $\geq 0.70$
	?	Dimensionality not known OR Cronbach's alpha not determined
	(-)	(Sub)scale not unidimensional OR Cronbach's alpha(s) $< 0.70$
Measurement error	(+)	MIC $>$ SDC OR MIC outside the LOA
	?	MIC not defined
	(-)	MIC $<$ SDC OR MIC equals or inside LOA
Reliability	(+)	ICC/Weighted Kappa $\geq 0.70$ OR Pearson's $r \geq 0.80$
	?	Neither ICC/weighted Kappa, nor Pearson's $r$ determined
	(-)	ICC/Weighted Kappa $< 0.70$ OR Pearson's $r < 0.80$
<i>Validity</i>		
Content validity	(+)	The target population considers all items in the questionnaire to be relevant AND considers the questionnaire to be complete
	?	No target population involvement
	(-)	The target population considers all items in the questionnaire to be irrelevant OR considers the questionnaire to be incomplete
<i>Construct validity</i>		
Structural validity	(+)	Factors should explain at least 50% of the variance
	?	Explained variance not mentioned
	(-)	Factors explain $< 50\%$ of the variance
Hypothesis testing	(+)	Correlation with an instrument measuring the same construct $\geq 50\%$ OR at least 75% of the results are in accordance with the <u>hypotheses</u> ) AND correlation with related constructs is higher than with unrelated constructs
	?	<u>Solely</u> correlations determined with unrelated constructs
	(-)	Correlation with an instrument measuring the same construct $< 50\%$ OR $< 75\%$ of the results are in accordance with the hypotheses OR correlation with related constructs is lower than with unrelated constructs

+ positive  
 - negative  
 ? indeterminate

MIC minimal important change  
 ICC intraclass correlation  
 SDC smallest detectable change

## Appendix 12: Additional Aspects of Utility Scoring Criteria and CEFIT Results

Questions for <u>Cost Efficiency</u>	Explanation/Evidence	Rating
1. What are the number of assessments (observations, questionnaires) needed to ensure reliable data?	Unknown	Poor
2. How long does an assessment take to complete	<15 mins	Excellent
3. What are the administrative costs of completing the assessment?	Some administrative resource but no specialist resource required.	Good
4. What is the cost to complete a reliable sample?	Moderate	Good
<b>TOTAL RATING</b>	<b>Average rating of all responses</b>	<b>GOOD</b>
<b>Questions for <u>Acceptability</u></b>		
1. Is there evidence of subjects understanding of the instrument/assessment?	Yes	Excellent
2. How many assessments are not completed?	No missing items and 100% response rate	Excellent
3. Has the instrument/assessment been tested in an appropriate context?	No, not yet	Poor
<b>TOTAL RATING</b>	<b>Lowest score counts</b>	<b>POOR</b>
<b>Questions for <u>Educational Impact</u></b>		
1. There is evidence of the instruments intended purpose being achieved (i.e. if aim is to enable hospital ranking for patient selection, is there evidence that the results are actually influencing patient choice?)	Explanatory and theoretical link between intended and actual QJ use, but no clear evidence.	Good
2. The scoring system is easily translated or available in an easy to use format?	Explicitly stated and easy to use	Excellent
3. The feedback from the results can be readily used for action where necessary?	Feedback is readily available in a format that enables necessary action.	Excellent
<b>TOTAL RATING</b>	<b>Lowest score counts</b>	<b>GOOD</b>

## Appendix 13: Approval to Use Published Paper One in Thesis

RE: Publication Re-Use



PermissionsUK <Permissions@sagepub.co.uk>

Fri 21/11/2014 15:23

To: Michelle Beattie;

• You replied on 21/11/2014 15:25.

Dear Michelle,

Thank you for your email.

You are welcome to include your article as part of your PHD thesis. We only have the article in a PDF format.

Best Wishes,

Ellie Hodge  
Permissions Assistant  
SAGE Publications Ltd  
1 Oliver's Yard, 55 City Road  
London, EC1Y 1SP  
UK

[www.sagepub.co.uk](http://www.sagepub.co.uk)

SAGE Publications Ltd, Registered in England No.1017514

Los Angeles | London | New Delhi

Singapore | Washington DC

The natural home for authors, editors & societies

Thank you for considering the environment before printing this email.



## Appendix 14: Approval to use Paper Two in Thesis



THE INTERNATIONAL JOURNAL OF  
PERSON CENTERED MEDICINE

### The International Journal of Person Centered Medicine (IJPCM)

#### License to Publish

Title of Article: Compassion or speed, which is a more accurate indicator of healthcare quality in the emergency department from the patients' perspective.....

Name of Corresponding Author: Michelle Beattie RN BSc (Hons) MSc  
.....

Name of Copyright Owner (if not Author's).....

Address of Copyright Owner: School of Nursing, Midwifery and Health, University of Sterling, Highland Campus, Centre for Health Science, Old Perth Road, Inverness, IV2 3JH.  
.....  
.....

By signing this agreement you as Corresponding Author and acting on behalf of any contributing author(s) or other copyright owner(s) agree to grant to the University of Buckingham Press (UBPL) the exclusive right to both reproduce and/or distribute your article (including the abstract) throughout the world in printed, electronic or any other medium, and in turn to authorise others (including Reproduction Rights Organisations such as the Copyright Licensing Agency and the Copyright Clearance Centre) to do the same.

You agree that UBPL may publish your article in the IJPCM and that we may sell or distribute it within the journal, on its own, or with other related material.

You warrant that you have gained the written agreement of all Authors to act on their behalf.

You warrant that the article is your, and your co-authors', original work, has not been previously published, and is not currently under consideration by any other journal. If it contains material which is someone else's copyright you promise that you have obtained unrestricted permission of the copyright owner (please attach evidence of this) and that the material is clearly identified and acknowledged in the text. You also warrant that the article does not, to the best of your knowledge, contain anything which is libellous, illegal, or infringes anyone's copyright or other rights.

You assert your Moral Right to be identified as the author(s) and we promise that we will respect your rights as the author(s). That is we will make sure your name(s) is/are always clearly associated with the article and, while you do allow us to make necessary editorial changes, we will not make any substantial alteration to your article without consulting you. Your article will be fairly reviewed by two referees, and UBPL will inform you of the outcome as quickly as possible. When the article is accepted, we undertake to publish it as soon as we reasonably can. If it is rejected, this agreement is automatically cancelled and all rights revert to you. On publication we will send you access to the electronic version to enable you to print

copies. UBPL will make all reasonable efforts to maximise the visibility of the journal and your article within it.

Copyright remains yours, and UBPL will acknowledge this in the copyright line which appears on your article. However, you authorise UBPL to act on your behalf to defend your copyright if anyone should infringe it and to retain half of any damages, after deducting costs. You also retain the right to use your article (provided you acknowledge the published original in standard bibliographic citation form) in the following ways, as long as you do not sell it or give it away in ways that would conflict with our commercial business interests. You are free to use your article for the internal educational or other purposes of your own institution or company; mounted on your own or your institutions website; posted to free public servers of preprints and or article in your subject area; or in whole or in part, as the basis for your own further publications or spoken presentations.

Signature of Corresponding Author..... *Ann Beattie* ..... Date *3 Sept 12*

Signature of Copyright Owner, if different

..... Date .....

## Appendix 15: Approval to Use Published Papers Three and Four in Thesis

Fri 04/03/2016 09:20  
Systematic Reviews Editorial <editorial@systematicreviewsjournal.com>  
RE: Use of Publication

To: Michelle Beattie  
You replied to this message on 04/03/2016 10:18.

Action Items [+ Get more ap](#)

Dear Ms Beattie,

Thank you for your email. Please feel free to include the two papers below in your PhD publication, all we would ask is that you cite them correctly.

Best wishes,

Rebecca Kirk  
Editorial Office Manager

**Biomed Central**  
Floor 6, 236 Gray's Inn Road  
London  
WC1X 8HB

email: [rebecca.kirk@biomedcentral.com](mailto:rebecca.kirk@biomedcentral.com)  
web: [MailScanner has detected a possible fraud attempt from "... claiming to be www.biomedcentral.com](http://www.biomedcentral.com)

---

**From:** Michelle Beattie [<mailto:michelle.beattie@stir.ac.uk>]  
**Sent:** 03 March 2016 11:18  
**To:** Systematic Reviews Editorial  
**Subject:** Use of Publication

Dear Colleague

I am writing to request permission to include the two papers below in my PhD by publication

**Beattie M, Lauder W, Atherton I, Murphy D. (2014)** Instruments to measure patient experience of healthcare quality in hospitals: a systematic review protocol *Systematic Reviews Journal.* 3; 4.

**Beattie M, Murphy D, Atherton I, Lauder W. (2015)** Instruments to measure patient experience of healthcare quality in hospitals: a systematic review. *Systematic Reviews Journal.* 4; 97.

Please could you let me know of any action I need to take.

