Creating learning environments for compassionate care (CLECC): a feasibility study

Bridges, Jackie, Pickering, Ruth M., Barker, Hannah, Chable, Rosemary, Fuller, Alison, Gould, Lisa, Libberton, Paula, Mesa-Eguiagaray, Ines, Raftery, James, Sayer, Avan Aihie, Westwood, Greta, Wigley, Wendy ORCID: https://orcid.org/0000-0002-7778-9818, Yao, Guiqing, Zhu, Shihua and Griffiths, Peter (2018) Creating learning environments for compassionate care (CLECC): a feasibility study. Health Services and Delivery Research, 6 (33). ISSN 2050-4349

This is the Published Version of the final output.

UWL repository link: https://repository.uwl.ac.uk/id/eprint/4161/

Alternative formats: If you require this document in an alternative format, please contact: open.research@uwl.ac.uk
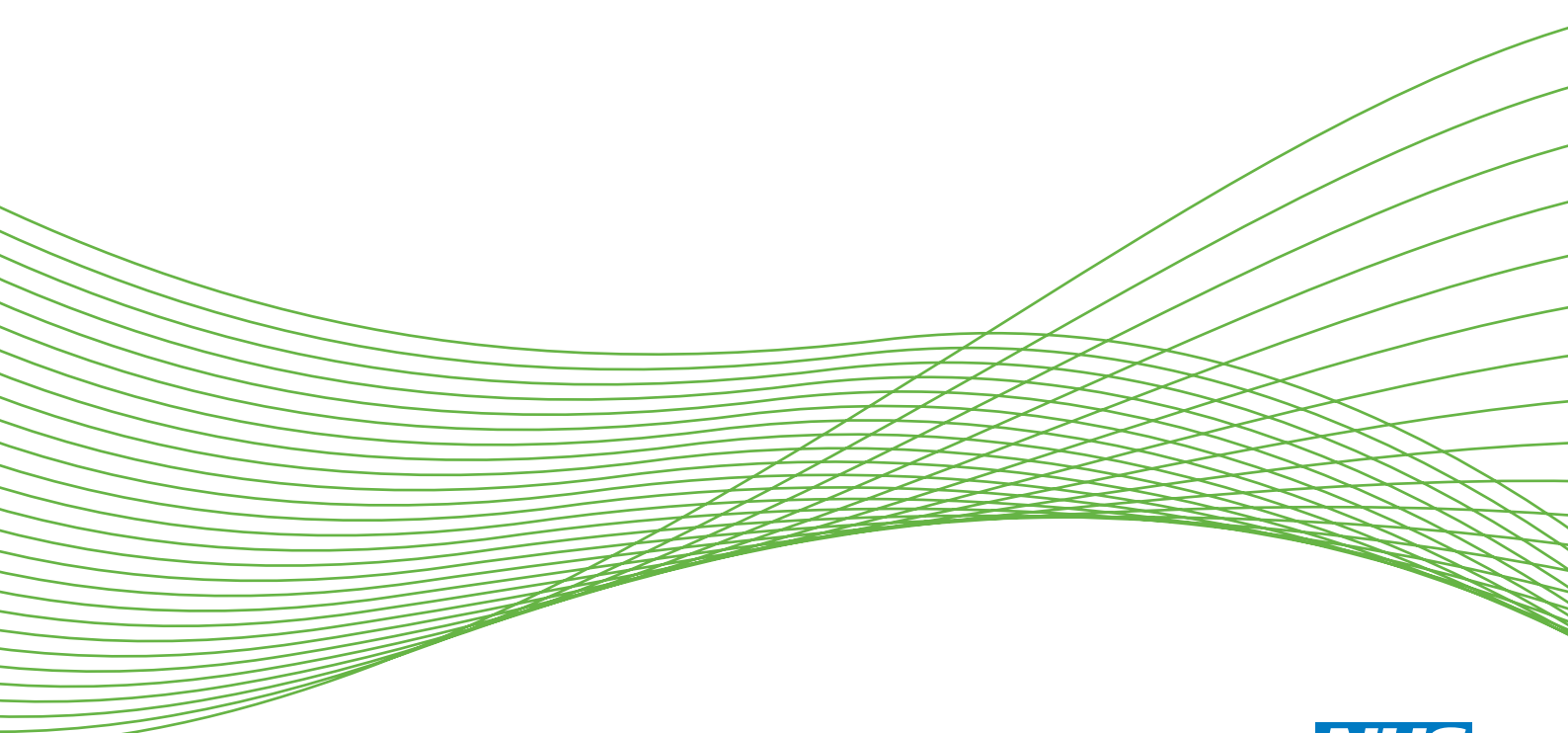
Implementing the Creating Learning Environments for Compassionate Care (CLECC) programme in acute hospital settings: a pilot RCT and feasibility study

Jackie Bridges, Ruth M Pickering, Hannah Barker, Rosemary Chable, Alison Fuller, Lisa Gould, Paula Libberton, Ines Mesa-Eguiagaray, James Raftery, Avan Aihie Sayer, Greta Westwood, Wendy Wigley, Guiqing Yao, Shihua Zhu and Peter Griffiths

NHS
National Institute for Health Research

Implementing the Creating Learning Environments for Compassionate Care (CLECC) programme in acute hospital settings: a pilot RCT and feasibility study

Jackie Bridges,[1,2]* Ruth M Pickering,[1] Hannah Barker,[1]
Rosemary Chable,[2,3] Alison Fuller,[4] Lisa Gould,[1]
Paula Libberton,[1] Ines Mesa-Eguiagaray,[1]
James Raftery,[1] Avan Aihie Sayer,[2,5,6,7]
Greta Westwood,[1,2,8] Wendy Wigley,[1] Guiqing Yao,[1]
Shihua Zhu[1] and Peter Griffiths[1,2]

[1]Faculty of Health Sciences, University of Southampton, Southampton, UK
[2]NIHR Collaboration for Leadership in Applied Health Research and Care
  (CLAHRC) Wessex, Southampton, UK
[3]Training, Development & Workforce, University Hospitals Southampton NHS
  Foundation Trust, Southampton, UK
[4]Institute of Education, University College London, London, UK
[5]NIHR Newcastle Biomedical Research Centre, Newcastle, UK
[6]Older People's Medicine, Newcastle upon Tyne Hospitals NHS Foundation Trust,
  Newcastle, UK
[7]Faculty of Medical Sciences, Newcastle University, Newcastle, UK
[8]Research and Innovation, Portsmouth Hospitals NHS Trust, Portsmouth, UK

*Corresponding author

This report should be referenced as follows:

Bridges J, Pickering RM, Barker H, Chable R, Fuller A, Gould L, *et al*. Implementing the Creating
Learning Environments for Compassionate Care (CLECC) programme in acute hospital settings:
a pilot RCT and feasibility study. *Health Serv Deliv Res* 2018;**6**(33).

Health Services and Delivery Research

Criteria for inclusion in the *Health Services and Delivery Research* journal
Reports are published in *Health Services and Delivery Research* (HS&DR) if (1) they have resulted from work for the HS&DR programme or programmes which preceded the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HS&DR programme

The Health Services and Delivery Research (HS&DR) programme, part of the National Institute for Health Research (NIHR), was established to fund a broad range of research. It combines the strengths and contributions of two previous NIHR research programmes: the Health Services Research (HSR) programme and the Service Delivery and Organisation (SDO) programme, which were merged in January 2012.

The HS&DR programme aims to produce rigorous and relevant evidence on the quality, access and organisation of health services including costs and outcomes, as well as research on implementation. The programme will enhance the strategic focus on research that matters to the NHS and is keen to support ambitious evaluative research to improve health services.

For more information about the HS&DR programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hsdr

This report

Abstract

Implementing the Creating Learning Environments for Compassionate Care (CLECC) programme in acute hospital settings: a pilot RCT and feasibility study

Jackie Bridges,[1,2]* Ruth M Pickering,[1] Hannah Barker,[1] Rosemary Chable,[2,3] Alison Fuller,[4] Lisa Gould,[1] Paula Libberton,[1] Ines Mesa-Eguiagaray,[1] James Raftery,[1] Avan Aihie Sayer,[2,5,6,7] Greta Westwood,[1,2,8] Wendy Wigley,[1] Guiqing Yao,[1] Shihua Zhu[1] and Peter Griffiths[1,2]

[1]Faculty of Health Sciences, University of Southampton, Southampton, UK
[2]NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC) Wessex, Southampton, UK
[3]Training, Development & Workforce, University Hospitals Southampton NHS Foundation Trust, Southampton, UK
[4]Institute of Education, University College London, London, UK
[5]NIHR Newcastle Biomedical Research Centre, Newcastle, UK
[6]Older People's Medicine, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle, UK
[7]Faculty of Medical Sciences, Newcastle University, Newcastle, UK
[8]Research and Innovation, Portsmouth Hospitals NHS Trust, Portsmouth, UK

*Corresponding author　Jackie.Bridges@soton.ac.uk

Background: Concerns about the degree of compassion in health care have become a focus for national and international attention. However, existing research on compassionate care interventions provides scant evidence of effectiveness or the contexts in which effectiveness is achievable.

Objectives: To assess the feasibility of implementing the Creating Learning Environments for Compassionate Care (CLECC) programme in acute hospital settings and to evaluate its impact on patient care.

Design: Pilot cluster randomised trial (CRT) and associated process and economic evaluations.

Setting: Six inpatient ward nursing teams (clusters) in two English NHS hospitals randomised to intervention ($n = 4$) or control ($n = 2$).

Participants: Patients ($n = 639$), staff ($n = 211$) and visitors ($n = 188$).

Intervention: CLECC is a workplace educational intervention focused on developing sustainable leadership and work team practices (dialogue, reflective learning, mutual support) theorised to support the delivery of compassionate care. The control setting involved no planned staff team-based educational activity.

Main outcome measures: Quality of Interaction Schedule (QuIS) for staff–patient interactions, patient-reported evaluations of emotional care in hospital (PEECH) and nurse-reported empathy (as assessed via the Jefferson Scale of Empathy).

Data sources: Structured observations of staff–patient interactions; patient, visitor and staff questionnaires and qualitative interviews; and qualitative observations of CLECC activities.

Results: The pilot CRT proceeded as planned and randomisation was acceptable to teams. There was evidence of potential contamination between wards in the same hospital. QuIS performed well, achieving a 93% recruitment rate, with 25% of the patient sample cognitively impaired. At follow-up there were more positive (78% vs. 74%) and fewer negative (8% vs. 11%) QuIS ratings for intervention wards than for control wards. In total, 63% of intervention ward patients achieved the lowest possible (i.e. more negative) scores on the PEECH connection subscale, compared with 79% of control group patients. These differences, although supported by the qualitative findings, are not statistically significant. No statistically significant differences in nursing empathy were observed, although response rates to staff questionnaire were low (36%). Process evaluation: the CLECC intervention is feasible to implement in practice with medical and surgical nursing teams in acute care hospitals. Strong evidence of good staff participation was found in some CLECC activities and staff reported benefits throughout its introductory period and beyond. Further impact and sustainability were limited by the focus on changing ward team behaviours rather than wider system restructuring. Economic evaluation: the costs associated with using CLECC were identified and it is recommend that an impact inventory be used in any future study.

Limitations: Findings are not generalisable outside hospital nursing teams, and this feasibility work is not powered to detect differences attributable to the CLECC intervention.

Conclusions: Use of the experimental methods is feasible. The use of structured observation of staff–patient interaction quality is a promising primary outcome that is inclusive of patient groups often excluded from research, but further validation is required. Further development of the CLECC intervention should focus on ensuring that it is adequately supported by resources, norms and relationships in the wider system by, for instance, improving the cognitive participation of senior nurse managers. Funding is being sought for a more definitive evaluation.

Trial registration: Current Controlled Trials ISRCTN16789770.

Contents

List of tables

List of figures

List of boxes

List of abbreviations

ADL	activities of daily living	MBI	Maslach Burnout Inventory
BPOP	Best Practice for Older People	MOP	medicine for older people
CC	Climate for Care	MRC	Medical Research Council
CEC	Carer Experiences of Care	NIHR	National Institute for Health Research
CI	confidence interval	NPT	normalisation process theory
CLECC	Creating Learning Environments for Compassionate Care	OR	odds ratio
		PDN	practice development nurse
CRT	cluster randomised trial	PEECH	Patient Evaluation of Emotional Care during Hospitalisation
EQ-5D-5L	EuroQol-5 Dimensions, five-level version		
		PPI	patient and public involvement
FECC	Factors that Enable Climate for Care	QALY	quality-adjusted life-year
FTE	full-time equivalent	QI Tool	Quality of Interactions Tool
HCA	health-care assistant	QuIS	Quality of Interactions Schedule
HRG	Healthcare Resource Group	RCT	randomised controlled trial
ICC	intraclass correlation coefficient	REC	Research Ethics Committee
IHOS	International Hospital Outcomes Study	RN	registered nurse
		SD	standard deviation
JSE	Jefferson Scale of Empathy		
MAQC	Matron's Assessment of Quality of Care		

Plain English summary

There is public concern about nurses' ability to care compassionately for older people, but there has been very little research about how to improve this situation. We have developed and studied a programme called Creating Learning Environments for Compassionate Care (CLECC). In the CLECC programme, all registered nurses and health-care assistants from participating teams attend a study day, with a focus on team building and understanding patient experiences. Then a senior nurse educator supports the team to try new ways of working on the ward, including regular supportive discussions on improving care. Each ward manager attends learning groups to develop their compassionate care leadership role. Two team members receive additional training in doing observations of care and feeding back to colleagues.

The CLECC programme was piloted on four wards in two English hospitals, with two other wards continuing with business as usual. We interviewed staff and observed CLECC activities to help us understand if it can be easily put into practice and if changes are needed in future. In addition, evaluation methods were tested, including ways to measure compassion, making sure that we could recruit enough older patients in a future study.

It was found that CLECC can be made to work with nursing teams on NHS hospital wards and that staff felt it improved their capacity to be compassionate. It was also learned that we could improve CLECC to help staff carry on using it, by, for example, helping senior nurses to understand their role in supporting staff to use CLECC.

It was found that observations of care were the best way to evaluate care quality for large numbers of older patients with dementia. We also gathered information about the best ways to test CLECC's value for money. These findings have given us a good foundation for designing an effective future evaluation of the CLECC programme.

Scientific summary

Background

Concerns about the degree of compassion in health care have become a focus of national and international attention. However, previous evaluations of compassionate care interventions have not provided robust assessments of their effectiveness in improving patient care, with limited use of experimental design and insufficient intervention description. Published qualitative evaluations do not examine the implementation process in depth or attempt to measure effectiveness. There is a need for high-quality mixed-methods evaluations to support health-care leaders in selecting appropriate interventions and to guide implementation.

Objectives

The Creating Learning Environments for Compassionate Care (CLECC) programme is a workplace educational intervention focused on developing sustainable leadership and work team practices theorised to support the delivery of compassionate care. This study aimed to assess the feasibility of implementing CLECC in acute hospital settings and to assess the feasibility of conducting a cluster randomised trial (CRT) with associated process and economic evaluations to measure and explain the effectiveness of CLECC.

The objectives were:

1. to determine the feasibility of implementing the CLECC intervention and sustaining the resulting work practices
2. to inform the design of a definitive evaluation of the effectiveness of CLECC
3. to inform the measurement of costs and benefits of CLECC in a definitive evaluation.

Methods

This mixed-methods study used two main approaches to assess feasibility: (1) a process evaluation to enable evaluation of the feasibility of implementing CLECC and (2) a pilot pragmatic CRT to inform a future evaluation of the effectiveness of CLECC. Ward nursing teams in two English NHS acute hospitals were included in the study; they were selected because they treat large numbers of older patients, and to ensure a mix of medical and surgical specialties. Six teams were randomised, with four allocated to the CLECC intervention and two to control conditions.

The Creating Learning Environments for Compassionate Care intervention

The CLECC intervention is a team-based educational programme focused on developing manager and team practices to create an expansive learning environment that enhances team capacity to provide compassionate care. Expansive (rather than restrictive) environments foster workplace learning and the integration of personal and organisational development. The implementation period of the programme is 4 months and is facilitated by a practice development nurse (PDN). CLECC is based on workplace learning theory with the ward conceptualised as a learning environment and ward team as a community of practice. It aims to embed ward-based manager and team practices including dialogue, reflective learning and mutual support, such that the team has the understanding and skills to continue to improve compassionate care following the end of the programmed activities. CLECC training consists of key activities that are combined to produce an integrated intervention over the implementation period: monthly ward leader action-learning sets; team learning activities, including local team climate analysis and values clarification; peer observations of practice and feedback to team by volunteer

team members; team study days focused on team building and understanding patient experiences; mid-shift 5-minute team cluster discussions; and twice-weekly team reflective discussions. Throughout the implementation period, ward leaders and their teams develop a team learning plan that includes a patient feedback plan and measures for continuing to develop and support leader and team practices that underpin the delivery of compassionate care.

Usual practice continued on control wards, that is, there was no planned team-based educational activity for staff.

Process evaluation

The feasibility of implementing CLECC into practice with the four intervention ward teams was assessed through a process evaluation using normalisation process theory as a framework. Qualitative interviews with nursing staff and managers during implementation and follow-up phases ($n = 33$ interviewees), observations of learning activities ($n = 7$) and ward leader questionnaires ($n = 12$) aimed to identify and explain the extent to which the CLECC intervention was implemented into practice, enabling an assessment of its workability and integration into existing work practices.

Pilot cluster randomised trial

In order to prepare for a definitive multicentre evaluation, the feasibility and piloted procedures for a pragmatic CRT of effectiveness were assessed. Cluster randomisation of staff and patients at ward nursing team level was undertaken. Outcomes were assessed at baseline and at 4 months after completion of the CLECC implementation period. The measurement of compassionate care was assessed across three complementary core outcomes: (1) researcher-rated observations of the quality of staff–patient interactions using the Quality of Interaction Schedule (QuIS), (2) patient-reported observations of emotional care using the Patient Evaluation of Emotional Care during Hospitalisation (PEECH) and (3) nursing staff self-reported empathy using the Jefferson Scale of Empathy (JSE). Baseline and follow-up data were also gathered on individual and ward team characteristics.

All trial analyses were carried out on an intention-to-treat basis. Possible QuIS ratings are positive social, positive care, neutral, negative protective and negative restrictive. The proportion of QuIS interactions rated for each of the five QuIS categories was analysed, including a further analysis for total positive ratings (the sum of positive social and positive care ratings) and total negative ratings (sum of negative protective and negative restrictive ratings). The frequencies of patients with the lowest (most negative) scores for each PEECH subscale were calculated. The differences between groups were tested using the chi-squared test. A three-level mixed-effects logistic regression model was fitted to investigate the effect of the CLECC intervention on the likelihood of a negative interaction. Predictive factors were included as fixed effects and presented as odds ratios (ORs) with 95% confidence intervals (CIs), after adjustment for baseline and ward consecutively. Mean PEECH and JSE scores were calculated by subscale and in total, and differences between groups at follow-up were tested using the Mann–Whitney U-test. Estimates of intracluster correlation were generated for each outcome measure.

Economic evaluation

The economic component of the study aimed to explore how costs and benefits might best be measured in a definitive evaluation. The feasibility of using EuroQol-5 Dimensions, five-level version (EQ-5D-5L), as a patient-based outcome measure at ward level was assessed. The likely training costs of the CLECC intervention and its implementation (through qualitative interviews with staff) were also explored.

Results

Feasibility of implementing and sustaining Creating Learning Environments for Compassionate Care

Staff were generally keen to participate and valued the positive contribution of CLECC not only to their own well-being but also to supporting good patient care. Many original CLECC practices were possible to implement as planned. Although practices did not always continue beyond the implementation period in their original form, staff reported that the philosophy and associated culture that CLECC had nurtured continued to guide their practice. Sustainability was strongly linked by staff to the extent to which the ward leader understood and valued CLECC.

Creating Learning Environments for Compassionate Care had some coherence for staff in that they appreciated its potential value, but their understanding was often limited to the concrete activities they had direct experience of. This may have then limited the development of participants' own practice in relation to CLECC, but interview data reflect extensive participation by staff, suggesting that engaging in CLECC was not limited by lack of coherence. Although it was often the concrete activities that were used by staff to explain CLECC, its role as a broader stimulus to action, and accompanying expectations that each team would use CLECC in their own way, developed cultures in which reflection, learning, mutual support and innovation were legitimised. In short, CLECC appears to have moved all of the participating teams further along the continuum to becoming more expansive learning environments.

In terms of cognitive participation, ward teams varied in the extent to which individual members saw it as their role to ensure that CLECC happened. Furthermore, there was uncertainty as to the role of matrons in supporting CLECC. Collective action to implement CLECC was dependent on the extent to which CLECC activities harmonised with the priorities of the wider organisation. Findings strongly reflect extremely busy hospital environments in which, without the right support for staff, care approaches tend to be very task focused. Staff flagging up what they valued about CLECC highlighted what nursing work can be like in contexts of this kind. The stress is not related only to barriers to satisfactory patient care. Ward staff valued CLECC because it refocused them on patients as people and because it involved sharing working time and space with other team members, promoting the feeling of being part of a team.

Our findings reflect the fact that, if the ways of working that CLECC promotes are not seen as valued or if this value is not indicated to frontline workers by managers, then these practices do not routinely occur. The findings also show, however, that it is possible to introduce practices at a local work team level that promotes relational ways of working between staff, albeit constrained in the absence of restructuring of the wider system. Findings point to refinements needed for CLECC and to the contexts in which it will be implemented to improve the prospects for its impact and sustainability. These focus on wider system restructuring to support work team conditions that enable the relational aspects of caring and working.

Informing future Creating Learning Environments for Compassionate Care evaluation

The findings from this study indicate that the use of experimental design to evaluate the effectiveness of compassionate care interventions within the context of a mixed-methods study is feasible, as is a focus on patient-based outcomes. Staff were amenable to the prospect of randomisation to either experimental condition. All wards recruited remained in the study throughout data collection, and all clusters randomised to the intervention went on to receive it. Blinding of patients and visitors to ward allocation appeared successful, although strategies to blind researchers gathering data need further development in a future trial. Evidence of pathways through which the CLECC intervention had the potential to influence practice in other wards in both of the participating organisations was found.

The recruitment rate for observations at baseline was 97% (i.e. 152 out of 157 approaches to eligible patients), and at follow-up was 90% (i.e. 157 out of 175 approaches). Some patients were approached and consented more than once, and some recruited patients were not observed. Overall, 273 patients were observed (i.e. 133 at baseline and 140 at follow-up). The mean age of patients was 82 years and 25% of patients observed had

evidence of cognitive impairment, suggesting that our sample was representative of the wider hospital population. Acceptability of the QuIS tool was high, and reliability between observers was acceptable. We did not find any evidence that staff changed their behaviour as a result of being observed. These findings support the selection of quality of staff–patient interaction, as measured by QuIS, as a candidate primary outcome in a future trial. With regard to clustering, there was a clear design effect apparent with QuIS at the observation session level.

The recruitment rate for patient questionnaires at baseline was 80% (i.e. 173 out of 217 eligible patients) and at follow-up was 75% (i.e. 186 out of 247 eligible patients). In total, 354 completed questionnaires were returned. Of these respondents, 83% were aged > 70 years and 12% had cognitive impairment. Most patients needed researcher help with questionnaire completion and the questionnaire was too long for some.

The recruitment rate for nursing staff questionnaires at baseline was 37% (i.e. 91 returned out of 249) and at follow-up was 35% (i.e. 87 out of 247). Overall, 178 questionnaires were returned. Respondents represented a range of ages, ethnic groups, job roles and experience. There was a perception that questionnaires were lengthy to complete and that staff were too busy.

Findings reflect a range of ward contexts at baseline, with similarities across some dimensions (e.g. bed numbers and staff views on relational care) and differences across others (e.g. staffing levels and duration of ward leadership). Using QuIS, staff–patient interactions observed at baseline were rated as total positive (73%), neutral (17%) or total negative (10%), but there was some variation in these proportions between wards. Using the PEECH questionnaires (with higher scores representing better experiences), patients at baseline tended to rate wards relatively positively {total mean PEECH score of 48.9 on a scale of 0–66 [standard deviation (SD) 11.7]}, although less so on the connection subscale [i.e. 1.66 out of 3 (SD 0.78)]. Results from the baseline nursing questionnaires showed variations between teams in nursing staff mean reported empathy levels (ward mean range = 107–120 out of a possible range of 20–140; higher mean scores indicate higher empathy).

At follow-up there were more positive (78% vs. 74%) and fewer negative (8% vs. 11%) QuIS ratings for the intervention wards than for the control wards. Once other variables were accounted for, the odds of a negative interaction were not significantly reduced because of the CLECC intervention (adjusted OR 0.30, 95% CI 0.07 to 1.32). In total, 63% of intervention ward patients indicated the lowest (i.e. more negative) scores on the PEECH connection subscale, compared with 79% of control group patients. However, the odds of a negative score were not significantly reduced because of the impact of CLECC once other variables were factored into the analysis (adjusted OR 0.47, 95% CI 0.14 to 1.59). Despite this, these are promising results given that data were gathered 4 or more months after the end of the implementation period, indicating that, if there is an effect, it is sustainable beyond the period in which CLECC is being actively facilitated. We found no evidence that nursing staff empathy may be improved because of CLECC, but these results have to be viewed in the context of a low response rate to nursing surveys.

Informing the measurement of Creating Learning Environments for Compassionate Care costs and benefits

Our findings have established the feasibility of estimating the cost of a CLECC-type intervention. Intervention costs were calculated as training costs (PDN time and staff time attending study day) and ongoing implementation costs (cost of staff engaging in CLECC activities on the ward). Findings show that, aside from initial CLECC training costs, the implementation of concrete CLECC activities by ward teams was not associated with additional resource use.

Use of the EQ-5D-5L was shown not to be feasible, mainly because different patients with different ailments and severity were involved at baseline and follow-up. We found that an impact inventory would provide estimates of both costs and benefits of CLECC with a focus on those associated with providing the intervention, but set within a wider context that includes effects on staff and on patients. Cost per change in each of the primary and secondary outcomes could also be estimated and compared with other studies.

Findings are not generalisable outside hospital nursing teams and this feasibility work is not powered to detect differences attributable to CLECC.

Conclusions

1. Compassionate care interventions, such as CLECC, should define the role of health-care leaders in mobilising structural capacity to support relational team working of staff in frontline caring roles.
2. The use of structured observations of staff–patient interaction quality is a candidate primary trial outcome measure but requires further testing and development.
3. A definitive evaluation of the implementation, effectiveness and cost-effectiveness of CLECC, drawing on experimental design in the context of a mixed-methods evaluation, is feasible.

Future work

Further funding is being sought to continue this research. In the first instance, this will focus on establishing the validity of QuIS in relation to people with cognitive impairment, and on establishing the organisational contexts in which CLECC is likely to achieve high impact and sustainability.

Trial registration

This trial is registered as ISRCTN16789770.

Funding

Chapter 1 Context

The study reported here aimed to assess the feasibility of implementing the Creating Learning Environments for Compassionate Care (CLECC) programme, a practice development programme aiming to promote compassionate care for older people in acute hospital settings, and to assess the feasibility of conducting a cluster randomised trial (CRT) with associated process and economic evaluations to measure and explain the effectiveness of CLECC.

In this chapter, we describe the background to the study, focusing in particular on the policy and practice context of the UK NHS. We also introduce the CLECC intervention, drawing on the international research literature to illustrate the rationale for designing and deploying this particular intervention.

In April 2013, the National Institute for Health Research (NIHR) Health Services and Delivery Research programme invited applications for funding research to support NHS organisations in responding to the Francis Inquiry analysis of care failures at Mid Staffordshire NHS Foundation Trust.[1,2] Acknowledging that all NHS organisations could learn from 'key system weaknesses' identified in the Francis Inquiry, the call specifically invited applications for 'robust evaluations of interventions to improve the leadership, organisational culture and quality of frontline care' (p. 1).[3] This report details a study funded through that call.

The NHS context

The need to strengthen the delivery of compassionate care in UK health and social care services, in particular to older patients, has been consistently identified as a high priority by policy-makers in recent years.[4] In addition to a series of investigations into high-profile failures, substantial and significant variations in the quality of hospital care for older people have been highlighted.[1,5] Variation exists between hospitals, but also between wards within hospitals and between staff within wards. Training, staffing levels, leadership, motivation and organisational culture are all implicated in failures of care. Although these issues are widely reported in the UK, there is evidence to suggest that they are relevant internationally.[6,7]

Care failures at the Mid Staffordshire NHS Foundation Trust in the late 2000s, and the inquiries that followed, were a watershed moment for the NHS. Over a period of some years, patient care in many wards and departments at the trust had been of very low quality, with, for instance, patients left in soiled bed clothes for lengthy periods, assistance not provided for patients who could not eat without help, and indifferent and unkind treatment by staff of patients, often older patients, and their families. Two inquiries led by Sir Robert Francis QC examined the causes of the lack of care and high mortality rates. The first inquiry focused on patient care at the trust and offered recommendations for improving practice at the trust.[2] The second inquiry focused on the systems of governance underpinning the care failures and offered recommendations for the NHS as a whole.[1] In the recommendations from the second inquiry, Francis called for a fundamental change in culture across the NHS towards a culture that puts patients first. Several of these recommendations focused on promoting compassionate nursing care. These recommendations focused on how to identify and promote desirable attributes (knowledge, skills, attitudes) in individual nurses. Many other recommendations focus on the systems needed to promote high-quality care and the responsibilities that should be held by key groups and organisations such as trust boards, NHS regulators, professional bodies and educational institutions. Although there is little detail about desirable systems and processes at a ward-team level, recommendations from both inquiries provide an outline of such measures to counteract the potential for the care failures encountered at the Mid Staffordshire NHS Foundation Trust. These include providing mechanisms through which health-care professionals can raise concerns about patient care with colleagues and with senior managers; the ongoing provision of training, support and supervision to nursing staff; and investing in ward leader roles that work alongside team members, providing role modelling and mentorship.

There have been significant changes to UK health-care provision since the establishment of the NHS in 1948. The improvement in medical treatments during this time has contributed to people living longer with more complex health conditions. Acute hospital inpatient beds are now predominantly populated by older people with multiple health conditions. However, recent years in particular have seen the adult social care system and pressures on primary care services affecting secondary care. Acute hospitals have struggled to meet their performance and financial targets. During the time that this study was taking place, health and social care in England was in the midst of unprecedented demand and financial challenges, with NHS providers having overspent by £2.45B by the end of 2015/16.[8]

It has been acknowledged that an increase in staffing numbers is required and a safe staffing guideline for nurses working on wards in acute hospitals has been published.[9] Although this was supported by the creation of 24,000 new nursing posts between 2012 and 2015, an increase of 8.1%, demand has outstripped supply with a deficit of 8.5% of the funded establishment recorded in April 2015.[10] This deficit is worse in adult acute nursing, with reported vacancies amounting to 9.7% of the establishment. It is common practice for wards to run with staff vacancies and for the staff complement to be made up of staff from nursing agencies. Recruitment drives targeted at overseas nurses have regained popularity, but this too is a temporary fix, with European Union nurses choosing to exercise their free movement rights.[11]

Through the development and tightening of systems for financial control and performance management, the NHS has seen an intensification of health-care work through higher patient numbers and time-based targets.[12–14] Use of staff without professional qualifications is increasing and nursing staff job satisfaction is low. As we developed this study, anecdotal evidence from a number of NHS acute hospitals indicated that the leadership and team practices, such as role modelling, mutual support, reflective learning and dialogue, required to support nursing staff in their caring role[15] were unlikely to be in place in most care settings.

Approach and definition of key terms

The literature is both confused and confusing in the way that compassion is used as a term. There are four key components of the narrative of compassion in nursing, and we have found these helpful to guide our thinking in this study about what compassion is.[16] The first is a set of ideas about the *moral attributes* of a 'compassionate' nurse, including wisdom, humanity, love and empathy.[17–19] These moral attributes are expressed through a kind of *situational awareness* in which vulnerability and suffering are perceived and acknowledged.[19,20] These perceptions underpin participation of the nurse in *responsive action* that is aimed at relieving suffering and ensuring dignity, and which involves the nurse in a participatory relationship in which the nurse exercises *relational capacity*[19,21–23] through which empathy is experienced and a caring pastoral relationship is constructed.[15,24,25]

Our systematic review of research reporting older patients' experiences of hospital care highlights the importance of this caring relationship to shaping experiences.[6] Older people want nurses, and others, to use social interactions to see the person behind the patient ('see who I am'), to establish a warm and human connection ('connect with me') and to establish understanding and involvement ('involve me').[6] A later review focusing on nurses' experiences indicates that registered nurses strive to achieve the caring relationship that is valued by patients, indicating that a perceived lack of compassion in nursing may not be attributable to a lack of the necessary moral attributes or situational awareness on the part of individual nurses. The findings reflected that nurses' relational capacity and capacity for responsive action can depend on ward-level conditions, and that there is a greater tendency for nurses with low relational capacity to avoid relationships with patients and to burnout, in spite of aspirations to a higher standard of care.[15] This study builds on these findings through the development and evaluation of an intervention targeted at improving the capacity of nurses to respond to patient vulnerability and suffering, specifically their relational capacity and capacity for responsive action.

The links between positive patient experiences, leadership, work team climate and the well-being of individual staff are becoming evident through research and so interventions that focus on developing these elements (leadership, work team climate, staff well-being) would appear to be worthwhile to support the development and exercise of relational capacity. A NIHR study on culture change and quality of acute hospital care for older people found that more positive patient and carer assessments of care were correlated with higher staff ratings of a supportive team climate and a shared philosophy of care.[26] In addition, ward leadership was a strong indicator of team members sharing a philosophy of care and feeling highly supported, a finding that, together with the qualitative data, highlighted the vital role of the ward manager in shaping a positive team climate for care.[26] These findings were mirrored in a second NIHR study that highlighted the key role of the ward leader in shaping the local ward climate of care, the importance of staff well-being, and, in particular, staff experiences of good local work-group climate, co-worker support, job satisfaction, positive organisational climate and support, and supervisor support as antecedents of positive patient experiences.[27]

Creating Learning Environments for Compassionate Care

Parts of this section are reproduced from Bridges J, Fuller A. Creating learning environments for compassionate care: a programme to promote compassionate care by health and social care teams. *Int J Older People Nurs* 2015;**10**:48–58,[28] with permission from John Wiley & Sons. © 2014 John Wiley & Sons Ltd. CLECC is a team-based implementation programme focused on developing leadership and team practices that enhance team capacity to provide compassionate care. Its objectives are to:

1. create an expansive workplace learning environment that supports work-based opportunities for the development of relational practices across the work team
2. develop and embed sustainable manager and team relational practices, such as dialogue, reflective learning and mutual support
3. optimise and sustain leader and team capacity to develop and support the relational capacity of individual team members
4. embed compassionate approaches in staff–service user interaction and practice, and continue to improve compassionate care following the end of programmed activities.

The CLECC intervention was designed for use by ward nursing teams in inpatient settings for older people, but is potentially transferable for use in other settings. The implementation programme takes place over a 4-month period but it is designed to lead to a longer-term period of service improvement. By envisaging the workplace as a learning environment and the work team as a community of practice, CLECC brings a distinctive approach to promoting compassionate care. It uses insights from workplace learning research to develop practices that enhance the capacity of the manager and work team to provide compassionate care within a complex and dynamic organisational context.

Fuller and Unwin's research on workplace learning and workforce development in a range of public and private sector industries demonstrates the importance of identifying and analysing both the organisational and pedagogical features that characterise diverse workplaces as learning environments.[29] They argue that this approach allows workplaces (e.g. hospital wards) to be located on the 'expansive–restrictive' continuum. Those at the expansive end are characterised by a range of features including the following: the knowledge and skills of the whole workforce (not just the most highly qualified or senior staff) are valued; managers facilitate workforce and individual development; teamwork is valued; innovation is important; the team has shared goals focusing on the continual improvement of services (or products); there is recognition of and support for learning from 'each other'; learning new knowledge and skills is highly valued; and the importance of planned time for off-the-job reflective learning is recognised. It follows that an expansive approach to workforce development is more likely to facilitate the integration of personal and organisational development. This has important implications for the design of learning interventions as it requires workplace learning to be perceived as something that both shapes and is shaped by the work organisation itself rather

than as a separately existing activity. Such an understanding highlights the importance of interventions that situate and integrate individual and team learning in the everyday life of the workplace (in this case the clinical unit/ward/team setting) as well as providing opportunities for off-the-job provision to foster reflection, consolidate learning and deepen understanding, thereby enhancing ownership and sustainability of new practices.

Our synthesis of qualitative research highlights the importance of the relational aspects of care to shaping older people's hospital experiences.[6] Being compassionate requires 'relational capacity' in practitioners, that is, capacity to experience empathy and to engage in a caring relationship.[24] Our research also shows that nurses' relational capacity can depend on ward-level conditions, and that there is a greater tendency for nurses with low relational capacity to avoid relationships with patients and to burnout, in spite of aspirations to a higher standard of care.[15] CLECC uses workplace learning principles to develop practices that enhance the capacity of the manager and work team to support the ongoing relational capacity of its individual members. This leadership and team capacity are key characteristics of the ward-level conditions needed to support nurses' relational work[15] and thus improve patient experiences, and are an important foundation for team activities, such as using service user feedback constructively.[30] By envisaging the workplace as a learning environment and the work team as a community of practice, CLECC brings a distinctive approach to promoting compassionate care, which enhances the capacity of the manager and work team to provide compassionate care within a complex and dynamic organisational context.[29,31] An overview of this programme theory for CLECC is shown in *Figure 1*.

During the 4-month implementation programme, CLECC learning activities are led by a senior (UK band 7) practice development nurse (PDN) or practitioner with strong influencing and interpersonal skills.[28] The PDN delivers the study days, facilitates cluster and reflective discussions, facilitates the action-learning sets and co-ordinates the peer observations of practice (see the following subsections for more detail on each of these activities). This individual is not part of the hierarchy of the ward team and this enables a distinction between CLECC activities and performance management. The activities themselves are characteristic of a practice development approach.[32] CLECC operates at two key levels: team and team manager. A focus on the team aims to develop team capacity to support team members to provide
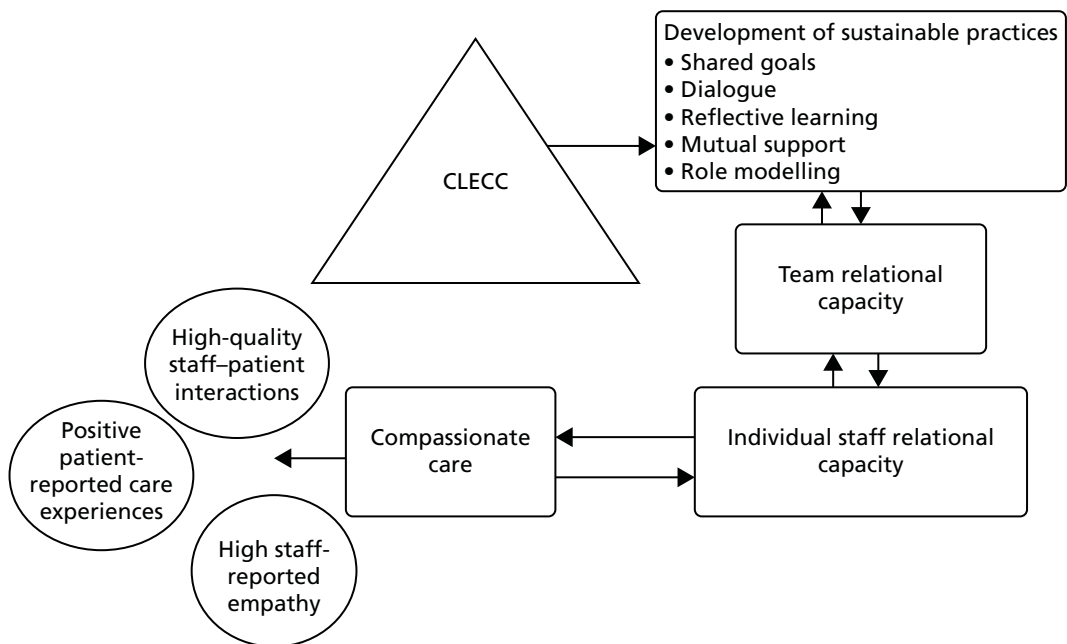
FIGURE 1 Overview of CLECC programme theory. Reproduced from Bridges J, Fuller A. Creating learning environments for compassionate care: a programme to promote compassionate care by health and social care teams. *Int J Older People Nurs* 2015;**10**:48–58,[28] with permission from John Wiley & Sons. © 2014 John Wiley & Sons Ltd.

compassionate care. An equivalent focus on the leadership capacity of the team manager (in ward settings this is the ward manager) aims to develop their role in leading the team, role modelling good practice and enhancing and embedding the desired team practices.

Although the programme draws on elements that have been piloted in other programmes, it is novel in combining these elements with an explicit focus on establishing reforms to routine practice and organisational resources that establish the basis for sustained changes in compassionate care. Although the implementation process is a key element, the essence of the programme is the ongoing processes of peer observation, daily cluster discussions, weekly reflective discussion and the use of evidence-based guidelines. *Table 1* sets out a typical schedule for the implementation programme.

Action-learning sets

The crucial role of the ward leader in influencing the caring culture and the work culture is well documented, with strong and visible leadership identified as an essential requirement for the delivery of dignified care.[26,33] In CLECC, ward leaders attend four 4-hour action-learning sets during the implementation programme. Action-learning sets have been used in other projects, including other development projects focused on dignity in care and/or care for older people, to provide an extended reflective space for individuals in a key position of influence to explore and develop their leadership role.[34–36]

The CLECC action-learning sets follow the McGill and Beaty model[37] for action-learning, that is, sets are made up of between four and eight members and are facilitated by an experienced facilitator. Set members may or may not work in the same organisation but often have similar work roles. Participants bring work problems of their own choosing to the session and other set members aid them in reflecting on the issue and drawing up an action plan to address it. In addition, each of the action-learning sessions is themed to encourage a focus on issues related to the manager's role in supporting the delivery of compassionate care. The first session focuses on establishing relationships among set members and agreeing ground rules. The themes for subsequent sessions are: (session 2) workplace climate/team values/valuing staff, (session 3) enhancing team

TABLE 1 The CLECC implementation programme schedule

Activity	Month			
	1	2	3	4
Ward leader action-learning sets	Session 1/setting up set, setting ground rules	Session 2/workplace climate/team values/valuing staff	Session 3/enhancing team capacity for compassionate care	Session 4/influencing senior managers
Team learning and service user feedback plan	Introduce and discuss	Discussion and draft by ward leader	Finalise, identify resources needed to support, present	Senior manager feeds back response to team plan
Peer observations of practice	Identify care-makers	Train care-makers	Observations of practice	Feedback on observations of practice
Study days	Team analysis of workplace climate/values clarification	–	–	–
Cluster discussions	Ongoing	Ongoing	Ongoing	Ongoing
Reflective discussions	'I feel valued at work when . . .' exercise	Team values clarification exercise; BPOP activities	BPOP framework activities; team learning and service user feedback plan discussions	Reflections on feedback from observations of practice

BPOP framework, Best Practice for Older People framework.
Reproduced from Bridges J, Fuller A. Creating learning environments for compassionate care: a programme to promote compassionate care by health and social care teams. *Int J Older People Nurs* 2015;**10**:48–58,[28] with permission from John Wiley & Sons. © 2014 John Wiley & Sons Ltd.

capacity for compassionate care, and (session 4) influencing senior managers. Reflecting on the results of other programme activities supports discussion in these themes. For instance, during the team study days, all staff will have been invited to complete a questionnaire on perceptions of ward climate. Reflecting on the results of these questionnaires in addition to the results of the 'I feel valued when . . .' exercise (see *Reflective discussions*) is encouraged in the second action-learning set.[36] In addition to this reflective learning set, participants facilitate each other in developing practical ways of dealing with some of the issues that arise during the programme, these issues being informed by the findings relating to ward leader strategies in an earlier dignity in care project.[38] Participants are encouraged to use the sets to devise a personal plan associated with their current and future role in promoting compassionate care, including planning clinical supervision sessions for themselves with a selected mentor and/or negotiating ongoing action-learning set access.

In addition to action-learning sets, ward leaders are also facilitated to further develop their relationship with their line manager as a way of accessing additional support. This includes a 1-hour meeting every 2 weeks during the 4-month implementation programme. These meetings provide an opportunity for the line manager to learn about the project and to explore opportunities to participate.

Team learning

Interventions to improve care quality at a ward or unit level may succeed, even if the wider organisation has features that inhibit service improvement on a wider scale.[26] Ward-level conditions can strongly influence nurses' capacity to build and sustain therapeutic relationships with patients.[15] Other work suggests that the work team can function as a buffer to stressors from the wider organisation, but that the team's capacity to do so depends on the extent to which the group perceives its role as supportive of the relational work of individual members.[39] Social structures and relationships within the team and the capacity of team members to support each other are a primary influence on how individuals learn emotional abilities and how tacit emotional knowledge is transferred.[40] Dialogue and reflection within the team, particularly with a focus on sharing experiences and narratives, appear to be linked with the development of individual emotional abilities, but these activities depend on the extent to which the workplace provides an environment in which staff feel safe to participate.[40] Other work indicates that expecting staff to, for example, use patient feedback constructively in the absence of team preparation to hear the patient feedback is unlikely to lead to service improvements.[30] A strong focus of the intervention is on the development of shared team goals and expectations, team dialogue, reflection and role modelling. Early activities in the intervention reflect a focus on developing a sense of security within the team,[41] with dialogue and reflective learning activities providing the forum for the development of individual and team relational capacity, and the creation by the team of sustainable practices and plans to support ongoing capacity through:

- commitment and role modelling by senior staff in the team – providing information, opportunities for discussion and involvement in goal-setting and decision-making
- creating facilitated collective and reflective 'spaces' through

    - mid-shift scheduled 5-minute cluster discussions, using trigger questions or observations as behavioural nudges in team members' planned work with patients
    - twice-weekly longer reflective group meetings, which should draw on a variety of toolkit materials to prompt dialogue and reflective learning, and give staff a regular opportunity to stand back from the demands of their operational practice

- building relationships in the team, using activities including team analysis of workplace climate
- critical reflections by the team on caring for and supporting each other, on team relational capacity and on delivery of compassionate care
- team values – the clarification and development of a shared vision
- developing shared ownership of compassionate care and understanding about how learning in the workplace can contribute to improved individual and team practice and 'expansive outcomes'
- development of a team learning plan, including a plan for hearing and responding to patient feedback.

Teams can be unidisciplinary or interdisciplinary, but an inclusive approach is essential, so for instance, use of CLECC with a nursing team includes the participation of all nursing staff, namely the ward manager, registered nurses, care assistants/health-care support workers and nursing students.

Peer observations of practice

Two staff volunteers from the team were selected to become 'care-makers', with their primary role being to undertake peer observations of practice for feedback to their colleagues. Care-makers receive 4 hours of training in peer observations of practice and undertake 8 hours of observation each during the programme. Peer observations are conducted using the Quality of Interactions Schedule (QuIS)[42] and findings are fed back at reflective discussion meetings (see *Reflective discussions*) with the help of the PDN. The results from the care-makers' observations of practice on the ward are shared to trigger discussions about how to build on existing good practice and improve practice when this is needed.

Study days

On each ward, three or four full study days are delivered by the PDN during the first month of the programme to enable all ward members to attend one study day. The purpose of the day is to prepare staff for the workplace elements (including cluster and reflective discussions) of the programme by providing opportunities to experience some of the techniques; to develop an understanding of underlying concepts; and to recognise an active role in their personal and team learning journey. Elements of the programme for classroom training are shown in *Box 1*.

Cluster discussions

Mid-shift cluster discussions commence during the first month (following the delivery of study days) and run daily throughout the implementation period. These 5-minute cluster discussions are facilitated initially by the PDN and all nursing staff on the ward at the time of the cluster discussion are encouraged to join the group discussion. The discussion focuses on establishing how the individual staff are at that moment in that context and provides opportunities for the group to offer help and support to members when difficulties are identified. Similar strategies have been used in other projects focused on developing dignity in care/compassionate care.[36,45]

BOX 1 The CLECC study day elements

Introduction to the BPOP framework.[43]

Life shield activity and group discussion: 'See who I am'.

Questionnaires and discussion on ward climate, dialogue and reflective learning on the ward.

Values clarification exercise about compassionate care.[44]

Videos, stories and discussion with service users: 'Involve me'.

Introduction to workplace learning activities and discussion on how to implement/support/sustain.

BPOP, Best Practice for Older People.
Reproduced from Bridges J, Fuller A. Creating learning environments for compassionate care: a programme to promote compassionate care by health and social care teams. *Int J Older People Nurs* 2015;**10**:48–58,[28] with permission from John Wiley & Sons. © 2014 John Wiley & Sons Ltd.

Reflective discussions

Twice a week, members of the team on duty at the time scheduled for a reflective discussion (usually the afternoon) arrange their work to enable their attendance at a group meeting facilitated by the PDN. To enable all staff on a shift to participate, two sessions may need to be held on the same day, both attended by the ward leader. This interaction is held in a comfortable meeting room on or near to the place of care, but away from the immediate distractions of care delivery. The meeting is for all team members, including senior members of the team and temporary team members, such as student nurses. The meetings involve a variety of group work tasks, some of which are repeated to enable the maximum number of team members to take part, whereas others will be unique. Tasks are aimed at opening up dialogue and reflective learning among those present, and so are selected to prompt personal reflections and narratives about experiences on the ward. They include:

- 'I feel valued at work when . . .' – those present are invited to complete this sentence to trigger discussions about valuing and supporting each other[36]
- team values clarification about compassionate care – drawing on collated results of the values clarification exercise in classroom sessions to develop shared vision[36,44]
- drawing on collated results of ward climate analyses to identify factors that need to be supported or changed[36]
- peer observations of practice – the results from the care-makers' observations of practice on the ward are shared to trigger discussions about how to build on existing good practice and improve practice when this is needed[36]
- the Best Practice for Older People (BPOP) framework – using resources and questions/prompts from the BPOP framework is an essential guide to generating discussion[46]
- a team learning plan – working with managers to draw up a team learning plan focusing on compassionate care and using patient feedback.

Best Practice for Older People framework

The BPOP framework is a set of evidence-based UK guidelines for nurses working with older people in acute settings.[43,46] Its successful use in development projects aimed at service improvement indicates that it is useful in guiding the practice of health and social care professionals working with other client groups (i.e. not just nurses working with older people). One example of this wider use is the City University Dignity in Care project at two London hospitals.[36,47] A resource has been published for use alongside the BPOP framework, providing teams with trigger questions and guidance aimed at generating dialogue and reflective learning in the team, and opening up conversations in which team members give and receive support and help with difficult matters, such as talking to patients about dying.[46] In CLECC, this resource is used to identify areas for support, action and learning in the team, and to inform the development of strategies to address these areas. Examples of trigger questions in this resource are:

- What kind of patients are most difficult to communicate with, and why?
- What kind of patients are most difficult to involve, and why?
- What subjects are hardest to talk to patients about, and why?
- What kind of relatives are most difficult to involve, and why?

The implementation stage of the programme takes 4 months, and during this time ward leaders and their teams develop a team learning plan that includes inviting and responding to patient feedback, and puts in place measures for continuing to develop and support manager and team practices that underpin the delivery of compassionate care. The team learning plan is presented to a senior trust manager, together with a case for support, and the relevant manager is invited to visit the ward team to discuss the plan and respond in person to the proposals.

In summary, the focus of the intervention is on creating an 'expansive' environment that supports work-based opportunities for the development of shared goals, dialogue, reflective learning, mutual support and role modelling for all members of the team at an individual and a group level.[29] The programme theory states that such an environment should facilitate staff to engage with and learn from service user experiences and their

own emotional responses, share positive strategies and support, and optimise and sustain personal and team relational capacity to embed compassionate approaches in staff–service user interaction and practice. Expansive outcomes are theorised to include high-quality interactions between service users and staff, and between care team members, positive care experiences reported by service users and staff reports of high empathy with patients and carers.

Introduction to the study

Findings from our systematic review, reported in *Chapter 2*, highlight the lack of definitive evaluation research on compassionate nursing care. Responding to a general absence of strong evidence for the effectiveness of service improvements related to compassion, and building on compelling evidence indicating that a strategy targeted at improving leadership and local ward team climate could improve patient experiences, the study reported here is a foundational step in addressing the need for well-designed and rigorous evaluation to understand what works best in improving care and patient experiences.

This study, conducted in two English hospitals during 2015–16 and reported in the chapters that follow, aimed to assess the feasibility of implementing CLECC and of conducting a CRT with associated process and economic evaluations to measure and explain its effectiveness. Conducting the study provided an important opportunity to assess the feasibility of a programme with unique characteristics designed to address the issues identified in other studies, and to design an evaluation that includes an assessment of its effectiveness. The process and economic evaluations aimed to provide important information about CLECC's workability, its integration into practice and to lay the foundations for establishing its value for money. The findings reported below (see *Chapters 7–10*) provide the basis for planning a larger, multicentre evaluation aimed at producing evidence that can be generalised more widely to other NHS acute care providers and, together with a refined intervention package, will be a valuable resource for change and improvement for NHS managers, practitioners and educators.

Chapter 2 Literature review

Although current definitions of compassion in nursing practice are imprecise and sometimes confused (see *Chapter 1*), there is intense interest in this problem both within and outside the profession of nursing. However, little is known about what strategies are effective in promoting compassionate care among nurses. To date there has been no rigorous critical overview of research into interventions designed to promote compassionate care among nurses in practice. This chapter aims to provide an overview of the evidence base on the evaluation of interventions for compassionate nursing care. It begins with an overview of qualitative research on compassionate care interventions. It then reports a systematic review of studies that evaluate the effectiveness of interventions to promote compassionate nursing care.

Qualitative research

Recent years have seen the use of qualitative research methods to underpin the development and evaluation of a number of interventions focused on improving compassionate care, or dignity in care, at the hospital ward level.[30,34,36,47–49] Interventions developed and evaluated in this way have typically been faciliated by a senior nurse, using reflective learning, action research and/or appreciative inquiry to work with ward-based nursing staff (often using patient stories and/or observations of practice) to strengthen support for existing good practice and to make changes when needed. These interventions are often shaped by a 'relationship-centred' philosophy in which achieving the well-being of all groups (patients, staff, family carers) is seen as fundamental to high-quality care.[41] They have used democratic and participatory processes involving patients, staff and sometimes family carers to articulate the patient's needs and to shape the practice changes made.

The accompanying qualitative evaluations have provided important information about the processes of change and the factors enabling and inhibiting sustainable change. Some of these evaluations have reported concrete practice changes resulting from the intervention,[36,47–50] while others report more variable success.[30,34] For instance, Dewar and Nolan[49] and Dewar[50] used appreciative inquiry and action research to involve older people, staff and relatives in developing compassionate relationship-centred care on an acute hospital ward. Methods used included participant observation, interviews, storytelling and group discussions. The findings indicated the value of appreciative caring conversations between staff, patients and relatives enabling all parties to discover 'who people are and what matters to them'[49,50] and 'how people feel about their experience'[49,50], with this knowledge enabling them to 'work together to shape the way things are done'.[49] In the resulting model, Dewar and Nolan[49] detail how older people, staff and relatives can work together to implement compassionate relationship-centred care.[49] In specifying 'how people can work together to shape the way that things are done here', Dewar[50] identified a number of important conditions for staff to feel able to express emotions, share experiences and ideas with each other, consider others' perspectives, take risks, use 'curious questioning' to examine situations and challenge existing practice, all identified as important actions to support the delivery of compassionate care. These conditions included transformational leadership, the level of support received from colleagues and senior staff, a shared set of principles for caring, open dialogue within the team and opportunities in which people had permission and space to reflect. These conditions echo the findings from other research as conditions at the team level that can support high-quality care.[26,27,51] Dewar[50] reports how these conditions developed and how compassionate caring practices became embedded in the work of the team over the course of the year-long project, providing valuable evidence that change of this kind is possible.

However, Dewar's project took place over the course of 1 year on an already high-performing ward with a strong leader.[50] The findings informed development work across a wider Leadership in Compassionate Care project implemented in a number of settings, but evaluation of the impact of these strategies elsewhere does

not report the influence of the ward climate or programme length on outcomes, so evidence is lacking that such strategies can be universally effective regardless of work team context.[52] In a contrasting study to Dewar's that explored the use of discovery interviews with older hospital patients as a way of improving dignity in care, Bridges and Tziggili[30] found that ward teams required strong and consistent leadership and intense preparation before they were able to hear and respond to patient stories about care. Both organisations involved in this dignity project experienced significant delays in the progress of the project and limitations in its impact because of a lack of leadership at the ward level and a lack of preparedness of the ward teams to engage in responding positively to patient feedback. One ward team with a strong leader was able to successfully engage with the patient stories, but only after some months of team preparation. These findings indicate that, while some wards may be ready to engage in programmes, such as Dewar's, others could benefit from a period of groundwork in which leadership and mutually supportive team practices are established.

The evaluative focus of these studies is the mechanisms for change used, particularly the democratic and participatory processes that involve patients, staff and sometimes family carers in articulating the patients' needs and shaping the practice changes made. These qualitative accounts often provide a fuller picture of the interventions deployed than the studies reviewed below, and often include an analysis of the enablers and barriers to change. However, they do not examine in depth the process of implementation itself and so fail to systematically identify the contexts in which successful implementation is more likely or, where contexts are not receptive, how resources, relationships and norms in the wider system may need purposeful restructuring to support implementation and sustain longer-term change.[53,54] In addition, as would be expected with a qualitative approach, there is only weak objective evidence of effectiveness of the interventions deployed in these studies in relation to impact on patient outcomes.

Review methods

The remainder of this chapter reports a systematic review of studies that evaluate the effectiveness of interventions for compassionate nursing care, using the four key components of the compassion narrative identified in *Chapter 1* to provide an operational definition. The objectives of the review were:

1. to systematically identify, analyse and describe studies that evaluate interventions for compassionate nursing care
2. to assess the descriptions of the interventions for compassionate care used, including design and delivery of the intervention and theoretical framework
3. to evaluate the nature and strength of evidence for the impact of interventions.

The review was conducted, guided by the Cochrane Collaboration methods[55] to assure comprehensive search methods and systematic approaches to analysis of the review materials. Sections of this review report are reproduced or adapted with permission from Blomberg *et al.*[16] and with permission from Elsevier. © 2016 The Authors. This is an open access article under the CC BY-NC-ND 4.0 license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Search strategy
A systematic search for primary research evaluating compassionate care interventions was undertaken on CINAHL, MEDLINE and The Cochrane Library databases (including the Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effectiveness, CENTRAL register of controlled trials, Health Technology Assessment Database and Economic Evaluations Database) in June 2015. No date limits were applied to the searches conducted.

Terminology in relation to compassionate care is problematic and as noted above, there is no one agreed definition of compassionate care. Instead, a number of terms are used interchangeably and inconsistently across the health-care literature. A broad and inclusive approach was therefore used in preliminary searches to scope and map the field. As many terms relating to compassionate care were identified and used as possible, but with a focus on identifying studies that reflected one or more of the key components of compassionate care outlined above. Through this mapping, relevant keywords were identified (e.g. professional–patients relations, dignity, person-centred care, relationship centred care, empathy, compassion, caring, and emotional intelligence). These keywords were used in final searches. Terms related to compassion were combined (AND) with terms related to relevant methods and occupational groups. Relevant index terms were included, which varied across databases (see *Appendix 1* for MEDLINE and CINAHL search strategies). Although no additional searches for unpublished (so called 'grey') literature were conducted, the sources used do index PhD theses (CINAHL) and some conference abstracts (CIHAHL, The Cochrane Library). Searches were limited to the English language.

Selection

An adapted PICO (population, intervention, comparison and outcome) framework was used to guide study selection.[56] We included primary research studies comparing the outcomes of an intervention designed to enhance compassionate nursing care (in any setting to any client group) with those of a control condition. Eligible designs were randomised controlled trials (including CRTs) or other quasi-random studies, interrupted time series and before-and-after studies (controlled or uncontrolled). Studies were excluded if they were focused exclusively on students, or if interventions were not directed at changing nursing staff behaviour.

The lack of conceptual clarity about compassion in the literature necessitated an inclusive approach to studies that were not necessarily labelled as addressing 'compassion'. We developed selection criteria based on the four elements of the compassion narrative described above (moral attributes of a 'compassionate' nurse including empathy, nurses' situational awareness of vulnerability and suffering, nurses' responsive action aimed at relieving suffering and ensuring dignity, and nurses' relational capacity) so that studies were included if they met one or both of the following criteria:

- explicit goal of the intervention was stated as improving compassionate nursing care (or a closely related construct, i.e. dignity, relational care, emotional care) (through addressing nurses' moral attributes, situational awareness, responsive action and/or relational capacity) and/or
- primary outcomes that assessed or evaluated either nurses' self-reports of compassion and/or ability to deliver compassionate care (moral attributes, relational capacity), and/or observed quality of interactions or other measure of compassion (situational awareness, responsive action), including patient reports of experienced compassion or a closely related construct.

The titles and abstracts from the search were screened against the inclusion criteria independently by individual review team members (JB, KB, PG and YW; see systematic review team members listed in *Acknowledgements*). During the screening process, frequent meetings were held by the research team in order to compare independent selections, resolve disagreements and make decisions. On independent rating (i.e. before discussion) reviewer pairs achieved between 80% and 90% agreement. In most cases of disagreement papers were excluded after discussion. Full-text papers were retrieved for all papers that screened positively in the first stage or about which a clear decision could not be taken (due to lack of information). Each full-text paper was reviewed independently by two team members followed by a decision to include or exclude in the final review. These reviews were followed by further team discussion to finalise inclusion into the data set. The reference lists of full-text papers included were scanned for further items. The search and selection process is summarised in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow chart (*Figure 2*). The number of duplicates removed was not recorded.
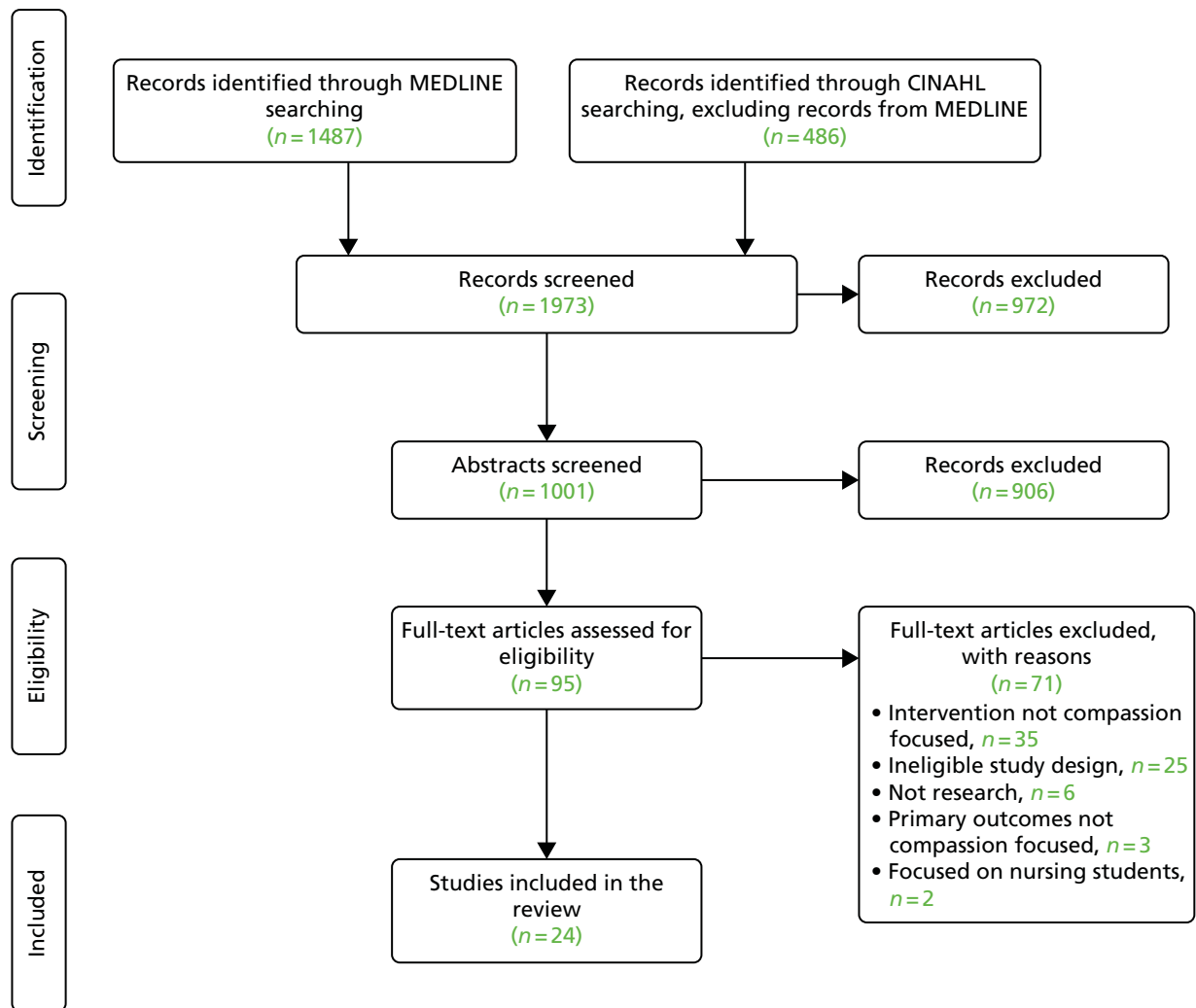
FIGURE 2 Flow diagram for systematic review searches. Adapted from Blomberg *et al*.[16] with permission from Elsevier. © 2016 The Authors. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Quality assessment

In order to effectively represent the variation in study quality evident in findings from the preliminary mapping phase, and to properly reflect the strength of evidence, we undertook a simple grading in order to categorise the strength of the underlying design of studies we retrieved.[57] Because of heterogeneity of study design identified in early scoping work, a rating of strong, medium or weak quality was allocated to each study depending on where the study design sat on the hierarchy of evidence for effectiveness in tandem with an assessment of its design and execution. The method selected was in line with the GRADE (Grading of Recommendations Assessment, Development and Evaluation) system as used by Cochrane for rating evidence to guide a broad assessment of individual study quality.[57,58] Studies were rated as high quality where outcomes were compared between treatment (intervention) and control groups, where allocation to groups was random, and where equivalence between groups was explicitly demonstrated. Study designs included here were randomised controlled trials (RCTs) and CRTs which met these conditions. Studies were rated as medium quality where outcomes were compared between intervention and control groups, and where equivalence between groups was demonstrated, but where other methodological issues weakened the design, for instance non-random allocation to groups or small sample size. Study designs included here were CRTs with small

numbers of clusters (for instance, $n = 2$) and controlled before-and-after studies with non-random allocation to groups. Uncontrolled before-and-after studies were rated as low quality as were other studies where other significant methodological shortfalls weakened claims of demonstrating effectiveness (e.g. controlled before-and-after studies where equivalence between groups is not demonstrated). These quality assessments were made by individual members of the review team, and checked with one other team member's ratings until consistent ratings were achieved.

An evaluation of quality of description of the intervention was also performed for each included study. The material used as the basis for this evaluation was the information provided in the paper about the intervention in addition to further information about the intervention accessed from sources referenced within the original paper. We did not otherwise seek out information about the intervention, wishing to test the extent to which the original paper and its referenced sources provided sufficient information to enable the intervention to be replicated. Each study was analysed against the criteria for description of group-based behaviour change interventions devised by Borek *et al.*[59] This framework provides a checklist for assessing the reporting of behaviour change interventions against 26 criteria covering intervention design, intervention content, participants and facilitators. Intervention design features assessed included intervention development methods; setting; venue characteristics; number, length and frequency of group sessions; and period of time over which group meetings were held. Intervention content assessed included change mechanisms or theories of change, change techniques, session content, sequencing of sessions, and participants' materials activities during sessions and methods for checking fidelity of delivery. Participant features assessed included group composition and size, methods for group allocation, and continuity of group membership. Facilitation features assessed included number of facilitators; facilitator characteristics and preparation including professional background, personal characteristics, training in intervention delivery and training in group facilitation; continuity of facilitator's group assignment, facilitator's materials and intended facilitation style. These assessments were conducted by one team member, and supplemented and refined in discussion with other team members.

Data analysis

A qualitative analysis was conducted across the different interventions reported to describe intervention types and contexts, and mechanisms for change. This analysis was conducted in smaller groups in the review team but further enriched through discussion of process and emerging findings among all group members.

Data were extracted for each study by Jackie Bridges and Karin Blomberg, including study design, sample and settings, summary details of intervention, outcomes and measurements, results and process issues. Results were tabulated and used to generate summary descriptions across key characteristics. Heterogeneity of studies in terms of interventions, methods and outcomes meant that a meta-analysis was not warranted, and so a more descriptive approach was merited. We considered the potential to pool studies using standard mean differences for measures, but this method requires that the instruments are measuring the same underlying construct and that the interventions have common mechanisms, but this was not clearly the case. The main intervention types were agreed through team discussion, as were key outcome types. Findings on effectiveness of individual interventions were plotted against key outcome types, and this was used as the basis for an analysis of evaluation strategies by intervention type and strength of evidence of effectiveness across intervention type and across the field as a whole. We recorded and tabulated both the direction of differences between groups (where reported) and statistical significance of differences. For controlled before and after studies, where there was no test of between group differences or groupby time interaction, this was categorised as a non-significant difference irrespective of a significant within-group difference. To inform the design of a future evaluation, we undertook a descriptive analysis of feasibility findings and other limitations identified in the medium- and high-quality studies we included.

Review findings

The review findings are presented here to address each of the review objectives in turn. First, we describe study characteristics to give an overview of studies that evaluate interventions for compassionate care. Second, we present an assessment of the quality of reporting of the interventions in the included studies, including their theoretical foundations. Third, we present evidence of effectiveness of the interventions in the included studies and analysis of the quality of that evidence.

Study characteristics

The final data set comprised 24 studies reporting 25 interventions. Twenty-two studies were published in journals[60–81] and a further two were doctoral theses.[82,83] Three types of intervention were identified. Staff training interventions ($n = 10$) focused on the development of new skills and knowledge in nursing staff, such as a training course in empathic skills communication. Care model interventions ($n = 9$) focused on the introduction of a new care model to a service, such as person-centred care. Nurse support interventions ($n = 6$) focused on improving nursing staff support and well-being through, for instance, the provision of clinical supervision.

Reports reflect a range of study settings including hospital ($n = 14$),[60,61,63,64,67,68,73,74,76,77,79,81–83] care/nursing homes ($n = 6$),[65,66,69–72] other community settings ($n = 3$)[62,75,78] and one study that used a range of health and social care settings ($n = 1$).[80] All but one of the staff training studies was conducted in a hospital setting, and six out of eight care model interventions were conducted in care home settings. Nurse support intervention studies were conducted in hospital settings ($n = 3$), district nursing services ($n = 1$), hospice at home ($n = 1$) and outpatient oncology service ($n = 1$). Eleven studies were conducted in the USA, with the other studies conducted in a range of other countries mostly in Europe but also including Australia, Canada, China and Turkey.

Study participants included nurses, nurse managers, patients and relatives. To evaluate the effect of the interventions a range of measurements was used, mainly self-reported instruments, but the effect was also proxy rated by researchers and using instruments based on researcher assessments of verbal communication and interaction. The outcomes measured in the studies varied widely, but could be classified into three types: nurse-based, quality-of-care and patient-based outcomes.

A table for each intervention type providing summary individual study characteristics and findings can be found in *Appendix 2*.

Quality of intervention reporting

Three types of intervention were identified: staff training, care model and nurse support. Interventions varied considerably in the extent to which they drew on an explicit theoretical foundation. Staff training interventions comprised training on verbal interactions, communication, communicating about spirituality and spiritual care, and empathy. Only four staff training interventions in the included studies had an explicit theoretical base. These were solution-focused brief therapy,[60] relationship-based care model/caring theories,[61] reminiscence theory and adult learning theory,[83] and the Tibetan Buddhist tradition.[62] Some interventions drew on definitions of particular concepts, such as empathy[63,64,82] and caring behaviours.[81] Other studies lacked an explicit theoretical foundation, referring only to results from previous research studies.

By contrast, all interventions introducing and testing a new care model were underpinned by an explicit framework. Most used theories or models developed in caring and nursing, except for one study using the *International Classification of Functioning, Disability and Health*[84] as the basis for an intervention to promote patient-centred communication with those living with aphasia/communication impairments.[65] Frameworks emphasised the person-centred care/environment/nursing,[66–68] the relationship between nurses and patients[69–71] or dignity in care.[72]

Nurse support interventions were based on reducing compassion fatigue, burnout and/or secondary traumatic stress;[73,74] and/or bolstering personal resources, such as compassion satisfaction, resiliency, empathy or sense of coherence.[73–75] Three were based on mindfulness theory.[76–78]

Reviewer ratings of the quality of intervention reporting in each study against each item in the Borek et al.[59] frameworkfor description of group-based behaviour change interventions are displayed in *Tables 2* and *3*. As is evident, the reporting of the interventions varied across all intervention types but was generally weak, with no intervention reports meeting all of the criteria deemed necessary for full intervention reporting. The design and the content of the interventions tended to be better described than details of the participants and the facilitators of the interventions. Overall compliance for intervention design reporting was 52% of criteria (shown in *Table 3* row labelled 'average % compliance by aspect of reporting'). The intervention design item with highest compliance (inclusion of details of the length of training sessions) was included in 73% (*n* = 16) of the 22 studies applicable here. The lowest was a specification of venue characteristics (*n* = 4, 17%).

For intervention content, highest compliance was reported for session content (*n* = 20 of the 21 applicable studies, 87%) and lowest for participants' materials (*n* = 8, 33%). Overall compliance for this aspect of intervention reporting was 50% of criteria. For reporting of participants, highest compliance was for description of group composition (*n* = 21, 88%) and lowest for continuity of participants' group membership (*n* = 3, 14% of 21 applicable studies). Overall compliance for this aspect of intervention reporting was 37% of criteria. For reporting of facilitators, highest compliance was for reporting facilitators' professional background (*n* = 12, 55% of 22 applicable studies) and lowest was for facilitators' personal characteristics and training in-group facilitation (both *n* = 1, 5% of 22 applicable studies). Overall compliance for this aspect of intervention reporting was 25% of criteria. On average, individual study compliance with the criteria was 42%, ranging from 8% to 65%. Of intervention types, care model interventions tended to be less well described than other types (average of 33% compliance).

Evidence of effectiveness

This section presents findings on the quality of evidence of effectiveness of the interventions in the included studies. Overall, methodological quality was low. Most studies either did not randomise to the groups and/or did not demonstrate equivalence between groups, weakening confidence in the findings. Only two studies were assessed as high quality and four as medium. The remaining 18 studies were assessed as low quality. Most studies (*n* = 16) were uncontrolled before-and-after studies. Four studies were before-and-after studies with separate intervention and control groups.[71,75,77,82] Four studies used a randomised controlled design. Three used a cluster randomised trial design, with clustering at unit or institutional level.[64,66,70] A further study was controlled but only included a post-test measure.[83]

Of the 24 studies, only eight studies included more than 100 participants. The largest sample included 115 nurses and 656 patients in an evaluation of an empathy-training programme.[64] The smallest sample included nine nurses in an evaluation of mindfulness based cognitive therapy for district nurses working with women with newly diagnosed breast cancer.[78] The number of clusters in controlled studies ranged from two to 38. Of the studies with experimental or quasi-experimental design,[64,66,70,71,75,82] just one[66] reported powering of sample size, but was not explicit about which outcome measure was the primary one used for these calculations.

Table 4 provides an overview of results from the individual studies against the range of outcomes used. Eighteen different types of outcomes were reported. For simplicity and brevity results for multiple measures using the same instrument or different instruments measuring same phenomena have been grouped together and treated as one. Across all studies and all outcome types results for 67 outcomes are reported. Further information on effect sizes is displayed in *Appendix 2* (see *Tables 41–43*).

TABLE 2 Systematic review completeness of intervention reporting: design and content

Intervention type	Study (first author and year)	Intervention source or development methods	General setting	Venue characteristics	Total number of group sessions	Length of group sessions	Frequency of group sessions
		Intervention					
		Design					
Training	Ançel 2006[63]	N	Y	N	Y	Y	Y
	Boscart 2009[60]	N	Y	N	N	Y	Y
	Glembocki 2010[61]	N	N	N	Y	Y	Y
	La Monica 1987[64]	Y	N	N	N	N	N
	Langewitz 2010[79]	Y	N	N	Y	Y	Y
	Puentes 1995[83]	N	Y	Y	Y	Y	N
	Searcy 1990[82]	N	Y	N	Y	Y	Y
	Taylor 2009[80]	Y	Y	NA	NA	NA	NA
	Wasner 2005[62]	Y	N	N	Y	Y	Y
	Yeakel 2003[81]	N	N	N	N	N	N
Care model	Brown Wilson 2013[69]	Y	Y	N	Y	Y	N
	Chenoweth 2014[66]	N	Y	N	N	N	N
	Finnema 2001[70]	N	Y	N	Y	Y	Y
	Ho 2016[72]	N	Y	N	N	N	N
	McCance 2009[67]	Y	Y	N	N	N	N
	McGilton 2003[71]	Y	Y	N	N	Y	N
	McGilton 2011[65]	N	N	N	N	Y	N
	Pipe 2010[68]	N	Y	Y	NA	NA	NA
Nurse support	Flarity 2013[73]	Y	N	N	Y	Y	Y
	Gauthier 2015[76]	Y	Y	Y	Y	Y	Y
	Horner 2014[77]	N	Y	N	Y	Y	Y
	Palmer 2010[78]	N	N	N	N	N	N
	Pålsson 1996[75]	N	Y	N	Y	Y	Y
	Potter 2013[74]	Y	Y	Y	Y	Y	Y
Compliance (%)		42	67	17	59	73	55
Average compliance by aspect of reporting (%)		52					

N, not reported; NA, not applicable to the intervention; Y, reported.
Adapted from Blomberg et al.[16] with permission from Elsevier. © 2016 The Authors. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

	Content						
Duration of the intervention	Change mechanism or theories of change	Change techniques	Session content	Sequencing of sessions	Participants' materials	Activities during the sessions	Methods for checking fidelity of delivery
Y	Y	Y	Y	N	N	Y	N
N	N	N	Y	Y	Y	Y	N
N	Y	Y	Y	N	N	N	Y
N	Y	Y	Y	N	N	Y	N
Y	Y	N	Y	Y	N	Y	N
N	Y	Y	Y	Y	Y	Y	N
Y	Y	Y	Y	N	Y	Y	N
N	Y	N	Y	N	Y	N	N
Y	Y	Y	Y	N	N	Y	N
Y	N	N	N	N	N	Y	N
N	Y	Y	Y	Y	N	Y	N
N	N	N	Y	N	N	N	N
Y	N	N	Y	Y	N	Y	N
N	Y	Y	Y	N	N	N	N
N	N	N	N	N	N	N	N
Y	Y	Y	Y	Y	Y	N	N
Y	Y	Y	Y	N	N	Y	N
N	Y	N	NA	NA	Y	NA	N
N	Y	N	Y	Y	Y	Y	N
Y	Y	N	Y	Y	Y	Y	N
Y	N	N	Y	N	N	Y	N
Y	N	N	N	N	N	N	N
Y	Y	N	Y	Y	N	Y	N
Y	Y	Y	Y	N	N	Y	N
54	71	46	87	39	33	70	4
	50						

TABLE 3 Systematic review completeness of intervention reporting: participants and facilitators

| Intervention type | Study (first author and year) | Intervention | | | | | |
| | | Participants | | | | Facilitators | |
		Group composition	Methods for group allocation	Continuity of participants' group membership	Group size	Number of facilitators	Continuity of facilitators' group assignment
Training	Ançel 2006[63]	Y	N	Y	Y	N	N
	Boscart 2009[60]	Y	NA	NA	NA	Y	Y
	Glembocki 2010[61]	Y	N	N	N	Y	N
	La Monica 1987[64]	Y	N	Y	Y	N	N
	Langewitz 2010[79]	Y	N	N	N	Y	N
	Puentes 1995[83]	Y	N	NA	NA	N	NA
	Searcy 1990[82]	Y	Y	N	N	N	N
	Taylor 2009[80]	Y	NA	NA	N	NA	NA
	Wasner 2005[62]	Y	NA	N	N	N	N
	Yeakel 2003[81]	Y	N	N	N	N	N
Care model	Brown Wilson 2013[69]	N	N	N	N	N	N
	Chenoweth 2014[66]	Y	Y	N	N	Y	N
	Finnema 2001[70]	Y	N	N	N	N	N
	Ho 2016[72]	Y	Y	N	N	N	N
	McCance 2009[67]	N	N	N	N	N	N
	McGilton 2003[71]	Y	Y	N	N	Y	Y
	McGilton 2011[65]	Y	N	N	N	Y	N
	Pipe 2010[68]	N	NA	N	N	NA	NA
Nurse support	Flarity 2013[73]	Y	N	N	N	Y	N
	Gauthier 2015[76]	Y	Y	N	N	Y	Y
	Horner 2014[77]	Y	N	N	N	N	N
	Palmer 2010[78]	Y	N	N	Y	N	N
	Pålsson 1996[75]	Y	Y	Y	Y	Y	N
	Potter 2013[74]	Y	N	N	N	Y	N
Compliance (%)		88	30	14	18	45	14
Average compliance by aspect of reporting (%)		37				29	

N, not reported; NA, not applicable to the intervention; Y, reported.
Adapted from Blomberg *et al.*[16] with permission from Elsevier. © 2016 The Authors. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Facilitators' professional background	Facilitators' personal characteristics	Facilitators' training in intervention delivery	Facilitators' training in group facilitation	Facilitators' materials	Intended facilitation style	Compliance (%)	Average compliance by intervention type (%)
Y	N	Y	N	N	Y	58	45
Y	N	Y	N	Y	N	57	
Y	N	Y	N	N	N	42	
N	Y	N	Y	N	Y	42	
Y	N	N	N	Y	N	50	
N	N	N	N	Y	Y	57	
N	N	N	N	N	Y	50	
NA	NA	NA	NA	NA	NA	50	
N	N	N	N	N	N	40	
N	N	N	N	N	N	12	
N	N	N	N	N	Y	38	33
Y	N	Y	N	N	N	27	
N	N	N	N	N	N	35	
Y	N	N	N	N	N	27	
N	N	N	N	N	N	8	
Y	N	Y	N	Y	N	62	
Y	N	N	N	N	N	35	
NA	NA	NA	NA	NA	NA	36	
N	N	N	N	N	N	42	46
Y	N	N	N	N	N	65	
Y	N	N	N	N	N	35	
N	N	N	N	N	N	12	
Y	N	Y	N	N	N	62	
Y	N	Y	N	N	N	58	
55	5	32	5	18	23		

TABLE 4 Systematic review summary of study results and statistical conclusions by outcome type

Study (first author and year)	Study quality	Outcome type								
		Nurse								
		Empathy	Compassion	Burnout	Stress	Mindfulness	Job satisfaction	Caring	Attitude	Other well-being
Training intervention										
La Monica 1987[64]	Medium	–								
Searcy 1990[82]	Medium	–								
Ançel 2006[63]	Low	▲								
Boscart 2009[60]	Low									
Glembocki 2010[61]	Low							▲		
Langewitz 2010[79]	Low									
Puentes 1995[83]	Low	▲							▲	
Taylor 2009[80]	Low	▲							▲	
Wasner 2005[62]	Low		▲		▲	▲			▲	▲
Yeakel 2003[81]	Low									
Care model intervention										
[a]Chenoweth 2014 (single)[66]	High									
[a]Chenoweth 2014 (combined)[66]	High									
Finnema 2001[70]	High									
McGilton 2003[71]	Medium									
Brown Wilson 2013[69]	Low							△		△
Ho 2016[72]	Low									
McCance 2009[67]	Low							▲		
McGilton 2011[65]	Low								▲	
Pipe 2010[68]	Low									
Nurse support intervention										
[b]Pålsson 1996[75]	Medium	–		–						–
Flarity 2013[73]	Low		▲	▲	▲					
Gauthier 2015[76]	Low		△	–	▲	△	△			
Horner 2014[77]	Low		–	–	–	–				
Palmer 2010[78]	Low	△				△				△
Potter 2013[74]	Low		△	△	▲		–			

▲, significant improvement; △, non-significant improvement; –, no change; ▽, non-significant deterioration.
PCC, person-centred care; PCE, person-centred environment.
a Chenoweth *et al.*[66] compared effectiveness of three interventions: two single (PCC and PCE) and one combined (PCC and PCE implemented together).
b Pålsson *et al.*[75] tested difference between the two groups, found differences in either direction, but not significant either within or between groups, and we have reported this as 'no difference'.
Adapted from Blomberg *et al.*[16] with permission from Elsevier. © 2016 The Authors. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

| Quality of care | | | | | Patient | | | |
Quality of interactions	Quality of relationship	Patient centredness	Continuity	Quality of care	Satisfaction/ experience	Behavioural (agitation)	Quality of life	Mood/ well-being
					–			▲
	–				–			
▲								
▲								
	▲				▲			
–						▲	△	△
▲						▽	△	△
				▲				
	▲		▲					
					△			△
							▲	
		▲						
	▲						△	▽
								▲
					△			

Studies of similar intervention types tended to use similar outcome types. Nurse support intervention studies primarily measured nurse-based outcomes. No nurse support studies used quality-of-care outcomes, and just one study used patient-based outcomes. In contrast, care model intervention studies primarily used outcomes related to quality-of-care and patient-based outcomes, but use of nurse outcomes was less common. Training intervention studies used the widest range of outcome types. Although the majority used nurse-based outcomes a small number drew on quality-of-care and patient outcomes.

Nineteen studies (79%) reported a significant positive difference in one or more outcomes (i.e. a beneficial effect of the compassionate care intervention). Only five (21%) of the 24 studies reported no significant difference in any of the outcome types measured. Of the 67 outcome types assessed across all studies, 32 (48%) showed significant positive effects for the intervention, with a further 18 (27%) showing positive but non-significant results. There were no significant negative differences and only three non-significant negative results. Patient outcomes were less likely to show significant differences, with only 5 out of 17 (29%) showing statistically significant differences.

Studies of low methodological quality were more likely to report outcomes in favour of the intervention, with low methodological quality studies reporting a mean of 92% of outcomes in favour of the intervention (significant + non-significant positives), whereas higher-quality (medium, high) studies report 55% of outcomes in favour of the intervention. Although an average of 76% of outcomes reported in studies of training interventions showed a statistically significant benefit, only 21% of outcomes for nurse support interventions were significant. Crucially, no intervention has been evaluated more than once.

Effects on patient-based outcomes

Six care model intervention studies reported patient-based outcomes. Of these, three showed statistically significant effects on a patient-based outcome. Of these, one was rated as high quality. In their CRT with 38 nursing homes, Chenoweth et al.[66] reported that the person-centred care intervention had a significant positive effect on reducing patient agitation, but the combined intervention (person-centred care plus person-centred environment) reported in the same study showed a non-significant effect of increasing patient agitation. This study fared poorly in terms of reporting of intervention description, meeting only 27% of criteria.

Three training intervention studies reported patient-based outcomes and, of these, two showed a significant positive effect. One medium-quality study reported significant positive effects on patient anxiety[64] and one low-quality study reported a non-significant positive difference in patient satisfaction.[81] A low-quality nurse support intervention study reported a non-significant improvement in patient satisfaction.[77]

Effects on quality-of-care outcomes

Four[60,79,81,82] training intervention and five[65–67,70,71] care model intervention studies (one of which[66] reported two interventions) examined effects on quality-of-care outcomes. Of these, eight[60,65–67,70,71,79,81] reported a statistically significant improvement in one or more outcomes. The combined person-centred care model intervention reported by Chenoweth et al.[66] was associated with a significant improvement in quality of interactions, but, although this finding is from a high-quality study, conclusions are tempered by the lack of intervention description noted above. In a CRT rated as high quality,[70] the authors reported a significant change in one dimension of quality of care following implementation of emotion-oriented care in nursing home settings, but the intervention description met only 35% of the criteria. A medium-quality evaluation of a relationship-enhancing programme of care in nursing homes[71] reported significant improvements in relational care, care providers' relational behaviour and continuity of care. A medium-quality evaluation of empathy training for hospital nurses[82] found no difference in interpersonal support. Other improvements in quality-of-care outcomes were reported in a range of low-quality studies.[60,65,67,79,81]

Effects on nurse-based outcomes

Seven[61–64,80,83] training, six[73–78] nurse support and three[65,67,69] care model intervention studies examined effects on nurse-based outcomes and, of these, 10[61–63,65,73,74,76,80,83,98] reported a significant improvement associated with the intervention. All of these 10 studies were rated as low quality. Three medium-quality studies[64,75,82] investigated nurse-based outcomes but none showed significant differences. No high-quality studies reported on nurse-based outcomes.

Feasibility findings

Findings from our analysis of feasibility issues and limitations documented in the reports of high and medium-quality studies ($n = 6$) are summarised here. The included studies were either experimental in the form of a cluster RCT or quasi-experimental with before-and-after measurements of intervention and control groups, but no randomisation to groups. Papers varied in the feasibility findings they reported and in the limitations identified but all were able to identify where improvements could be made in future research of this kind.

Two studies[64,82] in this subset were conducted in a hospital setting, both single-site hospital US settings. La Monica *et al.*[64] conducted a cluster RCT in four cancer units (two medical and two surgical) to determine the impact of a nurse empathy training programme on patient outcomes (anxiety, depression, hostility, satisfaction with care) and nurse outcomes (nurse empathy). The study was not focused on older people, and patients too ill or confused to complete the questionnaires were excluded. Baseline data were gathered over a 4-week period, followed by a 4-week empathy training delivery period, followed by a 4-week follow-up assessment period. La Monica *et al.*[64] reported that patients were not admitted for long enough to take part in both assessment periods of the study. Patient participation rate was reported to be 73% with reasons for non-participation including not feeling well enough, having a conflict with a treatment or personal schedule, being reluctant to rate the nurse, and generally not being interested. The research team also identified a number of issues with the outcome measures involving rating nurse empathy and satisfaction with care. The team noted that patients consistently rated nurses' empathy higher than nurses rated their own empathy, and speculated that it may be psychologically threatening for patients to rate nurses and nursing care, particularly while still in need of care.[64] In addition, at baseline and follow-up in both experimental groups, nurse- and patient-rated empathy scores were close to maximum, implying a ceiling effect.

Searcy[82] conducted a before-and-after study with an intervention group (one ward) and a control group (one ward) to determine the impact of an empathy education programme for hospital nursing staff on patient satisfaction with care, including perceptions of interpersonal support. Baseline data were gathered over a 6-month period, followed by a 2-week training period (consisting of two 1-hour sessions), followed by a 6-month follow-up assessment period. The patient survey was mailed 1 week after discharge to all patients discharged from the two participating units. All patients were adults and no exclusion criteria, such as dementia, were reported.[82] Survey response rate was reported as approximately 25%. Searcy[82] reported that baseline ratings were high, implying a ceiling effect to the chosen measure, and also found that older patients rated higher satisfaction than younger patients.

Three studies[66,70,71] in this subset evaluated the impact of care model interventions in care home settings in Australia, the Netherlands and Canada. All were focused on interventions targeted at improving care for older residents, with two particularly focused on dementia care.[66,70] One study involved different sets of residents in questionnaires and in observations.[71] The questionnaire subsample included people who were medically stable, able to understand English and cognitively able to provide consent and respond to questions.[71] Just five out of the original subsample ($n = 50$) did not participate in the follow-up assessment period (10 months after intervention start), suggesting a relatively healthy subsample.[71] The other subsample included people who required moderate to high levels of assistance with personal hygienic care. Residents recruited to this subsample were included in observations of care carried out at baseline and follow-up.[71] However, whether or not the subsamples in the two time periods were independent of each other was not reported, so attrition rates cannot be established.[71] The two other care-home studies reported high attrition

rates.[66,70] Chenoweth *et al.*[66] retained 36 out of 38 nursing home clusters, and 305 out of 601 nursing home residents remained in the study over 8 months. Most (i.e. 216 out of 296) attrition was attributable to residents dying. Finnema *et al.*[70] reported that 132 eligible nursing home family members completed a second questionnaire 8 months following completion of a first questionnaire by 241 family members. Forty-four people were not included in the second round because their relative had died.[70]

Two of the care home studies used observations of care but neither study reported feasibility issues with the observations.[66,71] Feasibility issues were, however, more apparent when self-report and/or proxy questionnaires were used, especially when cognitive impairment was more severe. As noted above, McGilton *et al.*[71] limited questionnaire distribution to relatively healthy residents and did not report any feasibility issues. Chenoweth *et al.*[66] used a range of instruments to measure resident well-being, and concluded that a lack of association between the intervention and quality of life, and depression, may be attributable to the difficulty in measuring these constructs through self-report and/or proxy in this population. Finnema *et al.*[70] invited family members to complete a newly developed questionnaire but not all items were completed, suggesting a need for further development and piloting.

The final study in this subset investigated the impact of clinical supervision (a form of nurse support) on burnout, empathy and sense of coherence in district nurses in Sweden.[75] Improving the care of older people and/or people living with a cognitive impairment was not a stated focus of the study. Measures used were all deployed through a written questionnaire completed by staff. In total, 33 out of 39 district nurses remained in the study over 28 months. The authors reflected that the high baseline nurse-reported empathy and sense of coherence, and low burnout, may indicate a ceiling effect to the selected outcome measures.[75] Empathy was measured using the Empathy Construct Rating Scale,[85] also used by La Monica *et al.*[64] in the hospital study reported above. The finding in both studies that there may be a ceiling effect to measuring empathy in this way suggests a limited capacity to measure improvements in empathy over time.

Of the six studies,[64,66,70,71,75,82] three reported it was not possible to mask nursing staff to experimental allocation.[64,66,75] Two studies reported concealing experimental allocation from the research team.[66,71] Two studies deployed measures to control contamination of control conditions[66,71] and one study identified pathways through which contamination may have occurred.[82]

Only Chenoweth *et al.*[66] formally measured fidelity to the intervention, and they reported variation in implementation between clusters, suggesting that the time frame of the study limited implementation (post-intervention assessment was at 6 and 14 months following planned start of intervention implementation). In one of the two hospital studies, La Monica *et al.*[64] reported that all eligible nurses attended the training but did not report if the training was delivered as planned. They speculated that the follow-up assessment period may have been timed too early (4-week assessment period following 4-week training period), and did not allow the results of the training to embed into practice. In the other hospital study, attendance at training or any aspect of fidelity went unreported.[82] Searcy[82] reported that staff attending the training fed back that 2 hours of training was insufficient. Pålsson *et al.*[75] and McGilton *et al.*[71] speculated that the intervention period may have been too short to effect change (10 months and 7 months, respectively). Pålsson *et al.*[75] did not report if clinical supervision was delivered as planned but did report number and length of sessions, and attendance at sessions (74%). McGilton *et al.*[71] reported that the protocol was adhered to in intervention delivery and 70% of care providers and 50% of supervisors attended the whole programme. Finnema *et al.*[70] did not report assessing if the staff training and supervision was delivered as planned, but did use findings from a staff self-assessment of nursing skills to infer that the observed increase in emotion-oriented skills meant that emotion-oriented care was applied more often than before. Five studies[64,66,71,75,82] speculated that the impact of the intervention may have been affected by other factors, such as leadership support of the intervention and by other initiatives taking place at the same time as the intervention.

In summary, previous experimental and quasi-experimental work in this field has raised a number of potential feasibility issues that can inform future study design and implementation. Hospital settings presented a number of validity issues including a possible reluctance of patients to rate nurses negatively while still in receipt of care and/or a ceiling effect to empathy rating instruments deployed to date. One of the hospital studies did not include patients who were 'confused'[64] and the other relied on a post-discharge written survey,[65] and so neither study illuminated how best to include people with a cognitive impairment in care evaluations. Response rates were markedly higher for the study surveying people while still an inpatient than for the study surveying people post-discharge. In care-home studies more inclusive of people with cognitive impairment, feasibility issues were identified with written questionnaires, and methods involving observations were not associated with feasibility issues. Most studies did not pay attention to fidelity of actual delivery of the planned intervention. Some studies reported that the intervention may have been delivered in too small a dose, and/or that follow-up assessment may have occurred too rapidly. Five[64,66,71,75,82] studies speculated on the importance of a variety of contextual factors in affecting the implementation process but data were not gathered to enable in-depth exploration of this potential influence.

Discussion

This chapter aimed to provide an overview of the evidence base on the evaluation of interventions for compassionate nursing care. It began by reporting on the qualitative work in this field, work that has focused on relationship-centred approaches to improving care, interventions that we would classify as a combination of care model and staff support intervention types. These studies often provide detailed information about the intervention and its inherent mechanisms for change, and often include analyses of the factors enabling and inhibiting sustainable change. However, they do not examine in depth the process of implementation itself and commonly pay only superficial attention to the influence of actors, networks and resources on the impact and sustainability of these interventions. In addition, as would be expected with a qualitative approach, there is only weak objective evidence of effectiveness of the interventions deployed in these studies in relation to impact on patient outcomes.

As identified in our systematic review, there is a wide range of intervention studies in which compassion has been addressed in a variety of ways, including through staff training, staff support or introducing a new care model. However, the overall strength of work in the field limits the conclusions that can be drawn to inform health-care policy and practice. No study reported in the systematic review included sufficient detail of its intervention to enable replication and further evaluation. This state of play limits the capacity of nurses and others to include effective strategies in their own practice, but also limits the construction of a coherent evidence base to guide managers and practitioners in improving services.[86–88]

In relation to the nature and strength of the existing evidence base, overall quality of the evidence was low and it appears that the few higher-quality studies are less likely to report positive results. No intervention was tested more than once and the majority of studies used before-and-after designs that are intrinsically weak. Patient-based outcomes were not routinely included, especially in relation to the evaluation of training interventions and nurse support interventions.

Our analysis did, however, highlight the feasibility issues associated with this field of research that have formed the backdrop to the design and implementation of the study reported in subsequent chapters.

Conclusions

Although there have been many published studies that appear to offer potential solutions to deficits in compassionate care, this is a body of literature that does not offer a definitive way forward for policy and practice in this important area of health care. This is especially challenging in a context in which the need for more compassion in health care is professed from national government to frontline practitioners.

Greater conceptual clarity, better designed and reported interventions, including a focus on implementation processes, and evaluations using stronger research designs are urgently required.

To date, no evaluations of initiatives of this kind have enabled a robust assessment of the effectiveness of interventions on patient care, linked with the use of theory-based interventions reported with sufficient clarity to support optimal implementation, impact and sustainability. The study reported in the following chapters is the first stage in designing a rigorous mixed-methods evaluation incorporating experimental design to understand what works best in improving care.

Chapter 3 Research objectives

This study aimed to assess the feasibility of implementing CLECC in acute hospital settings, and to assess the feasibility of conducting a CRT with associated process and economic evaluations to measure and explain the effectiveness of CLECC.

The objectives were:

1. to determine the feasibility of implementing the CLECC intervention and sustaining the resulting work practices
2. to inform the design of a definitive evaluation of the effectiveness of CLECC
3. to inform the measurement of costs and benefits of CLECC in a definitive evaluation.

As will be detailed in the chapters that follow, the first objective focused on exploring the extent to which the planned CLECC intervention could be made workable and integrated into routine practice, to enable conclusions to be drawn about how it can be optimised in the future to support sustained compassionate care delivery in acute settings. The second objective focused on gathering data to inform the future measurement of the effectiveness of CLECC in supporting compassionate care delivery. The third objective focused on identifying the optimal methods for measuring the costs and benefits of CLECC.

Chapter 4 Methodology

This chapter introduces the key methodological elements of the study, in particular the study design, feasibility parameters being tested and outcome measures assessed. Information is also given about patient and public involvement (PPI), ethics considerations and changes from original protocol. The study protocol can be accessed at: www.journalslibrary.nihr.ac.uk/programmes/hsdr/130748/#/.

Study design

This mixed-methods study used two main approaches to assess feasibility: (1) a process evaluation to enable evaluation of the feasibility of implementing the CLECC intervention and (2) a pilot pragmatic CRT to lay the foundation for a future evaluation of CLECC's effectiveness. The design draws on the Medical Research Council (MRC) guidelines for evaluating complex interventions.[89] The MRC guidelines highlight the importance of a robust theory-based intervention design coupled with a thorough understanding of the mechanisms of change and the contexts in which implementation is possible.[89] It is essential to undertake this design and evaluation prior to and during evaluation of effectiveness, so that the impact of the intervention can be optimised, and so that eventual findings on outcomes can be explained by how change happened (or not). In addition, the guidelines also highlight the importance of detailed groundwork on the implementation of experimental design and methods before a definitive evaluation of effectiveness is undertaken. This includes selection and testing of outcome measures; testing the feasibility of proposed methods for recruitment, data collection and analysis; and calculations of effect size to inform future sample size calculations.

Areas of feasibility that merited study here included implementation of the CLECC intervention into practice, contamination of practices from intervention to control wards, extent of intraward clustering of outcomes, rates of participation and attrition, and the performance of the selected outcome measures. Through piloting these procedures on six wards in two English hospitals, this study aimed to reduce uncertainty in designing and executing a definitive evaluation. We aimed to test a range of parameters to provide an evidence base from which to make decisions about evaluation design to be implemented in other centres in a future trial, including CLECC implementation, feasibility of ward-level randomisation, contributing to sample size calculation and selection of outcome measures.

The feasibility of implementing the CLECC intervention into practice was assessed through a process evaluation, using normalisation process theory (NPT) as a framework.[90] Qualitative interviews with nursing staff and managers during implementation and follow-up phases, observations of learning activities, and ward leader questionnaires aimed to identify and explain the extent to which the CLECC intervention was implemented into practice, enabling an assessment of its workability and integration into existing work practices. In addition, data were also gathered on the feasibility of conducting qualitative interviews with staff, patients and visitors to inform qualitative evaluation accompanying a future definitive trial, with the purpose of explaining trial outcomes.

An experimental design in the definitive evaluation is the most appropriate design to establish effectiveness.[91] In order to prepare for a definitive multicentre evaluation, we undertook a randomised pilot study to assess the feasibility and pilot procedures for a pragmatic CRT of effectiveness. In randomised pilot studies, all or parts of the future trial are conducted on a smaller scale to see if it can be done.[92] Given that the intervention was targeted at a group of staff rather than at the care of individual patients, cluster randomisation of staff and patients at nursing ward level was undertaken.[91]

Ward teams (i.e. clusters) were randomly allocated following baseline data collection to participate in the intervention or act as control. We theorised that implementing CLECC on four wards across two hospital sites was likely to provide sufficient diverse contexts within which its feasibility could be assessed, leading

to further refinement when needed. The inclusion of a small number of control wards aimed to generate insight into the likely acceptability of randomisation in the main trial and differential compliance with study procedures between intervention and control.

The CLECC intervention was implemented over a 4-month implementation period on four wards from June 2015, with two wards acting as control. Outcomes were assessed at baseline (2 months before intervention and before randomisation to groups) and 4 months after completion of the 4-month CLECC implementation period (i.e. follow-up was 8 months after randomisation). There is no single validated measure for compassionate care and we assessed its measurement across three complementary core outcomes: (1) researcher-rated observations of the quality of staff–patient interactions, (2) patient-reported evaluations of emotional care and (3) nurse-reported measures of empathy. Baseline and follow-up data were also gathered on individual and ward team characteristics.

The economic component of the study aimed to explore how costs and benefits might best be measured in a definitive evaluation. Standard health technology assessment economic evaluations rely on patient-based outcome measures, the feasibility of which was explored in this study by use of EuroQol-5 Dimensions, five-level version[93] (EQ-5D-5L), (a health status measure) at the ward level. In addition, the study explored the likely costs of both the CLECC intervention (training) and its implementation (qualitative interviews with nursing staff on intervention wards).

Data collection took place between March 2015 and May 2016.

Process evaluation

The process evaluation aimed to identify and explain the extent to which the planned intervention was implemented into existing nursing practices, to enable conclusions to be drawn about how CLECC can be optimised to support sustained compassionate care delivery in acute care settings. NPT was used to guide the process evaluation, shaping the interview topic guides and informing the framework for analysis.[90] NPT focuses on the dynamic processes that support the integration of innovative practices into everyday work, and so is a helpful way to evaluate what actually happens when complex interventions are introduced into practice, and how and why the desired outcomes are achieved (or not).[94] It explores social processes and the work that people do individually and collectively, in terms of cognitive and behavioural work. NPT has four core constructs: (1) coherence – making sense of the intervention, (2) cognitive participation – investing in the intervention, (3) collective action – the practical work of implementation and (4) reflexive monitoring – modifying and embedding the intervention.[95] These constructs were used to define the areas that formed the focus of the evaluation.

The process evaluation focused on:

1. exploring how and in what ways the new practice was initially received, how individually and collectively people practically conceptualise and make sense of it (coherence)
2. assessing the degree of ownership of and participation in the new practice by key individuals and teams (cognitive participation)
3. identifying the work that individuals and teams do to enact the new practice (collective action)
4. exploring the perceived impact of the new practice on staff work and on patient outcomes (reflexive monitoring).

Exploration of these areas was aimed at informing conclusions about how CLECC could be optimised for impact and sustainability beyond the implementation period.

Pilot cluster randomised trial outcome measures

There is no single validated measure for compassionate care and we assessed its measurement across three complementary core outcomes at cluster level and at individual participant level: researcher-rated observations of the quality of staff–patient interactions, patient-reported evaluations of emotional care and nurse-reported measures of empathy.

Quality of staff–patient interactions

The quality of staff–patient interactions was assessed using the QuIS. QuIS is a time-sampling tool that measures both the volume and quality of interactions, through observation, enabling a calculation of how many patients experience one or more negative interactions during an observation session.[42] QuIS interactions between staff and patients are coded as positive social, positive care, neutral, negative protective or negative (*Table 5*).

Although other observational tools have been developed for educational and service improvement purposes, they have not been validated as research instruments. QuIS has been used in a number of studies in NHS acute care settings for service improvement and evaluation, including use by the Health Advisory Service in their seminal evaluation 'Not because they are old'.[96] Other work has demonstrated that it is sensitive to changes in service quality.[42,97,98] The selection of this outcome reflects a concern to measure patient-based outcomes rather than, as is often the case in compassionate care intervention evaluations, staff-based outcomes or process measures.[16] Because it does not require any capacity to perform by patients, it is also potentially inclusive of people who would usually be excluded from research, for example people with dementia, people who do not speak English and people with communication difficulties.

Inter-rater reliability studies for QuIS have generally reported high levels of agreement, but these studies largely tested reliability by asking a second rater to categorise interactions based upon written descriptions by the first observer.[99] In contrast, Dean *et al.*[42] tested rating reliability between two observers present during interactions (in a long-term care setting) and found that agreement was more variable than reported in other studies, although still acceptable ($\kappa = 0.60$–0.91). Prior to our own work, no other studies have examined the reliability of the QuIS in acute care settings using a method similar to that used by Dean *et al.*[42] No studies prior to our own research in acute care have directly demonstrated a relationship between QuIS ratings and patient experience.

The QuIS was originally designed for long-term settings, and so extended acute care definitions of the original five QuIS categories were developed (see *Table 5*) and tested for their validity and reliability by the team in early piloting work, together with guidance for using the instrument in acute care and an associated training protocol (see *Appendix 3*).[99] The results from this early work are reported in *Chapter 9, Quality of Interactions Schedule*.

TABLE 5 Definitions of QuIS categories[42]

Category	Explanation
Positive social	Interaction principally involving 'good, constructive, beneficial' conversation and companionship
Positive care	Interactions during the appropriate delivery of physical care
Neutral	Brief, indifferent interactions not meeting the definitions of the other categories
Negative protective	Providing care, keeping safe or removing from danger, but in a restrictive manner, without explanation or reassurance in a way that disregards dignity or fails to demonstrate respect for the individual
Negative restrictive	Interactions that oppose or resist people's freedom of action without good reason, or which ignore them as a person

Patient-reported evaluations of emotional care

Patient-reported evaluations of emotional care were measured using the Patient Evaluation of Emotional Care during Hospitalisation (PEECH) survey tool.[100–102] Although a number of survey instruments are now available that measure patient experience, most are limited in their capacity to assess patient experiences of the more complex relational aspects of care.[100] Designed to address this gap, the PEECH tool focuses on the nature of interpersonal interactions with hospital staff and patient-reported assessment of the extent to which therapeutic emotional care has occurred.[101,102] Originally developed in Australia, PEECH has since been validated for use in English hospital settings and can be completed by patients during a hospital stay with or without assistance.[100] The subscales are security, knowledge, personal value and connection. PEECH is sensitive to changes in service quality and in ward environment.[103]

Nurse-reported empathy

Nurses' self-reported empathy was measured using the Jefferson Scale of Empathy[104] (JSE) (physician/ health professional version), a 20-item inventory in a seven-point Likert-type format ranging from strongly disagree to strongly agree, with higher scores reflecting a more empathic orientation.[104] Although caregiver empathy is recognised as an important component of compassionate care, the JSE is the only scale focusing on this concept that is designed for use in patient care contexts. Developed and validated for use by health-care workers, including nurses, the scale is sensitive to changes in individual empathy over time and context.[105,106]

Ward team characteristics

Baseline and follow-up data were also gathered to enable a description of the characteristics of the ward nursing teams involved in the study. These included qualitative interviews with nursing staff and the administration of a number of instruments through written survey. We assessed staff local working climate using Climate for Care (CC) and Factors that Enable Climate for Care (FECC) questionnaires,[26] 39-item and 19-item questionnaires, respectively, with answers on a 5-point Likert scale developed as part of a toolkit from a NIHR-funded project measuring culture change and quality of NHS acute hospital care for older people, with the ability to identify distinct nursing team climates.[26] We also administered the Matron's Assessment of Quality of Care (MAQC) and the Carer Experiences of Care (CEC) questionnaires from the same toolkit.[26] We assessed nursing staff perceptions of workload using items from the International Hospital Outcomes Study (IHOS) battery,[107,108] including (1) enough nurses on staff to provide quality patient care and (2) ratings of core care activities that were deemed necessary but left undone. This survey has been widely validated internationally and subjective ratings from it correlate with objective measures of both staffing and quality. We also measured levels of nursing staff burnout using the 22-item Maslach Burnout Inventory (MBI).[109]

Pilot cluster randomised trial randomisation

Ward teams (i.e. clusters) were randomly allocated following baseline data collection to participate in the intervention or act as a control. Randomisation was stratified by hospital and by ward type [medicine for older people (MOP) or surgical]. Randomisation was accomplished using the *ralloc* command in Stata® 12 (StataCorp LP, College Station, TX, USA) and conducted by the study statistician. We planned for three wards (one surgical and two MOP) in each hospital. Wards were allocated to achieve two intervention and one control in each hospital, with one of the surgical wards allocated to intervention and the other to control. This strategy was to ensure that intervention implementation could be tested in two ward contexts in each hospital, and also to ensure that we gained experience of intervention and control conditions in both MOP and non-MOP ward specialties. We performed simple randomisation. First, we allocated the surgical wards to control or intervention. The MOP wards of the hospital with the surgical ward allocated to control were both intervention wards and, therefore, we did a second randomisation to allocate the MOP wards in the second hospital to intervention or control.

Pilot cluster randomised trial allocation concealment

Ward teams (i.e. clusters) were identified and recruited before randomisation. Clusters were randomly allocated to a group following baseline data collection by team members not involved in data collection. At follow-up, researchers conducting observations of the quality of staff–patient interactions were recruited from outside the core research team and not informed of allocation. Researchers gathering questionnaire data at follow-up and involved in qualitative interviewing were aware of ward allocation. It was not possible to conceal allocation from ward team nursing staff. Patients and visitors were not informed of allocation. The success of allocation concealment was tested as part of the feasibility work.

Progression to a definitive evaluation

As outlined above, the study was designed to lay the groundwork for a future definitive evaluation of the CLECC intervention. In the first instance, we planned that the process evaluation would enable an assessment of the CLECC intervention's workability and integration into existing work practices, providing important information to guide its further refinement and implementation. Second, findings from the study would enable assessment of the feasibility of a future definitive evaluation against a number of important parameters. Results from the assessment of these parameters would then inform the design and implementation of a definitive evaluation, including sample size, level of clustering and selection of outcome measures.

The study took place on busy acute care wards and our target samples included frontline nursing staff and patients and visitors during the period of their hospitalisation. We set an explicit goal of maximising the participation of patients often excluded from research, that is, older people with complex needs including cognitive impairment and communication difficulties, as it is often people with these characteristics who are in greatest need of compassionate care. The study enabled us to develop and test a number of approaches to successful recruitment and participation. However, there is little in the literature to guide the recruitment of these groups in acute care settings, especially in relation to experimental studies. It was therefore not possible at the outset to quantify fully the study's success criteria in relation to the number of patients or staff to be recruited. We instead set target recruitment rates that were reviewed and refined as data were collected and analysed.

The success criteria set for the study were:

1. completion of a process evaluation into CLECC's workability and integration into existing work practices, sufficient to inform refinement and future implementation, and to inform future process evaluations
2. recruitment of sufficient wards ($n = 6$) to assess the feasibility of a CRT design to inform the design of a definitive evaluation, including information on participation and attrition rates, blinding strategies, mitigation of contamination, baseline rates and intracluster correlations for core outcomes, data collection and analysis procedures
3. recruitment or refinement of target numbers of staff, patients and carers to enable collection of data estimated to be sufficient to inform the selection and use of primary and other outcome measures in definitive evaluation.

Study progress against success criteria was externally monitored by the Study Steering Committee. A Study Advisory Group advised on the CLECC intervention implementation and a PPI group also advised the study team.

Patient and public involvement

Our consultative PPI work over 3 years on this topic confirms public recognition of the need for research of this kind and that the research addresses a topic of primary concern to the general public.

We have consulted with service users about this research since 2013, while we developed the CLECC intervention, and throughout this feasibility study. We recruited older members of the public who had experienced a hospital admission, or had been a carer of someone who had experienced a hospital admission. A PPI group ($n = 5$) started meeting in February 2013 to help develop the application and plan staff training for the feasibility study, and two members sat on the Study Steering Committee. A leading figure (Lesley Carter) from national charity Age UK (London, UK) also sat on our Study Advisory Committee. The initial PPI group input helped to identify priorities for the research outcomes, clearly showing that improvements to compassion in hospitals, and the involvement of patients and family in care were important outcomes for the PPI group. They also advised on the development of the CLECC intervention and its evaluation. Changes made as a result of PPI input include extending classroom sessions (part of CLECC intervention) to include registered nurses as well as care assistants, developing researcher guidance on approaching family carers to participate in research, and developing ways to include people with communication difficulties in the research. We also included PPI input in training our core research team. One of our PPI representatives (Jan Gollop) provided a half-day of training to research team members on conducting research on people with dementia. Her expertise from caring for her husband with severe dementia, and from her network of other carers, and experiences visiting hospitals gave her a unique insight into what researchers needed to know.

We also held a public consultation event in November 2015 to get feedback on the work to date, test our plans for definitive evaluation and recruit more PPI input. People who attended this event ($n = 6$) provided advice regarding the dissemination of feasibility findings, which informed the dissemination plan for both the feasibility study and the definitive evaluation proposed as the next step. People who came discussed the 'value' of a definitive study of CLECC, in particular its value for money in times of austerity, and concluded that compassion in hospitals was of such personal significance that a definitive evaluation did merit funding. These consultations have confirmed strong public support for the CLECC programme of work. The study addresses a topic of primary concern to the general public and our consultees have demonstrated very strong support for this work. Involving PPI representatives from both hospital sites has been particularly helpful in involving members of the public whose locality and experiences of the hospital in question help ensure not only patient involvement but also local involvement.

Ethics considerations

Our concern in this study was to keep the best interests of participating patients, visitors and NHS staff at the centre of what we did. We included a number of measures to help ensure this and carefully consulted with our PPI group and nursing representatives about our proposals over a 2-year period. We aimed for a proportionate approach that did not place undue burden on participants at any part of the process and also represented what we judged to be achievable as a research team working within limited resources. There was also the opportunity to pilot and adapt procedures within the feasibility framework.

One key ethics issue was the recruitment of and proposed data collection from patients, staff and visitors at what can be an already stressful time. We addressed this by developing concisely written information accompanied by verbal explanation and the chance to ask questions. We ensured that we included in written and verbal information the clear statement that people were not obliged to take part and that their care or treatment would not be affected in any way if they declined to take part or withdrew. We allowed people as much time as they needed to make up their minds about taking part. We ensured that research team members had the skills to identify distress caused by recruitment and/or data collection processes and had clear plans of action to follow if this happened.

Another ethics issue was the desire to include people who lack the capacity to make the decision to take part in the research. This is important because this group is often excluded from research and yet evidence suggests that they are most vulnerable to not experiencing compassionate care. We developed and implemented clear procedures to ensure that the principles of the *Mental Capacity Act 2005 Code of Practice*[110] and process consent were adhered to.[111,112]

A further potential issue was the participation of ward staff as research subjects and the concern that may be raised about their rights to refuse to take part or to withdraw from the study. Our communication strategy aimed to ensure that everyone who should know about the study was informed about it and their right to not take part. Researcher training emphasised issues relating to the anonymity of staff, visitors and patients. Clear procedures were developed to guide the reporting of unsafe practice.

Ethics approval for the study was granted by the national Social Care Research Ethics Committee (REC) (reference number 14/IEC08/1018) in December 2014. We originally applied for NHS REC approval, but were advised by the manager for the local REC that the study did not merit NHS REC scrutiny because the proposed intervention was not clinical. We appealed this decision, but it was upheld by the Health Research Authority regional manager. We requested an alternative, but equivalent form of review and were referred to the Social Care REC.

Research team and training for data collection

The core research team involved in data collection were the chief investigator (JB), two research fellows (LG and WW), the senior research assistant (HB) and the research assistant (PL). They conducted the qualitative interviews and were supported in screening, recruitment and data collection by others. Research nurses and clinical trials assistants ($n = 18$ over the course of the whole study) at the participating NHS trusts screened and recruited patients for observations. Research nurses, clinical trials assistants and eight other research assistants screened and recruited patients and visitors for the questionnaire survey and helped to complete questionnaires when this was needed. Core team members and research assistants undertook the patient observations. All staff involved in the research received classroom and field training, as set out below:

- recruiting for patient observations – 4 hours in the classroom and 4 hours in the field
- conducting patient observations – 4 hours in the classroom and 6 hours in the field
- recruiting for and conducting patient and visitor questionnaires – 4 hours in the classroom and 2 hours in the field.

Changes from original protocol

Two key changes were agreed with the Study Steering Committee and implemented. The first change was to the target sample sizes specified for questionnaire surveys with nursing staff, patients and visitors. This was made after baseline data collection when the feasible recruitment rates became clear. *Tables 6* and *7* illustrate original and revised target recruitment rates.

The second change was to the economic evaluation. As the study progressed and the membership of the health economics team developed, it became clear that the health economic dimension to the evaluation would be a more helpful foundation to the definitive evaluation if it focused on how the costs and benefits of CLECC could be measured in a future definitive evaluation.

TABLE 6 Target recruitment rates (original protocol)

| Participants | Qualitative | Observations at (*n*) | | Questionnaire survey at (*n*) | |
		Baseline	Follow-up	Baseline	Follow-up
Nursing staff	30	–	–	252	252
Patients	12	120	120	252	252
Visitors	12	–	–	96	96

TABLE 7 Target recruitment rates (revised)

Participants	Qualitative	Observations at (*n*)		Questionnaire survey at (*n*)	
		Baseline	**Follow-up**	**Baseline**	**Follow-up**
Nursing staff	30	–	–	84	84
Patients	12	120	120	96	96
Visitors	12	–	–	30	30

The original study objective related to the economic evaluation was:

> *To estimate the costs of the intervention and quality of life.*

This was changed to:

> *To inform the measurement of costs and benefits of CLECC in a definitive evaluation.*

Chapter summary

This chapter has included information on the study design, the feasibility parameters being tested and the outcome measures being assessed. PPI, ethics considerations and changes from the original protocol have also been explained.

Chapter 5 Data sources

This chapter presents information on sampling, recruitment and data collection. Data collection took place between March 2015 and May 2016. *Figure 3* presents an overview of which data were gathered when. Each of the methods reflected is described in detail below.

Ward sampling and recruitment

Six ward teams were included in the study. The sample size was determined by funding availability and the plan to run the study in at least two hospital organisations, and at least two ward specialties. Our previous research had highlighted the importance of organisational and ward context in determining nurses' relational capacity, and this informed a sampling strategy aimed at diversity.[15] Senior nurses at two English NHS trusts were invited, on behalf of their organisations, to take part in the study and agreed. They were invited to nominate three acute inpatient units each to take part in the study. They were guided to select three adult inpatient wards in each trust with the highest proportion of patients aged ≥ 65 years. In both trusts, wards with the highest proportion of older patients were MOP, so, to aid diversity of the sample, we also specified that one of the wards needed to be a surgical ward with a high proportion of older patients. Senior nurses were also encouraged to include those wards perceived as less effective in their general performance in addition to those perceived as high-performing wards, because there is a need to know if interventions of this kind can also work in contexts in which staff may not recognise the potential benefits of the intervention and/or do not prioritise improving compassionate care.

Critical care units were not eligible for inclusion, as previous research indicates that nurses' experiences of providing relational care in critical care environments are very different from their experiences of providing care in more general wards.[15] Wards were also excluded if the departure of the ward leader was anticipated
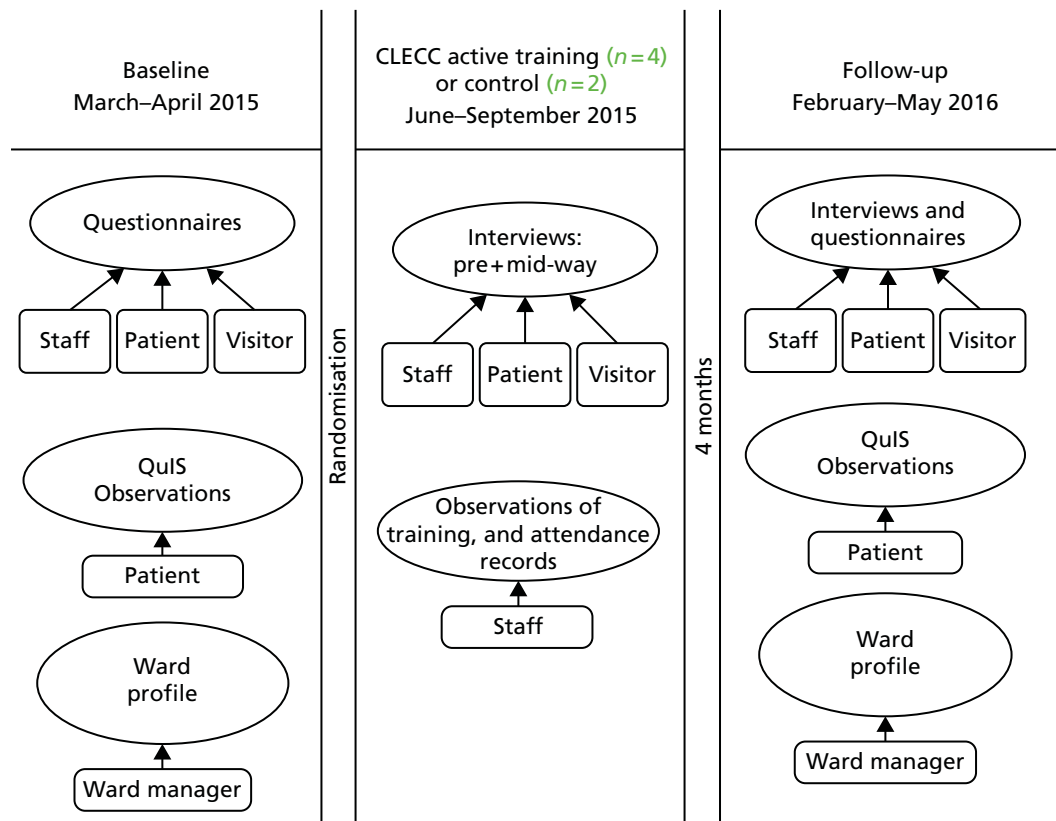
FIGURE 3 Data collection overview.

in the subsequent 6 months of attempted recruitment, as stable ward leadership was theorised to be an important influencing factor on implementation.

Following the circulation of written information about the study, the chief investigator met with the ward leader for each of the nominated wards. They were given a verbal explanation about the study and a chance to ask questions and discuss the implications of taking part. At the close of the discussion, they were invited to put their ward forward for the study or not, or, if they preferred, to take more time to make up their minds.

Once the wards were recruited, all nursing staff, including registered nurses (RNs) and health-care assistants (HCAs) employed to work in the participating ward teams, were eligible to take part in the study and recruitment processes were designed to maximise their inclusion. Meetings with ward leaders were designed to inform staff about the study. Researchers also offered to visit the wards to talk to the teams about the study and this took place on three wards. Written information sheets about the study were also given to ward leaders to distribute to staff. Posters about the study were displayed on the participating wards in public and staff areas.

Addressing the process evaluation and the pilot CRT in turn, the following sections outline sampling, recruitment and data collection.

Process evaluation

The process evaluation aimed to identify and explain the extent to which the planned intervention was implemented into existing nursing practices on the four intervention wards. Data were collected using a variety of methods in order to gather insights into different aspects of implementation from different viewpoints. The data collected were:

1. quantitative ward profile data generated by each ward leader at the outset of the intervention
2. qualitative interviews with nursing staff and ward leaders from the participating wards, the PDNs leading CLECC implementation, the matrons overseeing the wards and hospital senior nurses
3. field notes made by PDNs delivering the intervention
4. observations of a sample of CLECC classroom training days and CLECC action-learning sets for ward leaders
5. quantitative records of training delivered.

Data were collected between May 2015 and May 2016 to capture the period just before the implementation period, with follow-up for up to 12 months. Three main phases were used to guide the scheduling of data collection in the process evaluation: period 1 (baseline/early CLECC active training; May–August 2015), period 2 (mid-to-late CLECC implementation period; July–October 2015) and period 3 (after implementation period; October 2015–May 2016).

Process evaluation sampling and recruitment
All members of the participating nursing teams, including RNs and HCAs, were eligible to take part in the qualitative interviews and were invited to take part through posters, presentations and e-mails. We purposively sampled from those who volunteered to capture variations in staff grade and ward. Written consent was sought for the staff interviews and staff were given information about not being obliged to take part and their right to withdraw their consent at any time. We offered a payment of £15 shopping vouchers to individual staff who completed an interview. Individuals recruited early in the study were invited to a second and third interview so that variations over time could be tracked. When such individuals could not be contacted or declined further interviews, new individuals were recruited at the same grade and from the same ward to ensure that variation by ward and grade was maintained. All four ward leaders and two PDNs leading the CLECC implementation were also invited to three interviews each. At the final interview

round, all three matrons overseeing the participating wards and two further people in more senior nursing roles in the trust were invited to be interviewed.

Process evaluation data collection

Ward profile

Contextual data were gathered on intervention and control wards through the completion of a ward profile questionnaire by the ward leader (or other senior nurse on the team). These ward-level data included physical layout, specialty, bed occupancy, staffing, sickness rates, agency usage, turnover and shift length. These data were gathered at baseline phase and updated at follow-up phase.

Qualitative interviews

One-to-one face-to-face qualitative interviews were undertaken in three phases as outlined in *Process evaluation*. All but two participants opted to use a hospital meeting room for the interview. Two participants chose to be interviewed away from the hospital site. The interview schedules were designed to capture individual views and experiences, and to reflect the key NPT concepts of coherence, cognitive participation, collective action and reflexive monitoring. Period 3 interviews with intervention ward staff also included questions on resource implications of CLECC for the purposes of the economic evaluation. Schedules reflected the implementation stage at the time of interview. *Appendix 4* shows examples of interview schedules. The interviews lasted, on average, 46 minutes (range 17–70 minutes). Interviews were audio-recorded and transcribed verbatim, and transcripts were checked for accuracy by the interviewer.

The two PDNs kept detailed field notes of their experiences of delivering the intervention.

Observations of training activities

A researcher observed a sample of CLECC classroom training days and ward leader action-learning sets. Data were collected using unstructured non-participant observation, using event sampling, that is, recording observations of a set event rather than at regular periods over time.[113] The use of observation allowed for verbal and non-verbal interactions to be recorded in the form of field notes. The researcher adopted the role of complete observer and did not participate in the learning activities.[114] These observations were intended to complement the interview findings and the quantitative records of training delivered.

Quantitative records of training delivered

We also explored the feasibility of gathering data on the amount of training delivered through a register of attendance at classroom training and action-learning sessions, and a quantitative record of ward cluster discussions, reflective group discussions and cluster discussions.

Pilot cluster randomised trial

There were two main data collection phases for the pilot CRT and data were gathered from patients, staff and visitors. To enable us to gather data on baseline characteristics and assess against the selected outcomes, we used the following methods:

- observations of staff–patient interactions using QuIS[42]
- patient questionnaire survey comprising PEECH,[100] Picker Patient Experience Questionnaire (short-form)[100,115] to measure general care quality, EQ-5D-5L[93] and participant demographic details[100–102]
- nursing staff questionnaire survey comprising JSE, MBI, CC, FECC, selected items from the IHOS battery and participant demographic details[26,104,107–109]
- visitor questionnaire survey comprising CEC and participant demographic details[26]
- matron questionnaire survey comprising MAQC[26]
- ward leader ward profile
- qualitative interviews with nursing staff.

Copies of the questionnaires used are available on request.

We also piloted qualitative interviews about relational care on the wards with a small number of patients (*n* = 12) and visitors (*n* = 12) to inform a future process evaluation. An overview of the method and findings from the patient and visitor qualitative interviews is reported in *Appendix 5*.

Following baseline data collection in March and April 2015, the ward teams were randomised to intervention or control conditions. The CLECC intervention was implemented on four of the six wards from June 2015, starting on each ward with a 4-month implementation period. Follow-up data were collected on all six wards during February and March 2016.

Pilot cluster randomised trial recruitment

The procedures outlined here mirrored our envisaged procedures for a definitive trial. The opportunity was taken during the feasibility study to evaluate these procedures and further develop them when needed. Pre-screen and screening logs were developed to enable assessments of:

- the timeline of the introduction, approach, discussion and consent process
- the number of people assessed for eligibility
- the number of people approached to join the study
- the number of people recruited into the study
- the number of declined offers and the reasons for these decisions
- participation rates of older patients and patients with cognitive impairment
- the achievement of targets.

Patient recruitment to observations

All adult patients on participating wards were assessed for eligibility to be included in observations of care. Patients were excluded if they were unable to communicate their choices about taking part in the research and a consultee (as defined by the *Mental Capacity Act 2005 Code of Practice*)[110] could not be consulted. Patients who indicated either verbally or non-verbally that they did not wish to take part were excluded, as were patients who were unconscious or those for whom there were clinical concerns that may preclude them from being approached. Patients excluded for clinical reasons included people who were critically ill, at the end of life, or isolated because of a high risk of infection.

The patient sample for observations was determined by randomisation. Up to 24 hours in advance of each scheduled observation, all eligible patients on the ward at the time an observation was scheduled were identified. Patients were placed in a list in alphabetical order of names and each patient was then allocated a number in sequence. A random number generator was then used to select an index patient from the pool of eligible patients.

Each index patient was then approached and informed (verbally with accompanying written information and with aids when needed) about the planned observations. Recruiting researchers were trained to be person-centred and patient, allowing sufficient time for successful communication, and to make environmental modifications to optimise communication. If the index patient indicated verbally or non-verbally that they were happy for the observation to proceed, other eligible patients in the researcher's field of view were approached, informed about the planned observations, and if they indicated that they were happy for the observation to proceed their care was included in the observations. We did not ask for written consent from patients, but consent was instead recorded by the researcher. If the index patient declined to take part, another index patient was randomly selected, and approached and invited as before. The observations proceeded with data being collected on interactions with all patients who had agreed that the observations could proceed. When an assessment was made that a potential participant did not have capacity to make a decision about taking part in the research, advice was sought from a consultee, in line with the requirements of the *Mental Capacity Act 2005 Code of Practice*.[110] *Figure 4* outlines these processes.
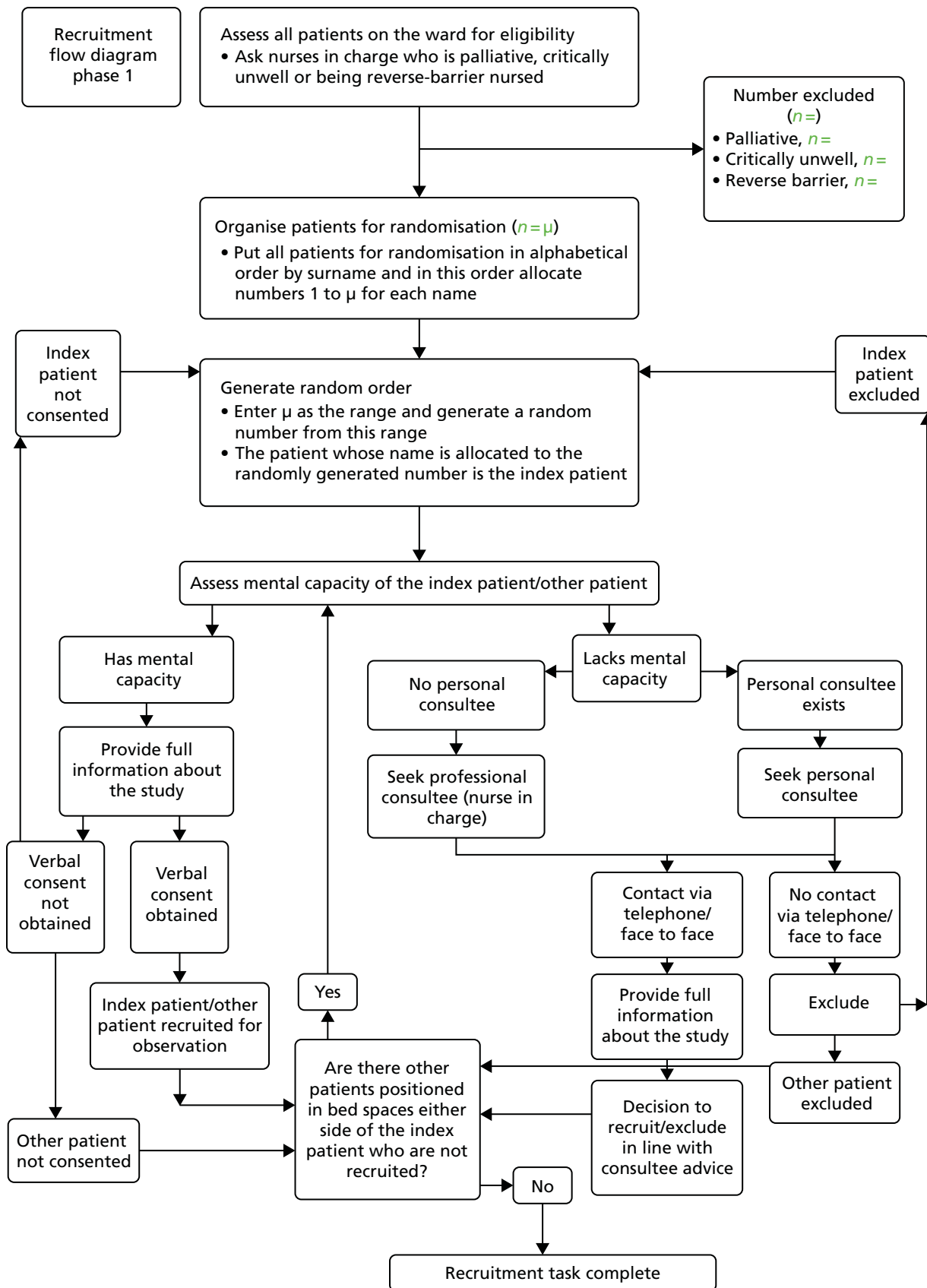
FIGURE 4 Recruitment process for observations.

Patient sample and recruitment to questionnaire survey

An adapted census approach was used with the patient questionnaire survey. Researchers aimed that all eligible patients on the ward on the scheduled day for questionnaire data collection would be invited to complete a questionnaire. Questionnaire data collection days were planned in advance for each ward to ensure researcher availability, but timing of planned data collection during that day meant that patients would be least likely to be involved in other activities, such as washes, meal times or doctors' rounds. Before distributing questionnaires, researchers ascertained with the nurse-in-charge which patients were able to be directly approached by a researcher to complete the questionnaire. Patients were excluded if they were critically ill, in receipt of palliative care or at high risk of infection. All eligible patients were then approached and assessed further for eligibility, particularly mental capacity. If the researcher was able to confirm eligibility, patients were informed about the research and invited to take part.

Staff sample and recruitment to questionnaire survey

We planned a census approach to the nursing staff survey in that all RNs and HCAs employed to work in the participating ward teams would be invited to complete a nursing questionnaire. Ward leaders were asked to provide a list of RNs and HCAs employed to work on their ward. Nursing questionnaires were then placed into individually named envelopes and given to ward leaders to distribute to staff.

Visitor sample and recruitment to questionnaire survey

An adapted census approach was used with the visitor questionnaire survey. Researchers aimed that all eligible visitors on the ward on the scheduled day for questionnaire data collection would be invited to complete a questionnaire. Questionnaire data collection periods were planned in advance to ensure researcher availability but also to coincide with visiting time on the individual wards. No exclusion criteria were set for visitors, and so any adult visitors on the ward during a data collection period were approached, informed about the research and invited to take part.

Staff recruitment to qualitative interviews

In addition to the qualitative interviews conducted as part of the process evaluation (which focused on the CLECC intervention implementation), intervention and control ward nursing staff members were also interviewed about ward characteristics, specifically relational care, teamwork and leadership at baseline and at follow-up. The purpose was to provide summary baseline data of group characteristics to inform the pilot CRT. Intervention ward staff recruited to interviews as part of the process evaluation were also questioned about these characteristics during their interviews (see *Pilot cluster randomised trial recruitment* and *Qualitative interviews*). A group of control ward staff was also sampled and recruited in the same way, with their interviews focusing solely on relational care, leadership and teamwork. All RNs and HCAs on control wards were eligible to take part in the qualitative interviews and were invited to take part through posters, presentations and e-mails. We purposively sampled from those who volunteered in order to capture variations in staff grade and ward. Written consent was sought for the staff interviews and staff were given information about not being obliged to take part and their right to withdraw their consent at any time. We offered a payment of a £15 shopping voucher to individual staff who completed an interview. Individuals recruited at baseline were invited to a second interview at follow-up. When such individuals could not be contacted or declined further interviews, new individuals were recruited at the same grade and from the same ward to ensure that variation by ward and grade was maintained.

Pilot cluster randomised trial data collection

Observations

All interactions with eligible patients over a 2-hour observation session were directly observed and coded by a trained researcher. Data gathered included the quality, length and frequency of all interactions between participating patients and staff during each observation session. Data collection was guided by a protocol for use of QuIS in acute settings developed in earlier feasibility work.[99] Contextual data were also gathered on the session (number of patients on the ward, staffing levels and skill mix), on the patients (age, gender, evidence of cognitive impairment, agitation at outset of interaction) and on individual

interactions (including number of staff, staff type and content of interaction). Patients were assessed as having cognitive impairment in a number of ways, although the research team did not have access to patient records at baseline and so could not look for evidence in the records. Clinical staff were asked about cognitive status before patients were approached. In some cases there was an indicator by the bedside. For instance, in one hospital patients with a known dementia diagnosis would have a magnetic flower mounted on the board by their bed to provide clinical staff with a subtle sign of their cognitive status. If such an indicator was present, this was recorded as evidence of cognitive impairment. In other cases, the researcher would detect signs of impairment as they talked to the patient or during the period of observation.

The platform developed and used for data collection was the Quality of Interactions Tool (QI Tool), a tablet-based interface developed during the feasibility study that enables users to enter data in real-time for subsequent wireless upload to an encrypted central database (see *Appendix 6*).

At each assessment period, researchers observed during time periods (10 × 2-hour observation sessions per ward per 3-week assessment period) randomly sampled over a 3-week period from Monday to Friday, 08.00–22.00. Observation sessions were balanced between wards and time of day. Follow-up data collection was conducted by researchers blinded to ward allocation.

Patient questionnaire survey

Questionnaire responses were written on a hard copy of the questionnaire. If patients agreed to complete a questionnaire, the researcher offered help with completing it and, if the patient was willing, the questionnaire could be completed straight away, taking as much time as was needed, or at a later point in time depending on patient preference. The length of time questionnaire completion took was recorded on each survey. Patients were offered help with interview completion to increase response rates, and whether or not this offer was taken up was also recorded.

Staff questionnaire survey

Questionnaire responses were written on a hard copy of the questionnaire. A post-box was placed on each ward for staff to return their completed questionnaires. Research team visits to the ward and e-mail feedback to ward leaders on completion rates were designed to encourage questionnaire completion. In addition, a prize of shopping vouchers was offered to the team in each hospital with the highest completion rate at each assessment period.

Visitor questionnaire survey

Questionnaire responses were written on a hard copy of the questionnaire. Completed questionnaires were gathered by hand by the researcher.

Completed questionnaires were collated and then individual responses were entered onto the SPSS version 13.0 (IBM Corporation, Armonk, NY, USA) database followed by 100% verification of data entered.

Matron questionnaire survey and ward leader ward profile

Matrons overseeing participating wards were e-mailed a survey, invited to complete it and to e-mail or post it back to the research team. Ward leaders were e-mailed a copy of the ward profile for completion and invited to e-mail or post it back when completed. Ward leaders were also offered help with completing the profile from the research team.

Nursing staff qualitative interviews

One-to-one face-to-face qualitative interviews were undertaken with nursing staff from intervention and control wards, as outlined in *Staff recruitment to qualitative interviews*, with the purpose of gathering qualitative data on baseline group characteristics. All participants opted to use a hospital meeting room for the interview. The interview schedules were designed to capture individual views and experiences as a member of their ward team, and focused on relational care, leadership and teamwork. The interviews

lasted, on average, 46 minutes (range 17–70 minutes). Interviews were audio-recorded, transcribed verbatim, and transcripts were checked for accuracy by the interviewer.

Economic evaluation

The economic evaluation drew on two main data sources outlined above. These were the EQ-5D-5L data from patient questionnaires and the qualitative interview data from staff that focused on the resource implications of CLECC. In addition, the cost of providing CLECC as an intervention was explored.

Chapter summary

This chapter has provided information about the key data sources used to address the study objectives, in particular sampling, recruitment and data collection in relation to the process evaluation, the pilot CRT and the economic evaluation. *Chapter 6* provides detail on data analyses.

Chapter 6 Data analysis

This chapter outlines the methods used for analyses of the process evaluation, pilot CRT and economic evaluation. To reiterate, the study objectives were:

1. to determine the feasibility of implementing the CLECC intervention and sustaining the resulting work practices
2. to inform the design of a definitive evaluation of the effectiveness of CLECC
3. to inform the measurement of costs and benefits of CLECC in a definitive evaluation.

Process evaluation

Data were analysed using systematic reading, familiarisation and open coding, undertaken independently by research team members and then in collaborative data analysis workshops. Coding discussions led to the development of two coding frames: the first to enable the exploration of themes related to relational care, ward leadership and teamwork across all six wards (for the CRT element), and the second, from intervention wards only, to focus on CLECC implementation and mechanisms of impact. All qualitative interview data from nursing staff were coded against these frames, with the use of constant comparative methods enabling the generation of new categories and the comparison of data in relation to these categories.

The relational care coding frame (*Figure 5*) was designed to support comparisons across the wards and, so, analysis focused on summarising and describing what people had to say about relational care, leadership and teamwork.

A deeper level of analysis was undertaken in relation to the CLECC implementation frame and this part of the analysis was the focus of the team's work. There was a conscious decision to avoid prematurely 'fitting' the data into NPT domains, and the use of the coding frame shown in *Figure 6* enabled preliminary descriptive coding in relation to implementation and mechanisms of impact. The constant comparison method was used to examine the codes generated against the framework, which was further developed as analysis progressed. Narrative data summaries and matrix/charting techniques were then used to facilitate comparison with the NPT framework and test and refine emerging theories of implementation processes.
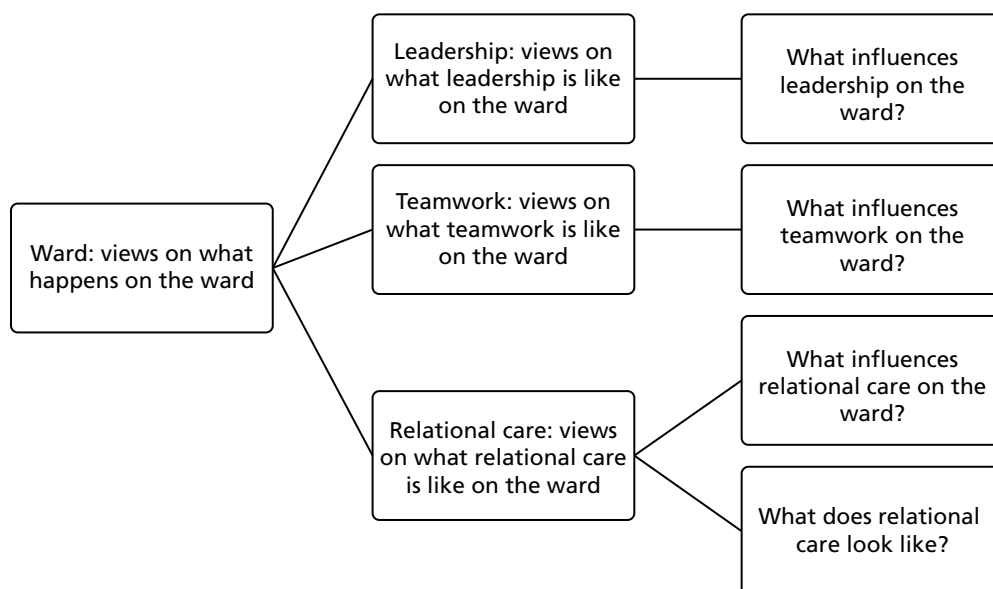
FIGURE 5 Relational care coding frame.

T2 and T3 data
- Ward managers
- Nurses
- HCAs

How is delivery achieved and what is actually delivered? Quantity and quality. Variations across, for example, trusts, wards, staff groups

Implementation

Views on CLECC

Was CLECC delivered as intended?

How much of CLECC was implemented?

What adaptations were made to CLECC?

What was the impact of adaptations?

How was implementation achieved?

What shaped implementation?a

Successful programmes
- Recruit with care
- Create expectations
- Demand effort
- Offer sounding boards and encouragement
- Build and test resilience
- Accommodate setbacks
- Explain themselves
- Share execution and control
- Offer onward external continuation

CONTEXT: anything external to CLECC that impedes/strengthens implementation

Mechanisms are agents of change
- Mechanisms describe how resources embedded in CLECC influence reasoning and, ultimately, behaviour of programme subjects
- CLECC will not work in and of itself; it is staff choices that make for success or failure

What do people say about what drove their behaviour/influenced their engagement with CLECC?

Mechanism of impact

How does the delivered intervention produce change?

Did CLECC work?

How did CLECC work?

What was CLECC's impact?

What shaped whether or not CLECC workeda

CONTEXT: anything external to CLECC that impedes/strengthens effects

FIGURE 6 Coding frame for analysis of CLECC implementation. Based on MRC guidance[89] and Pawson.[116] a, Anything external to CLECC that impedes/strengthens implementation.

Findings from the analyses of the qualitative interview data were triangulated with qualitative data from field notes kept by the PDNs leading the delivery of CLECC and researcher field notes from observations of CLECC training activities.

Most, but not all, of the team were researchers with a nursing background and had not met the research participants prior to the study. All were involved in interviewing and/or observing the staff in the study, and in data analysis. In early stages, more than one researcher coded the same data set, with comparisons between coding decisions creating opportunities for team discussion and the development of shared approaches. The deliberate use of reflective techniques during collaborative analysis workshops enabled individual views and assumptions to be surfaced and explored with a view to enhancing the quality of analysis. NVivo11 (QSR International, Warrington, UK) was used to support preliminary coding and consistency across the team.

Pilot cluster randomised trial

Data were analysed using SPSS 22[117] and the significant threshold (alpha) was set at 0.05 (two tailed). Stata 14 was used for multilevel logistic model. Exploratory data analyses were performed to check the data and identify inconsistencies. The primary purposes of tests of effectiveness were to pilot procedures for analysis and to inform effect-size calculations for a future study.

Baseline characteristics
Frequencies and percentages or mean and standard deviations (SDs) were calculated for baseline and follow-up patient and ward study participant flow, characteristics and contextual features.

Quality of Interactions Schedule data analysis
All trial analyses were carried out on an intention-to-treat basis with wards included on the basis of their planned CLECC intervention status, irrespective of the extent to which they actually adopted CLECC practices. Three approaches to the analysis of QuIS data as an outcome were used. The first analysed the proportion of interactions rated using each of the five QuIS ratings. The second calculated the proportion of all negative interactions (i.e. interactions rated as either negative protective or negative restrictive) per patient. The third used the data for the subset of patients observed for a full 2 hours and calculated the proportion of patients in this subset that experienced at least one negative interaction during the 2-hour observation session. Descriptive statistics were used to display the findings from all three approaches for baseline and follow-up assessment periods, and for intervention versus control wards. Analyses by individual ward were also performed. For the first and third approaches, differences between groups were then tested using a chi-squared test.

A three-level mixed-effects logistic regression model was fitted to investigate the effect of the CLECC intervention on the likelihood of an interaction being rated as negative (protective and restrictive combined). The individual interactions recorded between patients and staff were considered the lowest level of the model. Patient and observation session were included in the model as random effects making up the higher two levels of the model. Predictive factors were included as fixed effects and presented as odds ratios (ORs) with 95% confidence intervals (CIs), both before and after adjustment for the other predictors. Models were adjusted for baseline and ward consecutively.

In addition to tests of effectiveness, we also used baseline QuIS data to quantify and characterise staff–patient interactions and to identify the factors associated with negative interaction ratings. Details of the analysis are reported in our published paper, shown in *Appendix 7*.[118]

Patient Evaluation of Emotional Care during Hospitalisation data analysis
The PEECH tool assesses the degree of emotional care needed by patients in four subscales: security, knowing, personal value and connection. PEECH items are scored from 0 to 3, with 3 representing the

best possible score (0 = none of the time, 1 = some of the time, 2 = most of the time, 3 = all of the time). Subscale scores were calculated from the mean of items in the subscale and total PEECH score from the sum of all items when at least 75% of the items in each subscale were available. These were computed by individual ward, experimental group and assessment period. Differences in scores between groups at follow-up were tested using the Mann–Whitney U-test.

Using a different approach, subscale scores were dichotomised into either low scores (patients with an average score of ≤ 2) or high scores (average score of > 2). We calculated the frequencies of patients with low scores for each subscale by ward, experimental group and assessment period. Differences between groups were tested using a chi-squared test.

Further multivariate analyses using logistic regression were performed for findings in which a significant difference had been found. Unadjusted and adjusted models were fitted to predict the outcome with the CLECC intervention as the primary predictor. Models were adjusted for ward, baseline and patient characteristics (age, gender, ethnicity and education level) consecutively. ORs and their 95% CIs are presented.

Jefferson Scale of Empathy data analysis
Mean score, SD and range were calculated for total empathy score when at least 75% of the items were available. They are presented by individual ward, experimental group and assessment period. Differences between group means at follow-up were tested using the Mann–Whitney U-test.

Intracluster correlation
Intracluster correlation coefficients (ICC) for the three core outcome measures (QuIS, PEECH and JSE) were calculated to account for clustering by ward and observation period. ICCs for QuIS were calculated based on the proportion of patients with one or more negative QuIS interactions in a 2-hour period. ICCs were calculated using the *loneway* command in Stata 14.

Economic evaluation

The cost of the CLECC intervention comprised one-off training costs and ongoing implementation costs. Training costs were agreed with senior managers from participating hospitals, using standardised models for calculating the cost of staff time when relevant.

The CLECC implementation costs focused on the ward-based cluster discussions and the extent to which these required additional nursing staff time, with potential implications for nurse staffing ratios. Twenty-one transcripts from T3 qualitative interviews with intervention ward nursing staff were reviewed and annotated by James Raftery with a focus on data on resource implications of CLECC. These analyses were discussed with interviewers and with Jackie Bridges to validate emerging hypotheses as regards the implementation costs of CLECC.

To aid calculation of the benefits of CLECC in a future trial, the utility of the most used quality-adjusted life-year (QALY) outcome measures was explored, namely the EQ-5D-5L. The EQ-5D-5L was administered to patients in the relevant wards at baseline and at follow-up. Completion rates were estimated. EQ-5D-5L baseline and follow-up scores were translated to quality-of-life scores using the national tariff and compared by ward at each time point.[93]

Chapter summary

These chapter has outlined the various methods used for analyses of the process evaluation, pilot CRT and economic evaluation.

Chapter 7 Participant flow and baseline data

Study results are presented across four chapters. This chapter focuses on participant flow (numbers assigned to experimental groups and analysed) and on baseline data. *Chapter 8* presents the process evaluation findings on the implementation of the CLECC intervention. *Chapter 9* presents findings on the feasibility of evaluating the effectiveness of CLECC using the planned CRT design, including the feasibility of older people and people with cognitive impairment participating, outcome measure performance, and assessment of bias in the trial. It also addresses the feasibility of estimating CLECC costs. *Chapter 10* presents the results of outcome measurement in the pilot CRT.

The first part of this chapter describes the participation of individual ward teams (clusters) in the study and shares baseline characteristics of these teams, including ward specialty, staffing levels, skill mix, relational care, ward leadership and staff well-being. In the second part of the chapter, data are presented in relation to participant flow at the individual participant level in relation to each of the baseline data sets, and we include an overview of demographic characteristics for each data set. The chapter concludes with the findings from the baseline measures of quality of staff–patient interaction, patient-reported evaluations of emotional care and nurse-reported measures of empathy.

Recruitment and flow of ward teams (clusters)

Six ward nursing teams across two hospitals took part in the study, with three ward teams in each hospital. Ward leaders in all of the three nominated teams in Hospital A agreed to take part on behalf of their team. Two ward leaders in Hospital B agreed to take part, but a third ward leader nominated in Hospital B declined. Their team was about to embark on a quality improvement project and the leader was concerned that participating in the CLECC study as well would be too burdensome for the team. The matron concerned nominated an additional ward team caring for a high proportion of older patients, and this team's ward leader agreed to take part. Each of the ward leaders was consulted about the prospect of randomisation to intervention or control conditions and no concerns were raised about this feature of study participation, indicating that the randomisation strategy was acceptable to participating staff. Thus, six ward teams entered the study and all remained in the study until all data collection was complete.

Both participating hospitals were NHS trusts located in separate urban areas in the same geographical region in England. Hospital A is a university hospital and foundation NHS trust employing between 7000 and 8000 staff and providing services to > 1 million people plus a wide range of specialist services to > 3 million people. Hospital B employs > 6000 staff and provides acute services to > 0.5 million people living locally and provides a smaller range of specialist services than Hospital A. The urban areas served by both hospitals have pockets of high deprivation. Life expectancy at birth is marginally higher than the national average in Hospital A locality and lower in the Hospital B locality. Hospital A locality has a lower percentage of the population than the English average who are classified as being from the white ethnic group and Hospital B locality has a higher than average white population.

Individual ward characteristics

Table 8 shows an overview of the characteristics of individual wards in the study, taken from the ward profiles completed by each of the ward leaders at baseline. Four wards were MOP wards and two were surgical wards. All the wards had a similar number of beds, with Hospital A wards having fewer beds than Hospital B wards. There was more variation between the wards in mean length of stay, ranging from 6 days on one of the surgical wards to 19 days on the other surgical ward. All four MOP wards had a similar mean length of stay (13–14 days). Planned full-time equivalent (FTE) staffing levels (RNs and HCAs) and nursing skill mix (proportion of registered nurses in total of RNs and HCAs) varied between the wards.

TABLE 8 Individual ward characteristics (at baseline)

| Characteristics | Hospital | | | | | |
	A			B		
Ward	A	B	C	D	E	F
Allocation	CLECC	CLECC	Control	CLECC	CLECC	Control
Specialty	Surgery	MOP	MOP	MOP	MOP	Surgery
Beds (*n*)	29	29	28	30	32	31
Mean length of stay (days)	19	14	14	14	13	6
Planned staff FTE (*n*)	48.25	37.00	38.20	44.90	45.61	45.78
Staff in post (*n*)	47.00	36.00	34.00	49.00	51.00	50.00
Proportion of RNs in total (planned) nursing staff	65%	61%	60%	63%	63%	67%
Length of day shift: short (8 hours) or long (12 hours)	Mixed	Mixed	Mixed	Mixed	Mostly short	Mixed
Patients needing help with all ADL	17%	45%	25%	7%	0%	6%

ADL, activities of daily living; FTE, full-time equivalent.

The two surgical wards had higher planned staffing levels and skill mix than the MOP wards. In each ward in Hospital A, the numbers of actual staff in post were lower than planned, whereas for all wards in Hospital B, the staff numbers in post were higher than planned. All the wards used a mixture of short (approximately 8 hours) and long (approximately 12 hours) shifts for nursing work during the day.

There was substantial variation between the wards in the proportion of patients on the ward who ward leaders judged needed help with all of their activities of daily living (ADL), an indicator of patient dependency and thus need for help from nursing staff with these activities (0–45%). Wards in Hospital A had fewer beds and lower numbers of actual staff in post than Hospital B. Wards leaders in Hospital A also identified a higher proportion of patients needing help with all their ADL than Hospital B.

Ward leadership characteristics

All wards had a ward leader (senior sister) and deputy in post. *Table 9* shows the length of experience for each ward leader in that role on that ward, as a ward leader overall and in any role on that ward. Wards A, D, E and F all had ward leaders with at least 9 years' experience as a ward leader, whereas ward leaders for wards B and C were new or relatively new to the role. The ward leader for ward B took up post, their first at that level, at the beginning of the study. Only the ward F ward leader worked in a totally supernumerary role, with the others regularly or often being counted in the staffing numbers on a shift.

Quality of care

Table 10 shows the results of the MAQC completed through written matron's survey. A completed matron's assessment was received for each ward (*n* = 6). Higher scores indicate more favourable ratings. The wards have similar scores across both subscales with the exception of ward F, which had markedly lower scores for both subscales.

TABLE 9 Ward leadership characteristics by ward (at baseline)

Ward	Ward, allocation					
	A, CLECC	**B, CLECC**	**C, control**	**D, CLECC**	**E, CLECC**	**F, control**
How long has the ward sister been the senior sister/leader on this ward? (months)	108	0	10	16	69	4
How long is the ward sister's experience as a senior sister/ward leader? (months)	108	0	10	120	228	180
How long has the ward sister worked on this ward (in any role)? (months)	144	0	10	16	69	132
How often do ward sisters have patients allocated to them (i.e. they are counted in staffing numbers on a shift)?	Often (most shifts)	Regularly (at least once a week)	Regularly (at least once a week)	Often (most shifts)	Regularly (at least once a week)	Never

TABLE 10 Scores for MAQC (at baseline)

Subscale scores per ward	Ward, allocation						Mean (SD)
	A, CLECC	**B, CLECC**	**C, control**	**D, CLECC**	**E, CLECC**	**F, control**	
Meeting patients' needs score (range of scores = 6–30)	23	24	22	21	22	12	20.7 (4.4)
Looking to improve score (4–20)	15	17	14	15	16	10	14.5 (2.4)

Relational care

A sample of nursing staff ($n = 29$) from all wards was interviewed using qualitative methods to ascertain their understanding of relational care on the ward where they worked. There was no significant difference between wards and hospitals in the way that nurses generally described relational care. The term compassion was used with examples given of interactions that staff perceived as conveying compassion in their care delivery, for example offering a relative a cup of tea to make the patient more relaxed because patients see we are looking after their family as well (N002 staff nurse). There was a recognition among nurses of interactions that patients do not perceive as compassionate, for example getting on with a nursing intervention with a patient without first introducing yourself and what you intend doing (N001 staff nurse). Staff explained that this happened because of time pressure to complete all expected nursing activities. Nurses who admitted to not always introducing themselves or actively getting to know the patient perceived these episodes of care as lacking compassion. Time was reported as a significant factor in delivering compassionate care, with staff challenged by patients who appear to need more time than others to express their needs or concerns. However, staff felt that compassion can be transmitted even in few words, for example when someone is clearly distressed because they have been incontinent, taking care with tone of voice, saying 'it's OK' (N011 staff nurse).

Ward leaders reported incorporating discussion about compassionate care delivery with their ward team. They reported that a number of wards had patients who needed individual care, patients who tried to walk but, because of cognitive impairment, were unaware of their physical limitations to do this effectively and safely, and so their care required an increase in the usual staff complement. Agency staff were employed to provide this individual care and ward staff reported seeing them work without compassion, for example sitting and watching the patient rather than interacting with them. Ward leaders reported

actively engaging with all nursing staff who interact with patients on the ward, including agency staff, to promote staff expectations of compassionate care delivery.

Care on the ward was reported not to be focused solely on medical management of a patient but also on the incorporation of other interventions seen as important to an individual's well-being. Staff saw their role as offering choice to patients and this was reported to be led and supported by ward leaders, for example by offering the use of a dining table for patients who are able with or without assistance to get there (N034 ward leader). Relational care was talked about as an important dimension to the nursing role, and this was reflected in stories about connections developed between individual patients and individual staff members, and the sadness felt by staff when these patients died. This was especially the case when the staff member felt that they had been successful in providing good relational care, and that the relationship had been strong. All of the ward leaders considered role modelling and educating the nursing team to deliver good relational (compassionate) care to be an important part of their role.

Nurses on all wards shared an understanding of good relational care and felt that the general public perceived nurses as being less compassionate than they used to be. Examples were readily available of what constituted good relational care, as well as the barriers to and facilitators of the achievement of this standard. There was consensus that the degree to which care was compassionate varied, but that it was not ward specific, rather it was both time and staff dependent. Nursing staff were aware of a variety of initiatives that had been put in place to address the perceived lack of compassion in care delivery, citing the Friends and Family Test[119] most frequently. Staff generally were keen to find or be offered strategies to assist them in delivering good relational care more consistently. The following quotation reflects what ward leaders expected and what nurses expected to be able to deliver:

> I'm expecting them to look after the whole person, see what their needs are, help them with those needs and make them feel like the most important person they're dealing with at that point in time.
>
> *N031 ward leader*

Staff well-being

Table 11 displays findings from the baseline ward profile completed by the ward leader in relation to indicators of staff well-being. All the wards had new staff members joining over the previous 6 months, with wards D and E having the highest number of new starters. Sickness absence rates varied considerably,

TABLE 11 Staff turnover and absence by ward (at baseline)

Ward	Ward, allocation					
	A, CLECC	B, CLECC	C, control	D, CLECC	E, CLECC	F, control
How many of the team have joined the ward in the last 6 months? (number of people)	5	5	4	13	8	6
How many staff are currently on maternity/long-term sick leave? (number of people)	0	3	2	5	2	2
Average sickness absence rate	0.6%	1.4%	2.7%	10.7%	3.2%	5%
Rate of agency/bank staff booking over the last month	Daily	Daily	Daily	Daily	Daily	Several times a week
How often do staff not take their breaks because of work pressure?	Very rarely	Several times a week	Daily	A few times a month	Several times a week	Several times a week

from 0.6% on ward A to 10.7% on ward D. All the wards used bank or agency staff regularly, at least several times per week. On ward A, staff very rarely missed their breaks, whereas missing breaks happened more frequently on the other wards and on a daily basis on ward C.

Table 12 indicates the degree of burnout taken from baseline nursing survey responses (*n* = 91). Higher mean emotional exhaustion and depersonalisation scores, and lower personal accomplishment scores, indicate less favourable conditions and higher burnout. Wards A and E had a mean emotional exhaustion score lower than the overall mean, and wards A, D and F had a lower mean score for personal accomplishment. Wards A and D had a higher personal accomplishment mean score.

Table 13 shows the proportion of staff on each ward experiencing burnout calculated from nursing survey responses (*n* = 91). Emotional exhaustion varied from 27% of 22 staff on ward E to 50% of 18 staff on ward C. Depersonalisation ranged from 0% of 12 staff on ward A to 40% of five staff on ward B. Personal accomplishment ranged from 10% of 21 staff on ward F to 22% of 18 staff on ward C.

Other questionnaire results

Baseline and follow-up results for all questionnaire data are reported in *Appendix 8*.

TABLE 12 Scores for MBI (at baseline)

	Ward, allocation						
Subscale scores per ward	A (*n* = 12), CLECC	B (*n* = 5), CLECC	C (*n* = 18), control	D (*n* = 13), CLECC	E (*n* = 22), CLECC	F (*n* = 21), control	Total (*n* = 91)
Emotional exhaustion (0–54)							
Mean (SD)	20 (11)	24 (11)	26 (11)	24 (13)	19 (11)	23 (13)	22 (12)
Minimum to maximum	0 to 35	9 to 37	3 to 43	12 to 52	2 to 38	3 to 47	0 to 52
Depersonalisation (0–30)							
Mean (SD)	5 (3)	9 (7)	6 (5)	5 (7)	6 (5)	5 (5)	6 (5)
Minimum to maximum	0 to 11	0 to 17	0 to 15	0 to 20	0 to 14	0 to 17	0 to 20
Personal accomplishment (0–48)							
Mean (SD)	39 (8)	38 (6)	38 (7)	41 (8)	38 (8)	37 (6)	38 (7)
Minimum to maximum	25 to 48	29 to 44	25 to 48	24 to 48	11 to 48	23 to 47	11 to 48

TABLE 13 Staff experiencing burnout (at baseline)

	Ward, allocation						
Experiencing burnout	A (*n* = 12), CLECC	B (*n* = 5), CLECC	C (*n* = 18), control	D (*n* = 13), CLECC	E (*n* = 22), CLECC	F (*n* = 21), control	Total (*n* = 91)
Emotional exhaustion	4 (33%)	2 (40%)	9 (50%)	4 (31%)	6 (27%)	7 (33%)	32 (36%)
Depersonalisation	0	2 (40%)	3 (17%)	2 (15%)	4 (18%)	2 (10%)	13 (14%)
Personal accomplishment	2 (17%)	1 (20%)	4 (22%)	2 (15%)	3 (14%)	2 (10%)	14 (16%)
Values are frequencies (%).							

Individual participant flow

This section and the accompanying CONSORT diagram (*Figure 7*) focus primarily on individual participant flow in relation to the pilot CRT. Later, in *Nursing staff qualitative interviews*, but not reflected in the CONSORT diagram in *Figure 7*, data are also shared about participant flow in relation to nursing staff in the qualitative elements of the study.
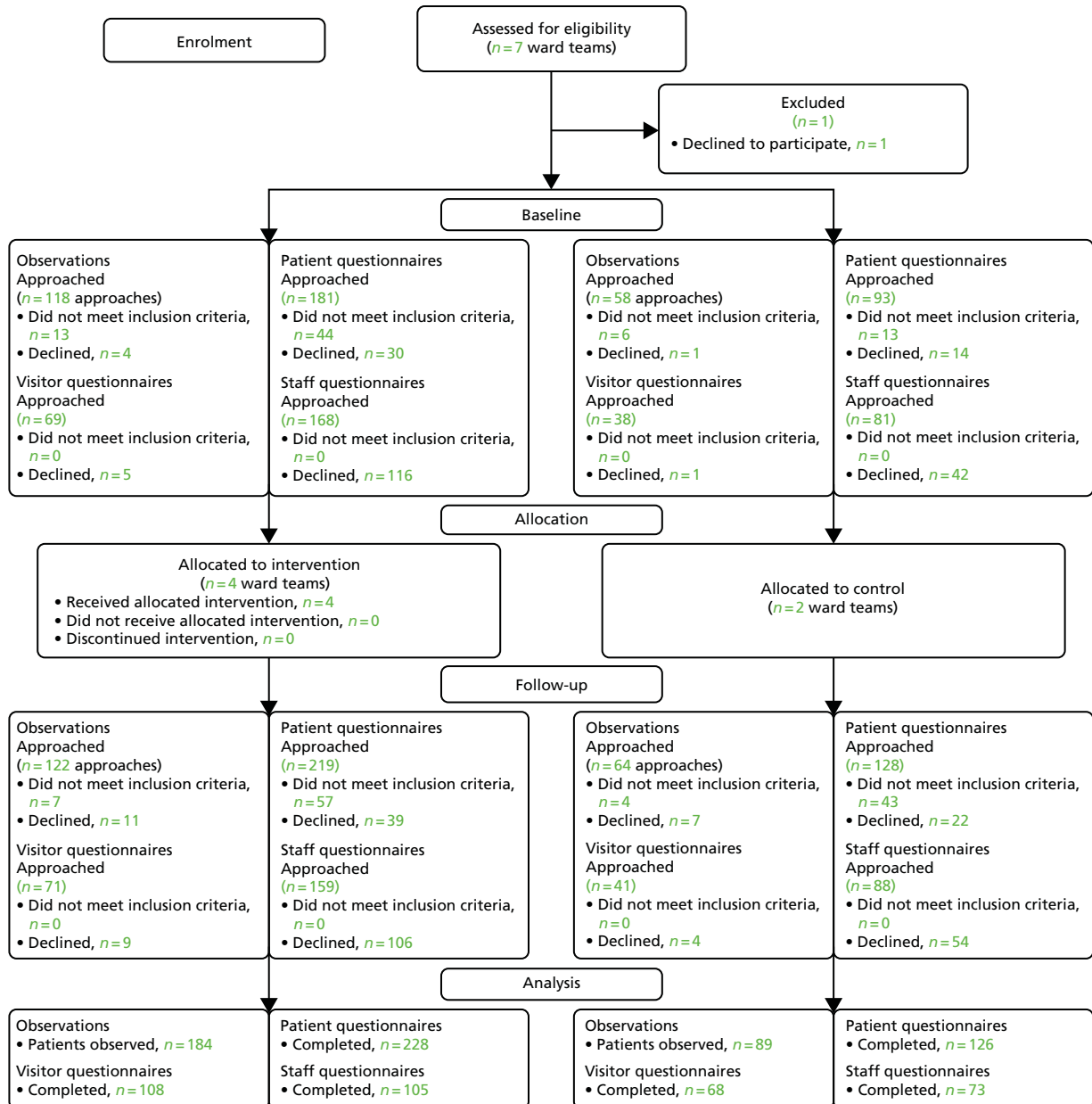
FIGURE 7 The CONSORT flow diagram for pilot cluster CRT. Reproduced from Gould *et al.* 2018.[120] © Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted. This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

All four wards randomised to the intervention went on to receive the intervention. *Figure 7* shows the flow of clusters and participants through the pilot CRT. Randomisation took place after baseline data collection, but the results are presented by allocation for baseline and follow-up data to enable comparisons between groups to be made.

A number of data from nursing staff, patients and visitors were gathered at baseline and repeated at follow-up to enable study outcomes to be assessed. Follow-up was at the cluster level rather than at the individual participant level. For instance, eligibility of patients and visitors was dependent on their being an inpatient during both assessment periods, which was an unlikely scenario given that they were 1 year apart. In addition, some turnover of staff employed on the ward teams was anticipated. Therefore, to maximise recruitment at each assessment point, recruitment for baseline and follow-up assessment periods was independent.

Observations

Recruitment was at the individual patient level for observation of staff–patient interactions. It was possible for individual patients to be involved in more than one observation session and the longer their stay on the ward, the more likely it was that they would be approached more than once to consent to their care being observed (provided they had consented on previous approaches). In total, 278 individual patients were approached, with a total number of 362 approaches. Observation recruitment data in *Figure 7* reflect the number of approaches rather than the number of individual patients.

The overall recruitment rate to baseline observations was 97%, that is, eligible patients agreed to take part in 152 out of 157 approaches. These rates were similar between intervention and control wards (96% vs. 98%). At baseline, patients consented to participate on 152 occasions. This corresponds to 123 individual patients because some were approached and consented more than once. Of the 152 occasions on which patients were consented they were observed 133 (88%) times. Some patients were consented and not observed because they were no longer available or eligible once observation was due to start.

Similar patterns were evident in the recruitment data for follow-up assessment, although recruitment rates were slightly lower. Overall recruitment to observations at follow-up was 90% (157 of 175 approaches). These rates were similar between intervention and control wards (90% vs. 88%). At follow-up, on 157 occasions, patients consented to participate. This corresponds to 114 individual patients as some patients consented more than once. Of the 157 occasions on which patients were consented, they were observed 140 (89%) times. Again, some patients were consented but not observed.

Across both assessment periods, in 93% (i.e. 309 out of 332) of approaches to eligible patients inviting them to participate in observations, patients agreed to take part, indicating a high acceptability. Recruitment rates at baseline were slightly higher than rates at follow-up. Overall recruitment rates were the same between intervention and control wards (93% vs. 93%).

Reasons recorded for patients declining participation included 'not feeling up to it' (17%, $n = 4$), 'too unwell' (4%, $n = 1$) and 'no reason' (8%, $n = 2$). No specific reason was recorded for 70% of approaches ($n = 16$). In 17% of approaches (i.e. 63 out of 362) the patient was approached and then assessed as not having the capacity to make the decision to take part in the research. In 67% (i.e. 42 out of 63) of these cases, researchers were able to contact a consultee for advice and in 100% of these cases the consultee advised that the patient should participate.

The mean age of patients observed was 82 years (84 years in the intervention group and 77 years in the control group). Patients in the control group were, on average, 7 years younger than patients in the intervention group, and this difference was statistically significant ($p < 0.001$). Most patients were female (77%) and 25% had evidence of cognitive impairment (31% at baseline and 19% at follow-up). There were no overall differences in gender and cognitive impairment between experimental groups.

Patient questionnaires

Potentially eligible patients ($n = 621$) were approached and invited to complete a written questionnaire, with researcher help if preferred, resulting in the completion of 354 questionnaires. Patients were approached having been screened as potentially eligible but, once approached, some patients (21%, 57 out of 274 approaches at baseline across six wards) were assessed as not meeting the inclusion criteria and were excluded from the study. A further 44 patients approached at baseline declined to take part. The overall recruitment rate to the baseline patient questionnaires was 80% (i.e. 173 out of 217 eligible patients agreed to take part). Recruitment rates were similar between intervention and control wards (78% vs. 83%). At baseline, questionnaires were completed for 97% (i.e. 168 out of 173) of the patients who consented.

Similar patterns were evident in the recruitment data for follow-up assessment. On approach, 100 patients in 347 approaches were assessed as not the meeting inclusion criteria. A total of 61 patients declined to take part. Overall recruitment to patient questionnaires at follow-up was 75% (i.e. 186 out of 247 eligible patients). Recruitment rates were similar between intervention and control wards (76% vs. 74%). At follow-up, a questionnaire was returned for 100% of the 186 patients who consented.

Across both assessment periods, 77% (i.e. 359 out of 464) of eligible patients agreed to take part in the questionnaire survey, indicating good acceptability. Although follow-up recruitment rates were slightly lower than at baseline, the number of patients approached was higher, so more patients were recruited overall to complete a questionnaire at follow-up than at baseline. Overall recruitment rates were similar between intervention and control wards (77% vs. 78%).

The most frequent reasons recorded for patients declining participation in the questionnaire survey were 'tired' (40%, $n = 12$) and 'questionnaire too difficult' (10%, $n = 3$). The most frequent reasons recorded for excluding patients were not having capacity (43%, $n = 48$) and being 'very cognitively impaired' (29%, $n = 32$).

All returned questionnaires were included in analyses. Most patients who completed questionnaires were female (70%), aged > 70 years (83%) and of white British ethnicity (97%). In total, 61% had other illnesses apart from the reason for hospital admission, 68% needed help with ADL while in hospital and 78% had been in hospital for > 3 days. Of all the patient questionnaires returned, 12% were completed by patients with cognitive impairment.

There were a small number of differences (not statistically significant) between intervention and control group patients who completed questionnaires at baseline (*Table 14*). At baseline, all of the intervention ward patients who completed a questionnaire were aged ≥ 61 years, whereas the control group, although including mostly patients aged ≥ 61 years, also included 13 people (22%) who were aged 31–60 years. At baseline, 26 (43%) of the control group were male compared with 25 (25%) of the intervention group. At baseline, 30 (53%) of the control group patients identified that they had other health conditions apart from the one that had brought them into hospital, compared with 62 (70%) of intervention ward patients. Twenty-eight (48%) of the baseline control group patients said that they needed help from others with ADL in hospital, compared with 74 (78%) of the intervention group. These differences about needing help from others were similar at follow-up. For the other responses identified here, there were no marked differences between groups at follow-up.

Visitor questionnaires

Visitors on the ward were approached and invited to complete a written questionnaire ($n = 219$), with researcher help if preferred. This resulted in completion of 176 questionnaires over both periods.

Of those approached across the six wards at baseline, six people declined to take part. The overall recruitment rate to baseline visitor questionnaires was therefore 94% (i.e. 101 out of 107 visitors agreed to take part). This rate was similar between intervention and control wards (93% vs. 97%). At baseline, questionnaires were completed for 88% (i.e. 89 out of 101) of the visitors who consented.

TABLE 14 Characteristics of patient questionnaire respondents

| | Time point | | | | | |
| | Baseline | | | Follow-up | | |
Variable	Control	CLECC	Total	Control	CLECC	Total
Age (years)						
18–30	0	0	0	3 (6%)	1 (1%)	4 (3%)
31–40	1 (2%)	0	1 (1%)	1 (2%)	1 (1%)	2 (1%)
41–50	5 (8%)	0	5 (3%)	1 (2%)	3 (3%)	4 (3%)
51–60	7 (12%)	0	7 (3%)	3 (6%)	5 (5%)	8 (5%)
61–70	3 (5%)	6 (6%)	9 (6%)	9 (18%)	6 (5%)	15 (9%)
> 70	44 (73%)	95 (94%)	139 (86%)	33 (66%)	94 (85%)	127(79%)
Gender						
Male	26 (43%)	25 (25%)	51 (32%)	16 (33%)	28 (25%)	44 (27%)
Female	35 (57%)	74 (75%)	109 (68%)	33 (67%)	84 (75%)	117 (73%)
Ethnic group						
Prefer not to say	1 (2%)	0	1 (1%)	0	0	0
White British	56 (95%)	95 (95%)	151 (95%)	53 (100%)	113 (98%)	166 (98%)
White Irish	2 (3%)	1 (1%)	3 (2%)	0	1 (1%)	1 (1%)
Other white	0	3 (35)	3 (2%)	0	1 (1%)	1 (1%)
Mixed ethnicity	0	1 (1%)	1 (1%)	0	0	0
Education level						
Primary school	5 (9%)	6 (6%)	11 (7%)	2 (4%)	8 (8%)	10 (7%)
Secondary school	31 (54%)	70 (71%)	101 (65%)	27 (57%)	71 (66%)	98 (64%
College	16 (28%)	17 (17%)	33 (21%)	12 (26%)	19 (18%)	31 (20%)
University	5 (9%)	5 (5%)	10 (7%)	6 (13%)	9 (85%)	15 (10%)
Other illness						
Yes	30 (53%)	62 (70%)	92 (63%)	26 (61%)	61 (58%)	87 (59%)
No	27 (47%)	27 (30%)	54 (37%)	17 (39%)	44 (42%)	61 (41%)
Need help from others						
Yes	28 (48%)	74 (78%)	102 (66%)	27 (52%)	88 (78%)	115 (70%)
No	31 (52%)	21 (22%)	52 (34%)	25 (48%)	25 (22%)	50 (30%)
More than 3-day stay?						
Yes	49 (82%)	81 (83%)	130 (82%)	38 (75%)	80 (73%)	118 (74%)
No	11 (18%)	17 (17%)	28 (17%)	13 (25%)	29 (27%)	42 (26%)

Similar patterns were evident in the recruitment data for follow-up assessment. A total of 13 visitors declined to take part. Overall recruitment to visitor questionnaires at follow-up was 88% (i.e. 99 of 112 eligible patients). These rates were similar between intervention and control wards (87% vs. 90%). At follow-up, a questionnaire was returned for 88% (i.e. 87 out of 99) of the visitors who consented.

Across both assessment periods, 91% of visitors approached (i.e. 200 out of 219) agreed to take part, indicating a high acceptability. Recruitment rates at baseline were slightly higher than rates at follow-up. Overall recruitment rates were similar between intervention and control wards (90% vs. 94%).

All returned questionnaires were included in analyses. *Table 15* shows the characteristics of visitor questionnaire respondents. The mean age for visitors who completed the questionnaires was 62 years and 63% were female. Most (98%) were of white British ethnicity. In total, 25% were either the spouse or the partner of the patient, and 42% were a son or daughter of the patients. A total of 58% identified themselves as a carer to the patient. At follow-up, 43% of intervention group visitors were visiting a

TABLE 15 Characteristics of visitor questionnaire respondents

| Variable | Time point | | | | | |
| | Baseline | | | Follow-up | | |
	Control	CLECC	Total	Control	CLECC	Total
Age of visitor (years)						
Mean (SD)	65 (11)	66 (12)	65 (12)	57 (18)	61 (14)	59 (16)
Minimum to maximum	32 to 89	30 to 91	30 to 91	17 to 93	22 to 89	17 to 93
Gender of visitor, n (%)						
Male	12 (34)	20 (39)	32 (37)	13 (39)	18 (35)	31 (36)
Female	23 (66)	32 (61)	55 (63)	20 (61)	34 (65)	54 (64)
Ethnic group of visitor, n (%)						
White British	33 (94)	52 (100)	85 (98)	32 (97)	54 (100)	86 (99)
White Irish	1 (3)	0	1 (1)	0	0	0
Other white background	1 (3)	0	1 (1)	0	0	0
Indian	0	0	0	1 (3)	0	1 (1)
Education level of visitor, n (%)						
Primary school	0	1 (2)	1 (1)	1 (3)	1 (2)	2 (2)
Secondary school	13 (38)	19 (38)	32 (38)	12 (38)	18 (35)	30 (36)
College	15 (44)	22 (44)	37 (44)	13 (41)	25 (49)	38 (46)
University	6 (18)	8 (16)	14 (17)	6 (19)	7 (14)	13 (16)
Relationship to the patient, n (%)						
Husband/wife/partner	12 (34)	11 (21)	23 (26)	9 (29)	11 (21)	20 (24)
Daughter/son	14 (40)	26 (50)	40 (46)	7 (23)	25 (47)	32 (38)
Father/mother	1 (3)	3 (6)	4 (5)	6 (19)	3 (6)	9 (11)
Daughter/son (in law)	3 (9)	2 (4)	5 (6)	0	3 (6)	3 (4)
Friend	3 (9)	2 (4)	5 (6)	1 (3)	3 (6)	4 (5)
Other	2 (6)	8 (15)	10 (12)	8 (26)	8 (15)	16 (19)
Are you a carer to the patient? n (%)						
Yes	16 (52)	32 (60)	48 (57)	16 (49)	33 (64)	49 (58)
No	14 (48)	21 (40)	35 (42)	17 (51)	19 (36)	36 (42)
Age of patient (years)						
Mean (SD)	79 (13)	86 (7)	83 (10)	70 (21)	83 (10)	78 (17)
Minimum to maximum	44 to 97	61 to 95	44 to 97	24 to 96	44 to 97	24 to 97
Patient diagnosed with Alzheimer's disease n (%)/other dementia?						
Yes	6 (18)	8 (15)	14 (16)	6 (18)	23 (43)	29 (34)
No	25 (73)	41 (77)	66 (76)	24 (73)	25 (47)	49 (57)
Do not know	3 (9)	4 (8)	7 (8)	3 (9)	5 (10)	8 (9)

patient with a dementia diagnosis, compared with 18% of control group visitors, and this was the only statistically significant difference between the two groups.

The most frequent reasons given by visitors declining participation were 'had already filled in a form that day' (16%, $n = 3$) and 'felt unable to answer accurately' (16%, $n = 3$).

Nursing questionnaires

All RNs and HCAs employed to work on the participating wards at the time of data collection ($n = 496$) were given a questionnaire to complete. This resulted in the completion of 178 questionnaires.

Of the 249 questionnaires distributed across the six wards at baseline, 158 were not returned. The overall recruitment rate to baseline nursing questionnaires was therefore 37% (i.e. 91 returned out of 249). Baseline return rates were lower on intervention wards than on control wards (31% vs. 48%). At follow-up the response rate of 35% (i.e. 87 questionnaires returned out of 247 distributed) was similar to baseline. Follow-up return rates were similar between intervention and control wards (33% vs. 39%).

Across both assessment periods 36% of questionnaires (i.e. 178 out of 496) distributed were returned completed. Overall return rates were lower on intervention wards than on control wards (32% vs. 43%).

All of the returned questionnaires were included in the analysis. As illustrated in *Table 16*, respondents represented a range of ages, ethnic groups, job roles/bands and years of experience.

TABLE 16 Characteristics of nursing questionnaire respondents

| Variable | Time point, *n* (%) | | | | | |
| | **Baseline** | | | **Follow-up** | | |
	Control	**CLECC**	**Total**	**Control**	**CLECC**	**Total**
Age (years)						
≤ 25	10 (26)	12 (24)	22 (25)	9 (27)	11 (21)	20 (23)
26–35	11 (29)	13 (26)	24 (27)	11 (32)	14 (27)	25 (29)
36–45	10 (26)	10 (20)	20 (23)	7 (21)	13 (25)	20 (23)
46–55	5 (13)	7 (14)	12 (14)	5 (15)	8 (15)	13 (15)
≥ 56	2 (5)	8 (16)	10 (11)	2 (6)	6 (12)	8 (9)
Gender						
Male	2 (5)	5 (10)	7 (8)	4 (12)	9 (17)	13 (15)
Female	36 (95)	46 (90)	82 (92)	30 (88)	43 (83)	73 (85)
Ethnic group						
Prefer not to say	4 (11)	5 (10)	9 (10)	4 (12)	4 (8)	8 (10)
White British	27 (71)	33 (65)	60 (67)	19 (58)	29 (58)	48 (58)
Irish	0	0	0	0	1 (2)	1 (1)
Any other white	4 (11)	6 (12)	10 (11)	7 (21)	9 (18)	16 (19)
White and black Caribbean	0	2 (4)	2 (2)	0	2 (4)	2 (2)
White and Asian	0	0	0	1 (3)	0	1 (1)
Any other mixed	1 (3)	0	1 (1)	0	2 (4)	2 (2)
Indian	0	1 (2)	1 (1)	1 (3)	1 (2)	2 (2)
Any other Asian	2 (5)	4 (8)	6 (7)	1 (3)	2 (4)	3 (4)

<div align="right">continued</div>

TABLE 16 Characteristics of nursing questionnaire respondents (*continued*)

	Time point, *n* (%)					
	Baseline			**Follow-up**		
Variable	**Control**	**CLECC**	**Total**	**Control**	**CLECC**	**Total**
Job title						
HCA	15 (39)	21 (42)	36 (41)	16 (47)	22 (45)	38 (46)
Staff nurse	16 (41)	24 (48)	40 (45)	13 (38)	21 (43)	34 (41)
Sister/charge nurse	5 (13)	5 (10)	10 (11)	4 (12)	4 (8)	8 (10)
Other	3 (7)	0	3 (3)	1 (3)	2 (4)	3 (4)
Current band						
2	17 (44)	20 (40)	37 (42)	17 (50)	22 (43)	39 (46)
4	0	1 (2)	1 (1)	0	2 (4)	2 (2)
5	16 (41)	25 (50)	41 (46)	13 (38)	23 (45)	36 (42)
6	5 (13)	4 (8)	9 (10)	4 (12)	2 (4)	6 (7)
7	1 (3)	0	1 (1)	0	2 (4)	2 (2)
Full time						
Yes	28 (74)	34 (68)	62 (71)	25 (76)	39 (77)	64 (76)
No	10 (26)	16 (32)	26 (29)	8 (24)	12 (23)	20 (24)
Years of career						
Mean (SD)	11 (10)	10 (10)	10 (10)	10 (9)	10 (8)	10 (8)
Median (LQ, UQ)	8 (3, 17)	7 (3, 18)	7 (3, 17)	9 (4, 15)	9 (2, 16)	9 (3, 16)
Minimum to maximum	1 to 40	1 to 35	1 to 40	1 to 30	0 to 30	0 to 30
Years on this ward						
Mean (SD)	4 (5)	4 (4)	4 (5)	3 (3)	3 (3)	3 (3)
Median (LQ, UQ)	2 (1, 4)	2 (1, 5)	2 (1, 5)	2 (1, 4)	2 (1, 5)	2 (1, 5)
Minimum to maximum	1 to 25	0 to 20	0 to 25	0 to 14	0 to 12	0 to 14

LQ, lower quartile; UQ, upper quartile.

Nursing staff qualitative interviews

All nursing staff on all participating wards were invited to participate in qualitative interviews, in addition to the two PDNs, two senior trust nurses and three matrons. Intervention ward staff and PDNs were invited to interview on three occasions (period 1, period 2 and period 3), and control ward staff were invited on two occasions (period 1 and period 3). At the final interview round (period 3), all three matrons overseeing the participating wards and two further people in senior trust nursing roles in the trusts were invited.

In total, 59 interviews were conducted, over three rounds, with 33 people. In total, 17 people were interviewed once, six were interviewed twice and 10 were interviewed three times. Eleven people who were interviewed once (at either period 1 or period 2) could not be contacted again or declined to participate in a further interview round, and one person who was interviewed twice did not take part in the third interview round. Six people who dropped out were HCAs and six were RNs.

Two senior trust nurses, two PDNs, 21 intervention ward staff and eight control ward staff were interviewed. All staff levels were represented at each interview round. All intervention wards were represented at each interview round and both control wards were represented at the first and third rounds.

All but one of the interviewees were female. In total, 11 interviewees were band 5 staff nurses, 10 were HCAs, six were ward leaders, two were deputy ward leaders, two were PDNs and two were senior trust nurses. A range of ages was represented, from the '29 years or under' to the '60–69 years' age band. RNs (including PDNs and senior trust nurses) had between 2 and 26 years' experience as a RN, with a mean length of experience equalling 11 years. HCAs had between 2 and 35 years' experience as a HCA, with a mean length of experience equalling 14 years.

Table 17 shows the characteristics of the nursing staff interviewed, not including the PDNs or the senior trust nurses. No significant differences between control and intervention ward staff were recorded.

TABLE 17 Characteristics of ward team nursing staff qualitative interviewees

Variable	Experimental group	
	Control	CLECC
Age (years), n (%)		
≤ 29	2 (25)	7 (33)
30–39	2 (25)	5 (24)
40–49	1 (13)	4 (19)
50–59	1 (13)	4 (19)
60–69	2 (25)	1 (5)
Gender, n (%)		
Male	0 (0)	1 (5)
Female	8 (100)	20 (95)
Job title, n (%)		
HCA	3 (38)	7 (33)
Staff nurse	3 (38)	8 (38)
Deputy ward leader	0 (0)	2 (10)
Ward leader	2 (25)	4 (19)
Current band, n (%)		
2	3 (38)	7 (33)
5	3 (38)	8 (38)
6	0 (0)	2 (10)
7	2 (25)	4 (19)
Years as RN		
Mean	9	12
Minimum to maximum	2 to 17	2 to 26
Years as HCA		
Mean	17	12
Minimum to maximum	11 to 25	2 to 35
Months on this ward		
Mean	82	52
Minimum to maximum	8 to 300	0.5 to 168

Baseline measures

This subsection reports the baseline scores for each ward on the study's core outcomes: quality of staff–patient interactions, patient evaluation of emotional care and nursing self-reported empathy. It also compares these outcomes by experimental group in the study, that is, intervention or control.

Quality of staff–patient interactions

The observed quality of staff–patient interactions using the QuIS is shown per ward in *Table 18*. All wards were observed for the same total time (20 hours each) balanced across days of the week (Monday to Friday) and times of day (08.00 to 22.00). The number of interactions across this 20-hour period varied between wards from 201 on ward F to 322 on ward B (see *Table 18*). The distribution of scores across the five available categories was broadly similar between wards, with most interactions (73% overall) scoring either positive social (13%) or positive care (60%). A total of 17% of interactions were rated neutral, 6% were rated as negative restrictive and 4% were rated as negative protective.

Table 18 also shows the number and proportion of all negative interactions, that is, the sum of negative protective interactions and negative restrictive interactions. The proportion of negative interactions ranged from 2% on ward F to 18% on ward C.

In total, 40% of patients observed for the full planned 2-hour period had one or more negative interactions, but the proportion on individual wards varied from 13% on ward F to 71% on ward A (see *Table 18*).

Table 19 shows the quality of interaction by experimental group. Distribution across the QuIS categories is similar for intervention and control wards. *Table 19* confirms that the proportion of negative interactions is the same (10%) for intervention and control wards. Of all patients observed, an average of 8% of interactions per patient were negative, and these proportions were similar across intervention (8%) and control (7%) wards. Of patients observed for the full 2-hour period, 45% of intervention ward patients had one or more negative interactions in a 2-hour period, compared with 30% of control ward patients.

In addition to tests of effectiveness, we also used baseline QuIS data to quantify and characterise staff–patient interactions and to identify the factors associated with negative interaction ratings. Findings are shown in our published paper displayed in *Appendix 7*.[118]

TABLE 18 Quality of staff–patient QuIS interactions by ward (at baseline)

Ward (number of interactions)	Ward, *n* (%)						Total (*n* = 1554), *n* (%)
	A (*n* = 257)	B (*n* = 322)	C (*n* = 210)	D (*n* = 313)	E (*n* = 251)	F (*n* = 201)	
Positive social	46 (18)	50 (15)	21 (10)	49 (16)	22 (9)	16 (8)	204 (13)
Positive care	127 (49)	196 (61)	116 (55)	205 (66)	144 (57)	139 (63)	927 (60)
Neutral	44 (17)	54 (17)	35 (17)	34 (11)	58 (23)	42 (21)	267 (17)
Negative protective	22 (9)	8 (3)	16 (8)	5 (2)	7 (3)	1 (1)	59 (4)
Negative restrictive	18 (7)	14 (4)	22 (10)	20 (6)	20 (8)	3 (2)	97 (6)
Negative protective + negative restrictive	40 (16)	22 (7)	38 (18)	25 (8)	27 (11)	4 (2)	156 (10)
Patients observed for the full 2 hours (number of patients)	17	20	23	20	21	16	117
Patients with one or more negative interactions	12 (71)	7 (35)	8 (35)	9 (45)	9 (43)	2 (13)	47 (40)

TABLE 19 Quality of staff–patient QuIS interactions by experimental group (at baseline)

QuIS rating	Trial group (N = 1554)	
	CLECC (n = 1143 interactions)	Control (n = 411 interactions)
Positive social n (%)	167 (15)	37 (9)
Positive care n (%)	672 (59)	255 (62)
Neutral n (%)	190 (17)	77 (19)
Negative protective n (%)	42 (4)	17 (4)
Negative restrictive n (%)	72 (6)	25 (6)
Negative protective + negative restrictive n (%)	114 (10)	42 (10)
Patients (n)	92	41
Percentage negative interactions per patient (minimum to maximum)	8% (0% to 53%)	7% (0% to 63%)
Patients observed for full 2 hours (n)	80	37
Patients with one or more negative interactions n (%)	36 (45)	11 (30)

Patient evaluation of emotional care

Patient evaluations of emotional care in each ward using the PEECH tool, administered through written patient surveys, are shown in *Tables 20* and *21*. Higher scores indicate more favourable ratings. *Table 20* displays mean and median PEECH scores for each subscale and total by ward. Scores tended to be broadly similar between wards for each subscale, with the connection subscale consistently scoring lower than the other three subscales on all wards. Total mean PEECH scores (sum of the mean of all items) ranged from 44.5 on ward C to 52.9 on ward E (out of possible total of 66). The distribution of scores was also similar between the two experimental groups (see *Table 21*).

TABLE 20 The PEECH scores by ward (at baseline)

PEECH ward (n)	Ward (N = 168), mean (SD); median (LQ, UQ)						Total (n = 168), mean (SD); median (LQ, UQ)
	A (n = 25)	B (n = 19)	C (n = 20)	D (n = 31)	E (n = 30)	F (n = 43)	
Security (0–3)	2.53 (0.45); 2.50 (2.25, 3.00)	2.29 (0.58); 2.33 (2.08, 2.75)	2.21 (0.59); 2.33 (1.96, 2.58)	2.30 (0.69); 2.40 (1.92, 3.00)	2.58 (0.42); 2.67 (2.50, 2.83)	2.43 (0.47); 2.50 (2.17, 2.83)	2.41 (0.54); 2.50 (2.17, 2.83)
Knowing (0–3)	2.42 (0.47); 2.33 (2.00, 3.00)	1.78 (1.02); 2.00 (1.00, 2.67)	1.96 (0.75); 2.00 (1.58, 2.42)	2.05 (0.97); 2.33 (1.08, 3.00)	2.35 (0.69); 2.50 (2.00, 3.00)	2.45 (0.67); 2.67 (2.00, 3.00)	2.22 (0.79); 2.33 (1.67, 3.00)
Personal value (0–3)	2.42 (0.53); 2.60 (2.15, 2.85)	2.08 (0.67); 2.20 (1.69, 2.70)	2.18 (0.65); 2.30 (1.75, 2.65)	2.25 (0.63); 2.40 (1.73, 2.70)	2.53 (0.40); 2.60 (2.35, 2.80)	2.43 (0.54); 2.60 (2.10, 2.80)	2.35 (0.57); 2.50 (2.00, 2.80)
Connection (0–3)	1.71 (0.61); 1.33 (1.33, 2.33)	1.43 (0.79); 1.33 (0.83, 2.00)	1.70 (0.85); 1.67 (1.33, 2.67)	1.63 (0.83); 1.33 (1.00, 2.08)	1.89 (0.71); 2.00 (1.33, 2.33)	1.56 (0.84); 1.67 (1.00, 2.00)	1.66 (0.78); 1.67 (1.00, 2.33)
Total PEECH (0–66)	51.5 (9.5); 52.0 (48.0, 60.0)	44.9 (13.9); 47.0 (38.0, 54.0)	44.5 (13.9); 43.5 (34.3, 52.8)	45.8 (13.9); 50.0 (32.5, 55.0)	52.9 (7.8); 54.0 (48.0, 58.5)	50.2 (10.7); 52 (45, 59)	48.9 (11.7); 52.0 (41.0, 59.0)

LQ, lower quartile; UQ, upper quartile.

TABLE 21 The PEECH scores by experimental group (at baseline)

| PEECH | Trial group ($N = 168$), mean (SD); median (LQ, UQ) | |
	CLECC ($n = 105$)	Control ($n = 63$)
Security (0 to 3)	2.48 (0.55); 2.50 (2.17, 2.83)	2.36 (0.51); 2.50 (2.00, 2.83)
Knowing (0 to 3)	2.18 (0.82); 2.33 (1.83, 3.00)	2.30 (0.72); 2.33 (1.67, 3.00)
Personal value (0 to 3)	2.34 (0.57); 2.50 (2.03, 2.80)	2.35 (0.58); 2.50 (2.00, 2.80)
Connection (0 to 3)	1.68 (0.74); 1.33 (1.17, 2.33)	1.61 (0.84); 1.67 (1.00, 2.33)
Total PEECH score (0–66)	49.2 (11.5); 51.0 (43.3, 57.0)	48.4 (12); 51.0 (42.0, 59.0)

LQ, lower quartile; UQ, upper quartile.

An analysis was also performed of frequency of low PEECH scores, that is, when patients scored on average 2 or below (*Table 22*). Ward D had the highest proportion of patients with lowest scores for each subscale. No one ward consistently had the lowest proportion of patients with low scores for each subscale.

Table 23 indicates similar frequencies of patients with low scores for each subscale between intervention and control wards.

Nursing empathy
Levels of self-reported empathy using the JSE from the nursing survey varied across the individual wards, with wards B and F scoring lower than the mean for all wards, and wards C, D and E scoring higher (*Table 24*). Higher scores indicate higher empathy.

Table 25 indicates that intervention wards at baseline had a higher mean and median empathy score than control wards.

TABLE 22 The PEECH frequencies of patients with low scores by ward (at baseline)

| Subscale | Ward, n (%) | | | | | | |
	A ($N = 25$)	B ($N = 19$)	C ($N = 20$)	D ($N = 31$)	E ($N = 30$)	F ($N = 43$)	Total ($N = 168$)
Security	5 (20)	5 (20)	3 (10)	11 (38)	3 (10)	8 (20)	42 (25)
Knowing	8 (33)	8 (32)	10 (35)	12 (43)	10 (36)	12 (29)	68 (41)
Personal value	5 (20)	5 (20)	4 (14)	9 (32)	4 (14)	10 (23)	44 (26)
Connection	16 (64)	16 (64)	17 (61)	20 (77)	16 (64)	30 (77)	123 (73)

TABLE 23 The PEECH frequencies of patients with low scores by experimental group (at baseline)

| Subscale | Trial group ($N = 168$), n/N (%) | |
	CLECC ($n = 105$)	Control ($n = 63$)
Security	23/100 (23)	15/59 (25)
Knowing	40/97 (41)	22/59 (37)
Personal value	25/100 (25)	17/63 (27)
Connection	67/93 (72)	43/58 (74)

TABLE 24 Staff mean empathy (JSE) by ward (at baseline)

Empathy (20–140)	Trial group (N = 91)					
	CLECC (n = 52)				Control (n = 39)	
Ward	A (n = 12)	B (n = 5)	D (n = 13)	E (n = 22)	C (n = 18)	F (n = 21)
Total score (SD)	113 (13)	112 (18)	120 (13)	115 (14)	115 (10)	107 (17)

TABLE 25 Staff empathy (JSE) by experimental group (at baseline)

| Empathy score (20–140) | Trial group (N = 91) | |
	CLECC (n = 52)	Control (n = 39)
Mean (SD)	115 (14)	110 (14)
Median (LQ, UQ)	117 (105, 127)	113 (102, 122)
Minimum to maximum	81 to 139	77 to 133

LQ, lower quartile; UQ, upper quartile.

Chapter summary

Six out of seven ward teams invited to take part participated in the study and all remained in the study until data collection was complete. Ward teams were randomised to intervention (n = 4) or control (n = 2), and staff were amenable to the prospect of randomisation to either experimental condition.

This chapter has described the six participating wards at baseline assessment across a range of characteristics including specialty, patient dependency, staffing, ward leadership, ward climate, staff well-being, care quality, patient safety and core study baseline measures. The findings reflect a range of ward contexts, with similarities across some dimensions, such as bed numbers. Some differences between wards were evident, including staffing levels and duration of ward leadership. Matrons' assessments of ward quality indicate that for all but one low-scoring ward, matrons assessed the care quality as good but not of the highest possible quality. Our qualitative interviews with staff did not identify particular differences between the wards, and, although all talked about the value of relational care, staff also reflected that limited time with individual patients constrained their capacity to be compassionate. Most staff–patient interactions observed were rated as positive. However, a proportion of interactions on all wards were rated negatively. In total, 40% of patients observed for 2 hours were rated as having one or more negative interactions during the 2-hour period. This proportion varied widely between wards (with intervention wards having a higher proportion than control wards) but indicates room for improvement on all participating wards. Staff survey results indicated that staff on all the wards showed signs of burnout; this was the case for over one-third of respondents. However, low response rates reduce certainty in the representativeness of these proportions in the staff group as a whole. Some individual staff on all the wards rated their empathy levels as low, others as high, and mean empathy levels for ward teams varied, with intervention wards having higher reported empathy. Low staff survey response rates suggest that these results need to be treated with particular caution. Overall, although these findings indicate some differences between the wards at baseline, there are also some remarkable similarities, particularly in relation to the quality of relational care, which is largely positive, yet with clear scope for improvement, and in relation to the presence of burnout in the staff group.

Although we were not successful in recruiting high proportions of nursing staff to take part in the written survey, the staff that did take part represented a wide range of characteristics in terms of work role, seniority, professional experience, age and ethnic group. Recruitment of patients and visitors was more successful, with the observations of staff–patient interactions attracting a particularly high proportion of recruits. Although not as high, recruitment to patient survey was still at a good level. Most patients involved in observations and the survey were older and female. One-quarter of patients who were observed had evidence of cognitive impairment, compared with 12% of questionnaire respondents. A high proportion of visitors approached to complete a questionnaire agreed to do so, with an average age of 62 years. The findings also indicate that the characteristics of participating staff, patients and visitors were equivalent across the experimental groups.

Chapter 8 Process evaluation results

This chapter reports the findings of the process evaluation that was undertaken as part of the feasibility study. The evaluation aimed to identify and explain the extent to which the planned intervention was implemented into existing nursing practices and to draw conclusions about optimising the sustainability of future CLECC interventions. The process evaluation was guided by the principles of NPT and used a mixed-methods approach.

In total, 47 interviews were conducted, over three rounds, with 25 people. Interviews were conducted with a range of staff members, including ward leaders ($n = 4$ people), deputy ward leaders ($n = 2$), staff nurses ($n = 8$), nursing HCAs ($n = 7$), senior hospital nurses ($n = 2$) and CLECC PDNs ($n = 2$). All four wards and all ward nursing staff roles were represented at each interview round. Senior hospital nurses were interviewed just once in the third round. In total, 13 people were interviewed once, two people were interviewed twice and 10 people were interviewed three times. The ward-based interviewees had worked on their current wards for between 2 weeks and 14 years, with an average value of 4 years. Staff nurses and deputy ward leaders had between 2 and 26 years' nursing experience, with an average of 12 years' experience. HCAs had between 2 and 35 years' experience working as a HCA, with an average 13 of years' experience. All but one of the interviewees were female. Two ward study days were observed in full, in addition to five ward leader action-learning sets.

Sections of this chapter are reproduced in Bridges *et al.*[121] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.

Implementation overview

The CLECC intervention was introduced to four ward teams in two hospitals. Its implementation in each hospital was led during a 4-month implementation period (beginning June 2015) by a PDN who was recruited and employed by the hospital and seconded from another role. Over the course of the implementation period, the role of the PDNs was to work with the ward leaders and the wider team to promote ownership of CLECC, to develop the learning activities into ways of working and learning that were seen to be feasible and helpful to the team and to develop a plan for sustaining what was good about CLECC after the departure of the PDN. The PDNs were mentored in their preparation for the implementation period by Jackie Bridges from the research team, working with a range of learning materials to prepare them for their role. The PDNs from each hospital communicated with each other regularly and also met up with Jackie Bridges for mentoring every 2 weeks at the outset, and then less frequently as time progressed.

Three ward teams ran the implementation period at the same time, and a fourth team started a few weeks later. The implementation period began for each ward team with the study days. Study days were exclusive to each ward team (i.e. they were not shared by participating teams from the same hospital). Each ward team had three study days led by the PDN and one-third of team members (RNs and HCAs, day and night staff) attended each of the days. This arrangement was intended to ensure that all team members had a chance to attend a study day. At both hospitals, a final 'mop-up' study day was held to include staff from both wards who had not made it to any of the previous study days. Ward leaders were invited to participate in all three study days and all did, with the exception of one ward leader who arranged for a deputy to attend. Study days included an introduction to CLECC principles. Staff were invited to complete CC and FECC questionnaires on the study day and to give a written response to the phrase 'I feel valued when . . .'. They also undertook a values clarification exercise on compassionate care, but the main focus of activities on the day was to create opportunities for staff to talk together and listen to each other about what is was like to work on the ward and to develop a shared vision for care.

Ward-based learning activities followed the study days. The PDNs initiated and trained staff in the cluster discussions. PDNs trained staff nominated for peer observations, arranged the observations of care and analysed the results. PDNs also analysed the material that staff had completed on the study days and they encouraged meetings between the ward leaders and matrons. Reflective learning sessions for staff were planned by the PDNs who arranged for a senior nursing manager to visit the ward at the end of the implementation period to hear about how the team had used CLECC and what their plans were for taking their ideas forward. They also arranged and facilitated the action-learning sets for ward leaders. The action-learning sets were held for staff from both hospitals jointly so that the set was larger than two members and the PDNs could share the facilitation.

The process of normalising Creating Learning Environments for Compassionate Care into practice

The findings that follow illustrate the amount of cognitive and behavioural work that staff needed to do, individually and collectively, to normalise CLECC into practice. Although many of the individual elements of CLECC were possible to implement during the 4-month implementation period, sustaining this work beyond this time was difficult for some ward teams to achieve, and the findings that follow explain why this was the case. Findings are presented under the four NPT constructs and draw together data from all data sources.

Coherence: Creating Learning Environments for Compassionate Care as a limited set of concrete practices versus underpinning philosophy

There was clear evidence from the interview and observation data that staff at all levels were able to distinguish the CLECC intervention from current ways of working. People were able to articulate activities associated with the CLECC intervention that were new to them. Staff valued the principles behind CLECC, appreciating the focus on staff well-being and the consequent impact on patient care quality:

> *I thought they [cluster discussions] were a really good idea, not just to bring up problems but to also say – actually we did this this morning, this went really well – for someone to say – oh, that was good, or thank you, just thank you for your hard work this morning, we were really under pressure this morning and everyone's worked really well as a team, thank you.*
>
> *N008 (HCA)*

For RNs in particular, the CLECC principles aligned with their aspirations for successful team working and high-quality patient care. However, for HCAs who had not always worked in this type of environment for long, this was a new and generally welcomed way of thinking about their workplace. Beyond the activities staff were directly involved in, they struggled to visualise the purpose and potential of CLECC. Staff tended to associate CLECC with the cluster discussions that took place part way through each shift, thus providing an opportunity to gather as a team and check on each other's well-being:

> *So, whereas before they might know that orange bay is heavier than green bay, they might not necessarily have volunteered to go and help. Now they are much more aware that if they are going – well actually we're struggling – well, we're not, we'll come and help you and I think that's because of the [clusters] and the fact that we're all sitting down and going – is there anything we can do to help you? And if they are going – well actually I've got a really poorly patient, so I've been struggling with the others – right – well then – we'll come and help you. And it's made them more aware of each other.*
>
> *N003 (HCA)*

All staff attended the study days and, on prompting, were able to link these sessions with CLECC. Staff were used to attending study days for a range of purposes but commonly attended as individuals. Participating in a study day in which only other team members were present was considered unusual and was generally welcomed. Staff saw the study day as a way of ensuring that they were working together and an opportunity to engage with the ward vision, which was not previously explicit. The most important

aspect of the study day was the chance to get to know each other, which staff reported they had not had the opportunity to do previously:

> *We had the study days and they were all very good and I found that I got to know the different people within those study days, or how they felt and I thought – oh, I didn't know that. So that was useful.*
> *N001 (staff nurse)*

The ward managers and PDNs charged with facilitating CLECC were involved in a wider range of CLECC activities. For ward leaders, additional CLECC activities included regular action-learning sets and meetings with their matron in which they discussed CLECC, and they were better able than other staff to note distinctions from usual practice. Ward leaders and PDNs were able to articulate the underlying philosophy of CLECC and to identify associated practices that were derived from CLECC but which fell outside the prescribed CLECC activities. In addition, senior nurse managers who were not directly involved in the CLECC intervention were able to confidently describe the underpinning principles based on what they had observed during the implementation period. These accounts, although expressed differently from the accounts of staff on the ward, reflected what staff gained from the CLECC intervention, even though ward-based staff struggled to define or explain the intervention themselves. The tools that staff required were acquired through reflecting on their own practice with patients and staff and developing support strategies as individuals and as a team:

> *To me CLECC is about giving staff tools to ensure that they support themselves to do a hard job. So it's about providing – a nurse with the knowledge of what they need to deliver . . . compassionate care or high-quality person-centred care, whatever you want to describe it as – every day, at a high-quality standard, is what we have to aim for, but also with you having some insight into how your behaviour affects both your patient and your staff.*
> *SN002 (director)*

> *Um . . . I don't know. Well I would say – it's about . . . reviewing your – your practice about the care that you're giving to patients and to the staff team that you're working with. It's to look at – I don't know – more supportive strategies and a way of working together. I think identifying stress and anxiety in yourself because the job is – is stressful, you know, we're not robots.*
> *N012 (nurse band 6)*

Another influencing feature on coherence was the extent to which team membership and team leadership were transient on the participating wards. High staff vacancy rates resulted in wards frequently being staffed by nurses who were not trained in CLECC or who did not view CLECC as part of their role. Ward leaders reported in the ward profiles that bank or agency staff were used daily. Indeed, team membership and leadership shifted throughout the 12 months under study. One ward leader was appointed as a CLECC PDN and so her ward needed a new ward leader at the outset. Part-way through the implementation period one ward team was split up and relocated because of ward refurbishment and was brought back together again at the end. The ward leader led a smaller team in a different location during refurbishment and this smaller team continued to work with the CLECC intervention, with efforts being made to integrate other team members back into CLECC working practices once the whole team was reunited. The ward leader of another ward retired just after the 4-month implementation period. Therefore, just one of the four participating wards had the same team leader in the pre-implementation period, throughout the implementation period and still in place after 12 months of CLECC. There was turnover of other team members too. On average, 36% of staff left over the course of the study, which is consistent with turnover on the control wards. One senior manager viewed staff turnover as a result of CLECC as a positive outcome:

> *[CLECC] exposed some practices, provided a culture where people could talk openly about how other members of staff made them feel; there has been a bit of a churn, so maybe some people that needed to go. People have now felt they've got a voice and, again, if people aren't doing what we need them to do, then they need to go.*
> *SN002 (director)*

But there was no provision for inducting temporary or newly arrived staff into CLECC, limiting their opportunity to make sense of CLECC.

In summary, although ward staff appreciated the potential value of CLECC, their understanding of CLECC was limited to and shaped by the concrete activities that they experienced. Additional knowledge about the underpinning principles of CLECC did not filter down to the team as a whole, with no evidence that coherence improved over time from the original induction into CLECC activities.

Cognitive participation: staff keen to participate but not sure who should drive it forward

Staff were generally keen to participate in the CLECC intervention, but it was not always clear whose responsibility it was to ensure that this happened. The PDNs led the implementation of the CLECC intervention as planned during the 4-month implementation programme period. Each PDN worked simultaneously with two wards in their allocated hospital. University staff prepared the PDNs for their role and provided ongoing mentorship. The PDNs organised specified CLECC activities as per the CLECC implementation period for the ward teams and were a visible presence to the ward staff throughout the implementation period. The PDNs experienced challenges in ensuring that all the staff were exposed to the CLECC activities. This was primarily on wards where some staff worked set shifts and were never on shift during the organised activities.

The approach of the individual PDNs to facilitating the CLECC activities influenced the degree of ownership of the CLECC intervention by the ward staff. One of the PDNs continued to be the stimulus for these activities throughout the 4-month period, whereas the other PDN deliberately undertook to transfer ownership for making things happen within the ward team:

> *The ward hadn't bothered doing their clusters . . . all I'd asked them to do was their 2 weeks of cluster meetings. They didn't do one.*
>
> *N036 (PDN)*

The PDN that encouraged the team to take ownership of the CLECC intervention actively worked with staff to make the initiative flexible and fit with the resource pressures:

> *It [cluster meeting] doesn't always stick to that time. It kind of depends how it's going. So we've had like busy days when stuff's been happening on the ward. At one point they [nursing staff] kind of ask permission to make it [cluster meeting] later, it's kind of sad. But I'm like . . . 'yeah, do it whatever time it works in the ward. If we can do it, that's a bonus'. So quite often it's – quite often it's the Health Care Assistants asking for it [cluster meeting].*
>
> *N035 (PDN)*

Although the PDN and ward leader/shift leader had initially been actively involved in originating the cluster discussions, as the intervention became embedded other team members took it upon themselves to call the cluster meeting in the absence of more senior leadership. In some cases it was the HCAs who led the discussions and on one ward the housekeeper became actively involved:

> *They [HCAs] will remind whoever is in charge of the ward, and say 'Are we having a check in today?' I've seen that quite a few times.*
>
> *N035 (PDN)*

The CLECC intervention gave staff the opportunity to see themselves as innovators. The study days provided a mechanism through which staff could articulate their ideas for improving practice on the ward and stimulate the team to set shared goals. Some ideas had been raised in the past but had not been achieved. The staff felt more empowered to act on the ideas and to set things up so that the change was sustainable.

Rather than working only at the ward level, staff raised their ideas with the matron, in addition to the ward leader, who was able, in theory, to provide the resource to make the initiative a reality. Previously the matron would not have been aware of the strong feeling held by the ward team about certain direct patient care-related issues as communication was usually through the ward leader alone:

Quite a few of the staff have got involved in various different things that have come out of the study days, what they wanted to change, and thought they could do better. And they've gone off and sort of little groups, or twos and threes, and are bringing that stuff back, passing it through the matron; putting a lot of that into place. It's things like patient's families not bringing in toiletries; [patients] don't have clothes to go home [in] . . . getting [relatives] to bring in a set of clothes and make them leave them in the locker. Some of [the staff] are taking more initiative themselves, rather than waiting to be almost told or suggestions. From that point of view, I think it's helped that way.

N030 (ward leader)

Not all the ideas were implemented in practice, and this appeared to be linked to uncertainty about whose role it was to realise them or to authorise them. This lack of follow-through was demoralising for the staff involved who were keen to put forward ideas and action plans:

So I think they – some of them felt a little bit disappointed that they'd made these suggestions and took their time to – to do them and then no one really followed it through or said – yes, we can use that or no we can't. It just got left.

N001 (staff nurse)

Everyone interviewed reflected that they saw CLECC as a way to build the team and improve care, and this underpinned their participation in prescribed activities. All ward team members attended the study days, and all ward leaders (except one who sent the same delegate each time) attended all of the action-learning sets. Cluster discussions were also reported as well attended, suggesting that people saw their engagement and participation as important. However, the participatory role of staff outside the ward team was less clear. Fortnightly meetings were encouraged between ward leaders and their matrons (usually the ward leader's line manager), but these did not go ahead for two of the ward leaders in one hospital, indicating a lack of clarity about the role of the matron in making CLECC happen:

Unfortunately this has been the only element of the programme to have failed the intervention wards and both ward managers felt that there has been a negative impact from the lack of support from the matron. It was felt that items identified by the nursing teams that were considered areas requiring improvement and determining a solution for implementation were unsupported and even in some instances rejected.

N036 (PDN field note)

Two of the ward leaders (at the second hospital) already attended regular meetings with their matron and chose to use this already established forum to address the CLECC intervention. These already established meetings appeared to be linked with a more proactive matron role in supporting the CLECC intervention:

So my matron's been very supportive the whole way through; we've kept in regular contact all the way through. She's been asking for updates, she's known about the interventions that we've done on the ward and has been really supportive.

N034 (ward leader)

Although the majority of staff were keen to participate in the CLECC intervention when invited, there was variation by work role at all levels and between hospitals and wards, in the extent to which individuals saw it as their role to make CLECC happen.

Collective action: participation shaped by organisational context

Whether or not the activities went ahead as planned was mediated by the extent to which the proposed activity harmonised with the priorities of the wider organisation and the resources made available to the ward team. A particular influence was the organisational priority afforded to material patient care activities over CLECC activities in the context of high patient care workloads. The flexibility of the CLECC intervention enabled staff to try other ways of implementing CLECC that partly overcame these barriers.

Staff reported that senior hospital managers had endorsed the work that had resulted from the CLECC intervention, suggesting that the benefits were visible and valued outside the immediate ward team:

> They seemed to be really positive about it and they said – 'if this is working for you, continue'. And she said 'it does seem to be working because you've got so many things put in place, like a suggestion box, etcetera' and she said – 'if this is working for you, you know – go for it.
>
> N009 (HCA)

In spite of this support, and also reflected in the senior nurse interviews, staff members' participation in planned learning activities was viewed as being of secondary importance to their role in providing patient care. The study days were designed in the original CLECC programme to be for the whole team, but managers were concerned that alternative staffing options for the ward would compromise patient care. Instead, ward teams were divided into three, with each third attending a separate study day. The option to deviate from the protocol, while still maintaining a degree of fidelity to the principles of CLECC, was essential to the implementation of the intervention from the outset:

> It wouldn't really be that compassionate to our patients if we left no staff whatsoever on the wards in order to train them. So, you know – I think that flexibility about how you do that – so I think whatever comes out needs to be about actually the important bit is that there's time and training and support.
>
> SN001 (director)

Cluster discussions proved possible to integrate into the working day and went ahead during the 4-month CLECC implementation period, yet they were less readily convened when patient care demands were very high and staffing resource was low. The consistency with which the cluster discussions took place varied with who was on shift and reflected the perceived priorities by the team of the organisation at that time.

Twice-weekly 1-hour reflective discussions as a group were planned, but, on all wards, the demands of patient care in relation to the staff available meant that it was not possible to release staff to attend them. In fact, staff other than the PDNs and ward leaders seemed unaware that the sessions were part of the CLECC intervention. Absence of the sessions meant that there was no forum in which to share with the team the results of the team climate analysis and values clarification exercise from the study days, or the peer observations of practice. Once it became clear that group sessions were not possible, PDNs experimented with different ways to encourage staff reflection. In one ward, the ward leader met with staff on a one-to-one basis. On other wards, results from peer observations were displayed for staff to see.

Interestingly, the part of the ward team relocated with the ward leader because of the refurbishment project (see *Coherence: Creating Learning Environments for Compassionate Care as a limited set of concrete practices versus underpinning philosophy*) found a focus on care and compassion much easier to integrate into their working day on the smaller ward and also found that the CLECC intervention had a significant impact on this as a choice of ethos by the ward leader:

> So there's a different ethos at the moment as well, because we're focusing on care and compassion.
>
> N033 (ward leader)

The ward leader's philosophy drove this focus, but more generous staffing levels and smaller patient numbers were seen as making it possible:

> So . . . because of the staffing levels, lack of, you know, less patients, it's much easier to be more caring and more compassionate and talk to people because you're not trying to – you know – wash somebody and finish quickly and move on to the next person. So you have got time for those nice little conversations, you know, talking to people about what they did before and – you've got time to spend with patients.
>
> *N033 (ward leader)*

Staff also felt less need for formal CLECC activities, such as cluster discussions, while on this smaller ward, as they created other, less formal, opportunities to catch up with each other throughout the shift. Once back as part of a larger team, cluster discussions were seen as necessary again and were resumed by the team. However, staff struggled to find time as lower staffing ratios resulted in them feeling too busy to stop.

As the most visible representative of hospital managers, the support role of the matron in relation to the CLECC intervention featured frequently in interviews. The extent to which the ward team perceived that they were supported in their endeavours by the matron was viewed as a strong mediator of whether or not CLECC was a success in influencing care. Some staff were accustomed to matron involvement only when there were problems, and these teams did not seem to be supported by their matrons to make changes that had arisen from the CLECC intervention. A lack of support with the CLECC intervention appeared to be linked with staff dissatisfaction with support from the matron in general:

> We don't hardly ever see a matron; the only time we ever see her is when she comes on and moans at us – or has something bad to say. She doesn't come on and praise – encourage or show that she's interested in the patients.
>
> *N008 (HCA)*

As noted earlier (see *Coherence: Creating Learning Environments for Compassionate Care as a limited set of concrete practices versus underpinning philosophy*), other matrons were reported to be very supportive, keeping in regular contact about the progress of the intervention. There appeared to be a disparity between senior management's expectations of the matron's role in the CLECC study and their actual involvement. At interview, one senior nurse manager reflected that their own role could have been more proactive in encouraging matron support for implementing CLECC:

> I assumed that my matron was working, was working with the ward managers on a weekly basis but I doubt it was what I expected it to be. So – I think we probably should have – should – could – have put more nursing leadership resources into it, just to provide that support and recognise it.
>
> *SN002 (director)*

Staff generally participated in CLECC activities when the opportunity arose, but this was shaped by the extent to which participation harmonised with the priorities of the wider organisation and that resources were available to enable participation. Planned staff learning activities took second place to direct patient care and transient team membership and leadership meant that including the whole team in the CLECC project was not possible. Matron support mediated the extent to which teams were able to implement CLECC. The CLECC property of plasticity enabled staff to develop and adapt practices that suited local circumstances, but what was possible was constrained by the available resources and priorities of the wider organisation.

Reflexive monitoring: valued by staff but challenging to sustain

The findings indicate that staff appreciated and were able to make use of the opportunities presented by CLECC, resulting in reported benefits to personal well-being and capacity to care. Staff spoke of engaging more consciously and deliberately with patients as people, and prioritising this engagement over the completion of tasks. They recognised that their practice was already compassionate at times but CLECC

had given them opportunities to value these practices and to make personal commitments to be more consistently compassionate:

> CLECC, for me, is about giving the staff the empowerment to feel like they can sit and do things with patients that are – compassionate rather than task orientated, so rather than just doing the obs [observations] and just doing the washes, having that – even if it's just 5 minutes, just having a chat with the patient about – their life, their family or sitting and doing an activity with them; those rather than just – well we've got to get the washes done, we've got to get the observations done – which do still need to be done but it's about – giving the staff that – yes – that empowerment of being able to say, well actually, let's do something a bit different.
>
> *N034 (ward leader)*

The CLECC intervention was associated with an improvement in staff morale and staff well-being more generally, which was viewed as having a positive impact on patient care. Cluster discussions were often cited as the means through which staff became aware of each other's needs and identified ways to give each other support. The cluster discussions provided the space to plan when and where additional support might be required, resulting in staff feeling that they were not on their own:

> So, whereas before they might know that orange bay is heavier than green bay, they might not necessarily have volunteered to go and help. They might have done but they might not, whereas now they are much more aware that if they are going – well actually we're struggling – well, we're not, we'll come and help you and I think – I think that's because of the check-ins [cluster discussions] and the fact that we're all sitting down and going – is there anything we can do to help you? And if they are going – well actually I've got a really poorly patient, so I've been struggling with the others – right – well then – we'll come and help you. And it's . . . I think it's made them more aware of each other.
>
> *N003 (HCA)*

Interestingly, some of the legitimacy for these practices seemed to come from the fact that they were part of the CLECC intervention. One interviewee cited an instance in which a senior manager visiting the ward came across a cluster discussion, which was also used by some teams to make sure that staff had a drink of water:

> I don't know who it was, but someone very high in the hospital [came to the ward] and was like, 'mmm why are people standing and drinking on the corridor?'. But the manager said, 'oh we have CLECC', so . . .
>
> *N025 (HCA)*

Once the manager was told that the cluster discussion was part of CLECC, she was reported to have then understood the purpose behind an activity considered to be unusual enough to be remarked upon.

The study days provided an opportunity for staff who had been working together for a significant time period to get to know each other, which facilitated better team working. Staff demonstrated that they could work together as a team without having to be directly led by someone more senior. The improved team working has reduced the work burden for some staff and has provided opportunities for staff to undertake activities that previously would have been rare occurrences:

> But because of the task orientated work – we've managed to go, right, we've finished, [they] haven't and then so we can go, right, we'll give you guys a hand and then we can all be finished together. And then that means we've got more time to do things that we might not be able to normally do, like – wash someone's hair, give them a nice – you know – do their nails and – yes. So it brings out that we – because of the cluster [discussions] – we can focus on who needs a hand and who doesn't and then we can all try and just do a bit more than – we might not be able to, like, you know, massage someone's feet, just the little extras that we might not be able to do normally.
>
> *N009 (HCA)*

Staff were generally in favour of CLECC, and, although some staff have taken ownership of particular aspects of CLECC, for example cluster meetings or doing things differently, there was reliance on the ward leader to facilitate continuation. Staff expressed concern that a change in leadership would result in the loss of a CLECC culture:

> *I hope it goes on for a long time because – I don't know – I don't want it to end; I like it because – and I think everyone else likes it. I think people would miss it, but I think – I don't know, I don't know. Things change once you have different managers and things like that, so hopefully it won't change when [ward leader] leaves but if we get a new manager and she doesn't like it, then – I don't know, I don't know, to be honest.*
>
> *N005 (staff nurse)*

The sustainability of CLECC was of concern to staff and the data highlighted what happened after the formal implementation period had come to an end. The principles that underpin CLECC appeared to be well embedded into the teams, but the activities that support these principles had not continued on all the wards 12 months after the start of CLECC. Individual wards varied in whether or not they had developed and followed through with a sustainability plan for CLECC. Ward leaders wanted to repeat the study days, but felt that they would not happen in future without the additional funding that came with CLECC as a research study. The ward leader action-learning sets, although reported to provide valuable learning opportunities, had not continued and were not missed. Ward leaders felt able to find the space and support offered by the action-learning sets elsewhere in already established systems. The formal reflective discussions were so sporadic that they did not feature in any CLECC sustainability plans. Peer observations provided a significant learning opportunity for those who did the observations, but without the space for formal reflective discussions to feed back the results to the team, these had quickly been forgotten as being part of the CLECC intervention.

The continuation or not of cluster discussions appears to be the most significant indicator of the sustainability of CLECC. All of the wards continued with the cluster discussions, but some were more sporadic than others. The ward that included a drink for staff in the cluster meeting had the highest meeting rate; however, all the wards were empowered to 'check in' with each other and not rely solely on handover. Attention to supporting each other appeared to increase the relational capacity of individual team members and the team as a whole. Although staff were not able to comment on whether or not patients and carers believed that care on their wards was more compassionate, staff had improved awareness of what allowed them to provide compassionate care and what hindered them. The cluster discussion was a tool that they did not have before that helped them to manage the challenges that they perceived as being out of their control.

Chapter summary

This chapter shared findings aimed at identifying and explaining the extent to which the planned CLECC intervention was implemented into existing work practices, to enable conclusions to be drawn about how interventions of this kind can be optimised to support sustained compassionate care delivery in acute settings. Findings reflect that some but not all CLECC activities were feasible to implement into practice, with a variety of factors influencing their impact and sustainability. Staff were generally keen to participate and valued the positive contribution of CLECC to their own well-being, to more cohesive team working and to supporting good patient care. Many original CLECC practices did not continue beyond the implementation period, but staff reported that the philosophy and associated culture that CLECC had nurtured continued to guide their practice. Sustainability was strongly linked by staff to the extent to which the ward leader understood and valued CLECC.

Findings indicate that CLECC had some coherence for staff in that they appreciated its potential value, but their understanding was often limited to the concrete activities of which they had direct experience. In terms of cognitive participation, staff were keen to participate, but ward teams varied in the extent to

which individual members saw it as their role to make sure that CLECC happened, and there was uncertainty about the role of matrons in supporting CLECC. There was strong evidence of collective action to implement CLECC, with team members generally participating when the opportunity arose. However, the ability to act or not was shaped by the extent to which activities harmonised with the priorities of the wider organisation. Although staff valued CLECC, its sustainability was linked to factors outside the direct control of the nursing team. These findings are discussed further in *Chapter 11*.

Chapter 9 Feasibility of evaluating effectiveness

A number of data were gathered and analysed to inform the design of a future study evaluating the effectiveness of CLECC in improving compassionate care. This chapter presents findings on the feasibility of evaluating effectiveness using a CRT design, including an assessment of bias and an outcome measure performance. It also addresses the feasibility of measuring CLECC costs and benefits.

Pilot cluster randomised trial assessment of bias

Selection bias

Selection bias was assessed by reviewing the demographic characteristics for each group that was recruited. These assessments focused on the representation of a wide range of nursing staff among the nursing questionnaire respondents, and on the inclusion of older patients and patients with cognitive impairment in observations and questionnaire completion.

Nursing staff

As noted in *Nursing questionnaires*, nursing questionnaire respondents represented a range of ages, ethnic groups, job roles/bands and years of experience. However, the low overall response rate to the nursing questionnaires (36%) reduces the certainty that the views of the respondents represent the views of all the staff working on the participating wards.

Participation rates of older patients and patients with cognitive impairment

Observations

In 17% (i.e. 63 out of 362) of approaches the patient was assessed as not having the capacity to make the decision to take part in the research. The care of 133 patients at baseline and of 140 patients at follow-up was observed. Mean patient age was 83 years at baseline and 80 years at follow-up.

At baseline 31% ($n = 41$) of patients showed evidence of cognitive impairment. This figure was lower at follow-up at 19% ($n = 27$). Overall, 25% (i.e. 68 out of 273) of patients observed had evidence of cognitive impairment.

Patient questionnaires

Across both assessment periods, of the patients approached to complete a questionnaire, 29% (i.e. 178 out of 621) had cognitive impairment. Of these, 117 (66%) were excluded. The most common reasons for excluding patients once approached were recorded as 'very cognitively impaired' ($n = 44$, 38% of 117 excluded) and 'no capacity' ($n = 59$, 50%). Of those patients with cognitive impairment who were assessed following approach as eligible for inclusion, 70% (i.e 43 out of 61) consented to take part and, of these, 98% (i.e. 42 out of 43) returned a completed questionnaire (*Table 26*).

Of the patients approached to complete a questionnaire at follow-up (baseline data not collected), 86% (i.e. 300 out of 347) were aged 65 years or older. Reasons for excluding these patients once approached were recorded as 'no capacity' (48%, i.e. 48 out of 101) and 'very cognitively impaired' (32%, $n = 32$). Of those assessed as eligible for inclusion, 79% (i.e. 156 out of 197) consented to take part and, of these, 91% (i.e. 142 out of 156) returned a completed questionnaire.

Detection bias

At follow-up it was not possible to conceal the allocation from nurses completing the questionnaires because they knew whether or not their ward had taken part in the intervention.

TABLE 26 Participation of people with cognitive impairment in patient questionnaires

Consent and return rates	Patient, *n* (%)
Baseline	
Approached	59 (23)
Excluded	35 (59)
Declined	9 (15)
Consent rate	15 (63)
Return rate	14 (93)
Follow-up	
Approached	119 (34)
Excluded	82 (69)
Declined	9 (8)
Consent rate	28 (76)
Return rate	28 (100)
Total	
Approached	178 (29)
Excluded	117 (66)
Declined	18 (10)
Consent rate	43 (70)
Return rate	42 (98)

For follow-up observations, a team of researchers ($n = 8$) who were otherwise not connected with the study were recruited and trained. They were not directly informed of ward allocation by the research team, although two reported that they learned the ward allocation from ward staff during data collection.

Researchers involved in distributing and helping patients and visitors with questionnaire completion were not blinded to ward allocation. Patients and visitors completing questionnaires were invited to speculate on whether or not they were on an intervention ward. At follow-up, 82% (i.e. 28 out of 34) of the control ward patients asked to reply 'yes' or 'no' thought that they were on an intervention ward, compared with the 64% (i.e. 55 out of 86) of intervention ward patients who correctly identified that they were on an intervention ward. Researcher field notes record that patients found this questionnaire item confusing and that researchers often had to reword it to aid understanding. For visitors at follow-up (who were also offered 'don't know' as an option), 16% (i.e. 5 out of 31) of control ward visitors and 35% (i.e. 18 out of 51) of intervention ward visitors thought that they were on an intervention ward. Most visitors (61% on control wards and 59% on intervention wards) said they did not know. This method of assessing blinding may not be valid but these findings suggest that study ward allocation was mostly concealed to patients and visitors completing questionnaires.

Contamination
One concern with interventions, such as CLECC, that aim for behavioural change, including collaborative behaviours, especially when intervention and control conditions are operating in parallel in the same organisation, is whether or not intervention practices 'contaminate' control ward practices. A comparison of staff names on each ward between baseline and follow-up did not reflect that any intervention ward staff had gone to work on either of the control wards. This finding, however, does not preclude the possibility that intervention ward staff joined the control ward and then left again before follow-up data collection. Researcher field notes reflect a conversation with a matron who oversaw an intervention and a control ward. The intervention ward leaders had reported the value of the CLECC cluster discussions and at a meeting of the ward leaders from that specialty (that included a control ward leader), a decision

was taken to adopt cluster discussions across all of the wards, with the exception of the control ward. There was no evidence that the control ward went on to adopt these practices anyway but the possibility cannot be excluded. In the other hospital, intervention and control wards were managed in different specialties and so similar mechanisms would not be in place, but in this hospital the director of nursing visited the intervention wards and reportedly expressed an interest in extending the CLECC intervention to other wards in the hospital. Thus, although no evidence was found of contamination, we did identify pathways within organisations through which this could happen.

Pilot cluster randomised trial outcome measure performance

This section reports on the performance of the three core outcome measures: (1) quality of staff–patient interaction using QuIS observations, (2) nursing staff self-rating of empathy using JSE in the nursing written survey and (3) patient evaluation of emotional care using the PEECH instrument in the patient written survey. It also reports on the performance of the EQ-5D-5L in measuring benefit as part of the economic evaluation.

Quality of Interactions Schedule

Using observations as a method of collecting data appeared to be highly acceptable to patients, staff and visitors. As discussed in *Observations*, a high proportion of patients approached agreed to take part. Staff and visitor consent was not explicitly sought, although they were invited through written study information to raise objections if they did not wish a planned observation to go ahead or if they wished for the observation to be halted. Patients were also informed that they could halt the observations at any time. No planned observations were cancelled or halted because objections were raised. Although inviting staff or visitors more explicitly to share their views may have led to some sessions not going ahead, no data were gathered that indicated this.

All but one of the 120 planned observation sessions went ahead as scheduled. On one occasion, an evening observation could not proceed because the researcher could not gain access to the locked office containing the computer tablet needed for data collection and so that session was rescheduled for the same time of day the following week. Observers were able to observe staff–patient interactions of up to four patients simultaneously. On some occasions (11%, i.e. 27 out of 273 patients), individual patients were observed for less than the planned 2 hours because they left the ward during the session and the protocol dictated that researchers did not follow them. If they returned to the ward during the 2-hour period, observation resumed. Data were gathered from Monday to Friday, from 08.00 to 22.00.

At the close of each observation session, researchers were asked to record if staff being observed reported changing their behaviour because of the observation. Of 120 observations sessions, researchers did not ask staff on 29% ($n = 35$) of occasions, recorded that staff reported changing their behaviour on 4% ($n = 5$) of occasions and reported staff not changing their behaviour on 67% ($n = 80$) of occasions.

In a separate exercise to the main data collection, inter-rater reliability was tested over six 2-hour observation sessions, involving three researchers. Each session included two of the three researchers. The nature of the QI Tool software design meant that it was possible to calculate the reliability of the number of interactions recorded by each observer in each session, but not the reliability of the quality of interaction rating. The ICC for the number of interactions recorded per observation session was 0.93 (95% CI 0.607 to 0.990; $p = 0.001$), indicating high reliability. In earlier feasibility testing in acute hospital settings similar to those in the main study, using manual methods for recording quality of interaction, we found close agreement between observers in relation to the number of interactions observed (ICC = 0.97) and moderate to substantial agreement on the quality of interactions (absolute agreement 73%, $\kappa = 0.53$–0.62, depending on weighting scheme).[99,122]

Because main study observation and questionnaire data were gathered from different patient groups, it was not possible to test the validity of QuIS ratings against patient-reported experience. In our earlier feasibility work, however, 17 patients who had been observed using QuIS were asked to rate interactions,

and 18 patients were able to complete a shortened version of PEECH. Patients without the capacity to consent were excluded. We found 79% agreement (weighted κ 0.40: $p < 0.001$; indicating fair agreement) between patients and QuIS observers over whether interactions were positive, negative or neutral.[99] This earlier work also found a significant correlation between the percentage of QuIS interactions that were rated positively and patient responses to the individual PEECH item 'exceeded expectations' on the personal value subscale (Spearman's $r = 0.603$; $p = 0.008$).[99] We found a moderate (but not statistically significant) association between the percentage of positively rated interactions and PEECH's 'facial expression' (Spearman's $r = 0.426$; $p = 0.088$) and 'social conversation' (Spearman's $r = 0.402$; $p = 0.098$).[99]

The QI Tool software performed well, enabling the accurate collection and transfer of interaction and contextual data.

Jefferson Scale of Empathy

The JSE was administered as part of the nursing survey. The whole questionnaire (not just the JSE) took a mean of 37 minutes to complete (SD = 50 minutes, minimum to maximum = 4 to 400 minutes). Completion time was calculated from inviting respondents to record the time at which they started the questionnaire and the time at which they finished. The longer completion times experienced by some individuals may have been because they were undertaking other activities rather than focusing solely on the questionnaire, but we did not gather any data to substantiate this. Completed questionnaires did not reflect that respondents had any difficulties completing them and there was no feedback about problems from ward staff. However, there was a perception that they were lengthy to complete and, therefore, hard for staff to find the time to complete them, and this may have affected response rates.

Patient Evaluation of Emotional Care during Hospitalisation

The PEECH instrument was administered as part of the patient questionnaire survey and, as noted in *Patient questionnaires*, had high acceptability with patients. Researchers helped patients to complete the survey in 68% ($n = 242$) of cases. The whole questionnaire (not just the PEECH) took a mean of 27 minutes (SD = 27 minutes, minimum to maximum = 4 to 330 minutes) and researcher field notes reflect that some patients found it too long. Field notes also reflect that having multiple response options presented by the PEECH's four-point Likert-type scale was confusing, although this was improved when researchers presented a separate board with the responses written in larger type and invited patients to point to their chosen response. Patients completed their questionnaire in or by their bed because many had mobility difficulties and it was not practical for researchers to relocate them for the purposes of questionnaire completion. Because researcher help was required for most patients to complete their questionnaires, this location meant that answers could not be given without the risk of being overheard by others and researcher field notes reflect that some patients seemed concerned by this. Giving patients the option of pointing to their preferred response rather than verbalising it seemed to alleviate these concerns. Researchers reported that some patients had difficulty understanding some of the questions and so researchers had to reword them in order to get a response.

EuroQol-5 Dimensions, five-level version, health status

The EQ-5D-5L was administered as part of the patient questionnaire survey to inform the economic evaluation. The instrument was initially located on the patient questionnaire, before the last item which asked for demographic details, but researchers found that starting the interview with asking about demographic details relaxed patients and was a good way to assess mental capacity. As a result, the EQ-5D-5L was often the last part of the patient questionnaire to be completed when researchers were helping with questionnaire completion.

Researcher field notes reflect that patients had difficulties understanding the EQ-5D-5L measure, especially the visual analogue scale, which requires subjects to mark on a numbered scale (0 to 100) how good their health is today. These concerns are reflected to an extent in the analysis of missing EQ-5D-5L data shown in *Table 27*. In total, 89% (i.e. 150 out of 168) of baseline patients and 76% (i.e. 141 out of 186) of follow-up patients gave responses in all five domains of the EQ-5D-5L.

The EQ-5D-5L scores ranged widely among patients, from 0.39 to 1 as shown in *Figure 8*. A total of 13% (i.e. 19 out of 149) of patients had a negative EQ-5D-5L score (< 0). There was no strong correlation between EQ-5D-5L score and age or gender.

Mean EQ-5D-5L index values across six wards and two experimental groups (*Tables 28* and *29*) are lower than general population by age (0.75 for age > 75 years). A lower index value represents a lower quality of life, with 1 representing the best health possible and 0 representing death. Lower than average values in this population are expected given that respondents were in hospital. The EQ-5D-5L score (0.412) in the intervention group was lower than that in the control group (0.502), but there was no statistically significant difference in EQ-5D-5L score between both groups (mean 0.09, 95% CI –0.026 to 0.205). EQ-5D-5L scores were different at follow-up. Given that different patients were involved at each point, these changes, which were both positive and negative, cannot be interpreted as changes attributable to CLECC. Only if the different groups of patients in each ward had the same ailments and severity at baseline and follow-up could changes be attributed to CLECC.

The results of the QALY analysis based on the EQ-5D-5L data showed large differences between wards at baseline and follow-up. This was not surprising given that different patients with different ailments and

TABLE 27 Summary of EQ-5D-5L data (at baseline and follow-up)

Time	Participants	EQ-5D-5L domain (*n*)					All (*n*)
		Mobility	Self-care	Activity	Pain	Anxiety	
Baseline	Missing	11	10	16	9	9	18
	Recorded	157	158	152	159	159	150
	Total	168	168	168	168	168	168
Follow-up	Missing	19	17	23	17	18	27
	Recorded	167	169	163	169	168	159
	Total	186	186	186	186	186	186
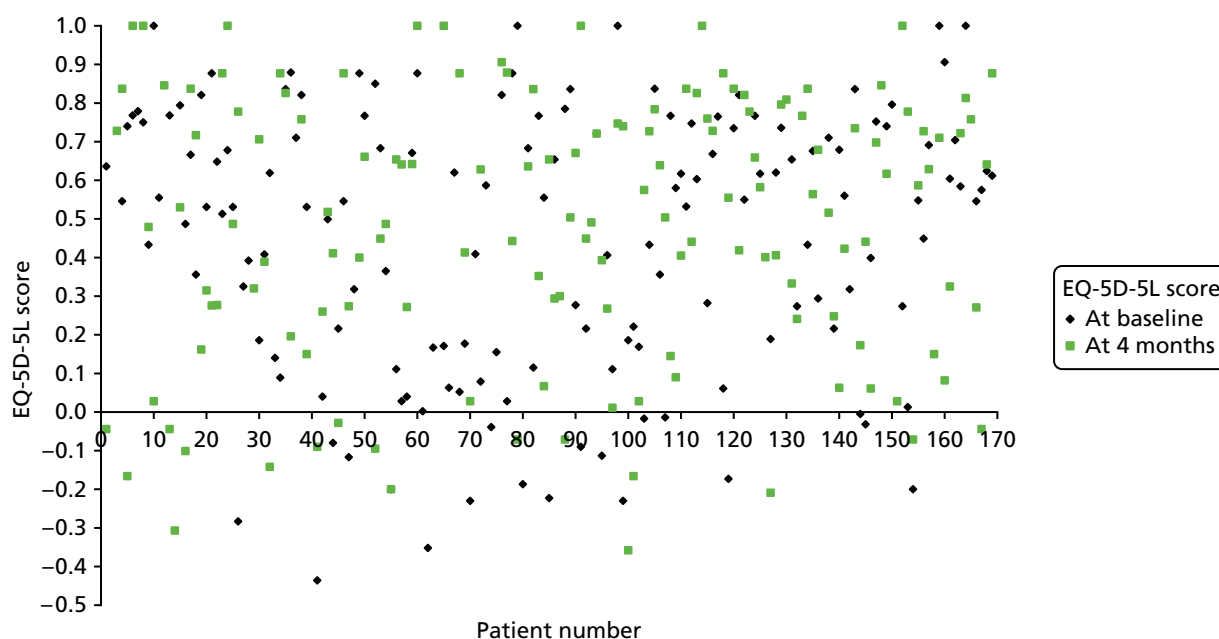
FIGURE 8 Scatterplot of EQ-5D-5L scores (at baseline and follow-up).

TABLE 28 The EQ-5D-5L scores by ward and experimental group (at baseline)

	Trial group							
	CLECC					Control		
EQ-5D-5L score	A (*n* = 25)	B (*n* = 19)	D (*n* = 31)	E (*n* = 30)	Total (*n* = 105)	C (*n* = 20)	F (*n* = 43)	Total (*n* = 63)
Mean	0.337	0.472	0.452	0.395	0.412	0.551	0.476	0.502
Lower 95% CI	0.189	0.281	0.332	0.262	0.344	0.376	0.353	0.404
Upper 95% CI	0.485	0.663	0.571	0.527	0.48	0.726	0.599	0.6
Minimum	−0.283	−0.436	−0.223	−0.2	−0.436	−0.352	−0.23	−0.352
Maximum	0.877	1	0.877	1	1	0.879	1	1

TABLE 29 The EQ-5D-5L scores by ward and experimental group (at follow-up)

	Trial group							
	CLECC					Control		
EQ-5D-5L score	A (*n* = 32)	B (*n* = 33)	D (*n* = 29)	E (*n* = 29)	Total (*n* = 123)	C (*n* = 31)	F (*n* = 32)	Total (*n* = 63)
Mean	0.314	0.545	0.463	0.419	0.433	0.651	0.572	0.602
Lower 95% CI	0.195	0.4	0.337	0.285	0.368	0.479	0.466	0.513
Upper 95% CI	0.433	0.69	0.589	0.553	0.497	0.823	0.677	0.692
Minimum	−0.2	−0.166	−0.2	−0.358	−0.358	−0.307	−0.209	−0.307
Maximum	0.846	1	1	1	1	1	1	1

severity were involved in each ward at baseline and follow-up. Consequently, although patients could be classified into Healthcare Resource Groups (HRGs; the basis on which NHS hospitals are paid), this does not include any measure of severity. Data on the quality of life of patients would need to be adjusted for severity if changes in EQ-5D-5L were to be interpretable. No agreed severity measurement exists and none is routinely collected in the NHS. Apart from diagnosis, procedure and demographic data, no other relevant data are collected. As a severity casemix measure was not used in this study, it is not possible to explain the differences in QALYs that were found.

Feasibility of estimating costs of Creating Learning Environments for Compassionate Care

Initial training costs were feasible to calculate and primarily focused on the cost of ward staff time attending the study day and the cost of PDN time to facilitate the implementation period (*Table 30*).

Future calculations of cost-effectiveness will depend on the length of time for which any benefits of CLECC are sustained beyond the initial investment of these training costs. If, for instance, the training costs are one-off costs with the effects sustained over an infinite period of time these costs become negligible when assessing cost-effectiveness and only ongoing implementation costs are relevant over time. Qualitative findings suggest that CLECC practices can be sustained for up to 12 months beyond the start of the implementation period. The cost of attaining and sustaining the benefits of CLECC over a 12-month period are therefore £53 per day per ward (£19,503/365), or £2 per bed-day (based on 30 beds per ward) [£19,503/(365 × 30)]. Were the benefits sustained over a 24-month period with no further training required, these costs would be halved.

TABLE 30 The CLECC training costs

Cost	£	Comments for future use
Cost of whole ward team attending an 8-hour study day (2014/15)	6646	Study day can be covered in 7 hours. No travel costs as the study day venue is at the usual place of work
Cost of 0.5 FTE × band 7 PDN to support CLECC implementation period on one ward. Over 6 months to include training for PDN and preparation for implementation period (2014/15)	12,857	May be able to reduce pre-implementation period phase to < 2 months
Total	19,503	

In terms of ongoing costs of implementation, the results from qualitative interview data indicated that nursing staff implemented CLECC cluster discussions on the ward in short 5- to 10-minute sessions. Those who commented on whether or not extra time needed to be scheduled for such sessions suggested that it could be met within existing time schedules. Sessions remained within the short 5- to 10-minute slots, which was not seen as onerous. When wards became particularly busy, the sessions did not take place. CLECC sessions were popular with all the nursing staff interviewed. CLECC 'champions' were seen as important in ensuring that sessions took place. The nursing staff interviewed reported that the implementation of CLECC did not require additional nurse time. They considered that ward leadership was more important to ensure that cluster discussions happened, especially at busy times.

Chapter summary

This chapter has reported a number of important findings that can inform the design of a future evaluation. Findings indicate no selection bias inherent in our study recruitment processes and methods. The methods employed were inclusive of all staff levels and also inclusive of older patients with a range of cognitive abilities, ensuring excellent representation from a traditionally hard-to-reach group who are often excluded from research but who are prone to more negative experiences of hospital care. There was some evidence that observer researchers could find out the experimental group of the wards on which they were observing and this will need careful attention in a future trial. In addition, the findings of pathways for contamination beyond the intervention wards mean that any future trial design will need to avoid running intervention and control conditions in the same organisation at the same time. A more positive interpretation of these same findings is that the CLECC intervention appears to have the potential for impact beyond the target ward teams and future evaluations should aim to explore this potential.

The QuIS tool was highly acceptable to patients recruited and had the highest rate of participation by patients with cognitive impairment of all methods. We demonstrated that our recruitment and data collection plans for QuIS use were feasible in busy ward environments. The validity and reliability of QuIS was acceptable. The QI Tool software performed well, enabling straightforward data collection and upload.

Patient questionnaires, despite being very acceptable to patients, had a lower participation rate than QuIS for people with cognitive impairment. Most patients needed researcher help and there were complaints that it was too long. The PEECH and EQ-5D-5L scales were hard for some patients to use.

The response rate to nursing questionnaires was very low, particularly at the follow-up stage on the intervention wards, perhaps owing to research fatigue. There was a perception that questionnaires were lengthy to complete and that staff were too busy. The low response rate means reduced certainty that the responses to the questionnaire represent the views of the nursing staff as a whole.

A patient outcome measure, such as EQ-5D-5L, does not appear to be feasible to use to evaluate value for money for two reasons. First, given that some two-thirds of patients needed help to complete the questionnaires, it could be used in a full study only as a proxy measure. Second, and more importantly, interpretation of EQ-5D-5L scores as measures of health improvement is not possible as different patients, with different ailments and severity were involved at baseline and at follow-up. Without an adjustment for the various ailments and their severity, differences in EQ-5D-5L scores cannot be interpreted.

Through qualitative interviews with staff, we were able to establish that, aside from initial CLECC training costs (cost of staff time at team study day and cost of employing CLECC PDN), the implementation of concrete CLECC activities by ward teams was not associated with additional resource use. Nursing staff reported that the CLECC cluster discussions typically took 5–10 minutes each shift. They were dropped when wards became too busy to have them. We were also able to identify other candidate activities associated with supporting CLECC that merit attention in a future evaluation.

Chapter 10 Pilot trial outcomes

The focus of this chapter on reporting the core outcomes of the pilot CRT is intended to inform the design of a future definitive evaluation, including the measurement of CLECC economic benefits. This chapter focuses on a comparison between experimental groups on the core outcomes for the study: quality of staff–patient interactions, patient evaluation of emotional care and nursing self-reported empathy. Intracluster correlation is also reported for each outcome. These data will enable sample size calculation and inform outcome measure selection and use in a future trial.

Quality of staff–patient interactions

The observed quality of staff–patient interactions, assessed using QuIS scores, between experimental groups at follow-up is shown in *Table 31*. The distribution of interaction scores across the five available categories is broadly similar to baseline distribution, with most interactions rated as positive care and fewest interactions rated as negative protective for each experimental group.

There are more positive (social plus care) and fewer negative (protective plus restrictive) scores for intervention wards than for control wards at follow up (78% vs. 74%, 8% vs. 11%). Chi-squared testing of these results suggested a significant difference between experimental groups ($p = 0.017$), but the results shown in *Table 32* indicate that, once other variables are taken into account in the analysis, the odds of a negative interaction are not significantly reduced because of the effect of the CLECC intervention. Results are in the direction of an effect favourable to CLECC, that is, there were fewer negative interactions on intervention wards, but this was not a statistically significant difference.

The proportion of negative interactions per patient was calculated on all of the QuIS data ($n = 140$ patients at follow-up). The results in *Table 31* indicate that, although some patients had no negative interactions, others had up to 67% of their interactions rated negatively. On average, 10% of interactions per patient at follow-up were rated as negative.

TABLE 31 Quality of staff–patient interactions QuIS score by experimental group (at follow-up)

QuIS rating (*n* interactions)	Follow-up (*N* = 1555)	
	CLECC (*n* = 1119)	Control (*n* = 436)
Positive social	243 (22%)	64 (14%)
Positive care	632 (56%)	260 (60%)
Neutral	151 (13%)	62 (14%)
Negative protective	36 (3%)	21 (5%)
Negative restrictive	57 (5%)	29 (7%)
Positive social + positive care	875 (78%)	324 (74%)
Negative protective + negative restrictive	93 (8%)	50 (11%)
Patients (*n*)	92	48
Percentage negative interactions per patient (minimum to maximum)	8% (0% to 56%)	12% (0% to 67%)
Patients observed for full 2 hours (*n*)	85	44
Patients with ≥ 1 negative interactions	38 (45%)	21 (48%)

TABLE 32 The QuIS multilevel logistic regression results: OR of a negative interaction

Variables	Model		
	1, unadjusted OR (95% CI) (*n* = 3111)	2, adjusted OR (95% CI) (*n* = 3111)	3, adjusted OR (95% CI) (*n* = 3111)
CLECC effect	0.72 (0.35 to 1.51)	0.47 (0.17 to 1.29)	0.30 (0.07 to 1.32)
Time period (Baseline vs. follow-up)		0.56 (0.22 to 1.43)	0.38 (0.11 to 1.32)
Ward			
A (CLECC)			1.00
B (CLECC)			0.60 (0.20 to 1.83)
C (control)			0.80 (0.21 to 3.05)
D (CLECC)			0.75 (0.24 to 2.35)
E (CLECC)			0.61 (0.19 to 1.90)
F (control)			0.23 (0.05 to 1.02)
Variance component estimates (95% CI)			
Observation session level (*n* = 120)	2.13 (1.25 to 3.62)	2.09 (1.23 to 3.55)	1.96 (1.14 to 3.37)
Patient level (*n* = 273)	0.51 (0.23 to 1.13)	0.51 (0.23 to 1.13)	0.51 (0.23 to 1.13)

In total, 92% of patients (i.e. 129 out of 140) at follow-up were observed for a full 2-hour period. For patients observed for the full 2-hour period, the number with one or more negative interactions during that 2-hour period was calculated (*Table 31*). Findings did not differ significantly by experimental group (chi-squared *p* = 0.744).

Table 33 illustrates the total negative interactions by individual ward. At the individual ward level, some wards (*n* = 3) appear to have improved their QuIS ratings and some appeared to have deteriorated (*n* = 3). The proportion of patients experiencing at least one negative interaction in a 2-hour period varies between individual wards from 25% to 64%. Two of the intervention wards showed an improvement between baseline and follow-up (i.e. the proportion of patients with negative interactions decreased), two showed a deterioration and the two control wards also showed deterioration.

TABLE 33 Negative staff–patient QuIS interactions by ward (at follow-up)

Interaction	Trial group (*N* = 1555 interactions)					
	CLECC (*n* = 1119 interactions)				Control (*n* = 436 interactions)	
Ward (*n* interactions)	A (*n* = 282)	B (*n* = 388)	D (*n* = 210)	E (*n* = 239)	C (*n* = 233)	F (*n* = 203)
Negative interactions, *n* (%)	16 (6)	31 (8)	39 (19)	7 (3)	29 (12)	21 (10)
Change from baseline	↓	↑	↑	↓	↓	↑
Patients observed for full 2-hour period (*n* patients)	CLECC (85)				Control (44)	
	21	22	18	24	23	21
Patients with ≥ 1 negative interaction, *n* (%)	7 (33)	14 (64)	11 (61)	6 (25)	14 (61)	7 (33)
Change from baseline	↓	↑	↑	↓	↑	↑

↑, deterioration (i.e. negative ratings increased); ↓, improvement (i.e. negative ratings decreased).

Patient evaluation of emotional care

Patient evaluations of emotional care using the PEECH tool, administered through written patient survey, did not differ significantly by experimental group (Mann–Whitney *U*-test, $p > 0.05$), although small non-significant differences in total score and three out of four subscales favoured CLECC (*Table 34*).

Fewer patients in the CLECC group than the control group had low scores in the connection subscale (63% vs. 79%) (*Table 35*), but this was not adjusted for any potential differences at baseline or in patient characteristics or ward effects.

The results shown in *Table 36* indicate that, once other variables are taken into account in the analysis, the odds of low connection scores are lower on the intervention wards, but not significantly so. Model 2 has been adjusted for ward, Model 3 for ward and baseline and Model 4 for ward, baseline and patient characteristics (age, gender, ethnicity and education level).

The results are in the direction of an effect favourable to CLECC, that is, CLECC may be associated with a reduction in the odds of having a low score in the connection subscale of PEECH. However, this association is no longer significant when we adjust for baseline and patient characteristics.

Patient evaluations of emotional care in each ward at follow-up are shown in *Table 37*. Higher scores indicate better patient-reported experiences. As at baseline, connection consistently scores lower than the other subscales. Total mean PEECH scores range from 47.6 on ward C to 53.8 on ward B (out of possible total of 66). There was a small improvement in total mean PEECH score summed for all the wards from baseline to follow-up [48.9 (SD 11.7) to 49.9 (SD 10.8)]. There are variations in the direction of change over time between individual wards, with ward E showing deterioration across all subscales and total PEECH score, and wards B, D and C showing improvement in all subscales and total PEECH score.

TABLE 34 Mean PEECH scores by experimental group (at follow-up)

PEECH score	Trial group (*N* = 186)		
	CLECC mean (SD) (*n* = 123)	Control mean (SD) (*n* = 63)	*p*-value[a]
Security (0 to 3)	2.48 (0.50)	2.46 (0.48)	0.653
Knowing (0 to 3)	2.19 (0.88)	2.26 (0.66)	0.800
Personal value (0 to 3)	2.43 (0.57)	2.31 (0.57)	0.071
Connection (0 to 3)	1.81 (0.82)	1.71 (0.63)	0.350
Total PEECH score (0 to 66)	50.6 (11.3)	48.5 (9.8)	0.116

a Mann–Whitney *U*-test.

TABLE 35 Frequencies of patients with low PEECH scores by experimental group (at follow-up)

PEECH score	Trial group (*N* = 186)		
	CLECC (*n* = 123)	Control (*n* = 63)	*p*-value[a]
Security	27/117 (23%)	11/55 (20%)	0.650
Knowing	42/112 (38%)	24/52 (46%)	0.293
Personal value	30/117 (26%)	14/56 (25%)	0.928
Connection	73/115 (63%)	41/52 (79%)	0.048

a Chi-squared test (not adjusted for baseline or patient characteristics).

TABLE 36 Logistic regression results: OR of a low PEECH connection subscale score

Variables	Model			
	1, unadjusted OR (95% CI) ($n = 318$)	2, adjusted OR (95% CI) ($n = 318$)	3, adjusted OR (95% CI) ($n = 318$)	4, adjusted OR (95% CI) ($n = 273$)
CLECC effect (CLECC vs. control)	0.60 (0.37 to 0.98)	0.47 (0.22 to 1.00)	0.51 (0.17 to 1.51)	0.47 (0.14 to 1.59)
Time period (baseline vs. follow-up)		0.72 (0.34 to 1.53)	0.75 (0.30 to 1.82)	0.76 (0.27 to 2.10)
Ward				
A (CLECC)			1.00	1.00
B (CLECC)			0.87 (0.39 to 1.98)	0.97 (0.36 to 2.61)
C (control)			1.04 (0.36 to 2.97)	0.56 (0.17 to 1.89)
D (CLECC)			1.59 (0.69 to 3.64)	2.10 (0.80 to 5.54)
E (CLECC)			1.15 (0.51 to 2.57)	0.97 (0.37 to 2.52)
F (control)			1.41 (0.55 to 3.60)	1.30 (0.45 to 3.75)
Patient characteristics				
Age (years)				
≤ 50				1.00
51–60				1.22 (0.19 to 7.90)
61–70				1.22 (0.21 to 7.00)
> 70				0.72 (0.16 to 3.16)
Gender (female vs. male)				2.08 (0.11 to 1.19)
Education				
Primary school				1.00
Secondary school				0.37 (0.11 to 1.19)
College				0.55 (0.15 to 2.06)
University				1.04 (0.20 to 5.42)

Nursing staff self-reported empathy

Levels of self-reported empathy using the JSE scores from the nursing staff survey (RNs and HCAs) varied across the individual wards at baseline and at follow-up. Higher scores indicate greater empathy.

Empathy scores by experimental group at follow-up are shown in *Table 38*. Mean and median empathy scores were similar across experimental groups at follow-up. A Mann–Whitney *U*-test confirmed no significant difference between groups ($p = 0.800$).

There was a small reduction in mean empathy score for all wards from baseline to follow-up (113 vs. 112). Mean empathy scores decreased on four wards from baseline to follow-up and increased on two wards (*Table 39*).

TABLE 37 The PEECH scores by ward (at follow-up)

| PEECH score | Trial group (N = 186) | | | | | |
	CLECC (n = 123)				Control (n = 63)	
Ward (n)	A (n = 32)	B (n = 33)	D (n = 29)	E (n = 29)	C (n = 31)	F (n = 32)
Security (0–3)						
Mean (SD)	2.37 (0.57)	2.59 (0.41)	2.45 (0.54)	2.54 (0.48))	2.48 (0.52)	2.44 (0.46)
Median (LQ, UQ)	2.67 (1.83, 2.83)	2.67 (2.17, 3.00)	2.67 (2.00, 2.83)	2.67 (2.17, 3.00)	2.50 (2.00, 3.00)	2.45 (2.17, 2.83)
Change from baseline	↓	↑	↑	↓	↑	↑
Knowing (0–3)						
Mean (SD)	2.23 (0.85)	2.40 (0.70)	2.08 (0.99)	2.05 (0.97)	2.33 (0.67)	2.21 (0.66))
Median (LQ, UQ)	2.33 (1.83, 3.00)	2.67 (2.00, 3.00)	2.33 (1.17, 3.00)	2.33 (1.67, 3.00)	2.17 (2.00, 3.00)	2.33 (2.00, 2.67)
Change from baseline	↓	↑	↑	↓	↑	↓
Personal value (0–3)						
Mean (SD)	2.32 (0.68)	2.63 (0.42)	2.43 (0.46)	2.32 (0.65)	2.25 (0.66)	2.36 (0.50)
Median (LQ, UQ)	2.60 (1.70, 2.93)	2.78 (2.38, 3.00)	2.50 (2.00, 2.80)	2.55 (2.03, 2.80)	2.60 (1.70, 2.73)	2.60 (2.10, 2.70)
Change from baseline	↓	↑	↑	↓	↑	↓
Connection (0–3)						
Mean (SD)	1.74 (0.85)	2.11 (0.82)	1.64 (0.85)	1.75 (0.72)	1.81 (0.55)	1.63 (0.68)
Median (LQ, UQ)	2.00 (1.00, 2.67)	2.33 (1.33, 3.00)	1.67 (1.00, 2.33)	2.00 (1.33, 2.33)	1.67 (1.67, 2.08)	1.67 (1.25, 2.00)
Change from baseline	↑	↑	↑	↓	↑	↑
Total PEECH score (0–66)						
Mean (SD)	48.6 (13.5)	53.8 (8.5)	50.1 (10.2)	49.6 (12.3)	47.6 (10.8)	49.2 (9.1)
Median (LQ, UQ)	52.5 (37.0, 62.0)	55.0 (49.0, 59.0)	53.0 (40.0, 58.5)	52.5 (40.0, 61.5)	47.0 (39.5, 57.0)	52.0 (43.0, 56.0)
Change from baseline	↓	↑	↑	↓	↑	↓

LQ, lower quartile; UQ, upper quartile.
↓, deterioration (i.e. positive ratings decreased); ↑, improvement (i.e. positive ratings increased).

TABLE 38 Staff empathy by experimental group (at follow-up)

| Empathy score (20–140) | Trial group (N = 87) | |
	CLECC (n = 53)	Control (n = 34)
Mean (SD)	112 (17)	113 (13)
Median	115	115
LQ, UQ	102, 125	104, 122
Minimum to maximum	57 to 133	79 to 135

LQ, lower quartile; UQ, upper quartile.

TABLE 39 Staff empathy (JSE) by ward (at follow-up)

| Empathy (20–140) | Trial group (*N* = 87) | | | | | |
	CLECC (*n* = 53)				Control (*n* = 34)	
Ward	A (*n* = 10)	B (*n* = 10)	D (*n* = 15)	E (*n* = 18)	C (*n* = 16)	F (*n* = 18)
Mean (SD)	108 (12)	113 (16)	114 (20)	112 (18)	113 (9)	113 (16)
Median (LQ, UQ)	109 (99, 117)	115 (99, 128)	120 (102, 129)	116 (105, 126)	114 (105, 120)	116 (104, 126)
Change from baseline	↓	↑	↓	↓	↓	↑

LQ, lower quartile; UQ, upper quartile.
↓, deterioration (i.e. reported empathy decreased); ↑, improvement (i.e. reported empathy increased).

Intracluster correlation

At ward level, the ICCs for QuIS, PEECH (subscales and total) and JSE were low (< 0.1). ICC was high at the observation session level for QuIS (*Table 40*).

Economic evaluation

The implications of the findings reported in this chapter are briefly explored in relation to an economic evaluation. None of the outcomes reported here seems appropriate for use in cost-effectiveness analysis. Such analysis typically takes the form of cost per unit of a particular outcome, usually the primary outcome. The primary outcome here, QuIS, is not expressed as a single number, which rules out any use of cost per QuIS unit. The same broadly applies to PEECH. Although nurse empathy is expressed as a single score, interpretation of a cost per unit of nurse empathy would be difficult. Given these difficulties, the most productive way forward would involve an impact inventory of the sort recommended in a recent authoritative review.[123] Besides being good practice for all economic evaluations, this would ensure that relevant data were presented on the range of costs and benefits involved.

TABLE 40 Values for ICC for primary outcome measures

Outcome measure	Level	ICC (95% CI)
QuIS	Ward	0.071 (0.000 to 0.164)
QuIS	Observation session	0.411 (0.264 to 0.558)
PEECH security	Ward	0.011 (0.000 to 0.050)
PEECH knowing	Ward	0.023 (0.000 to 0.073)
PEECH personal value	Ward	0.027 (0.000 to 0.079)
PEECH connection	Ward	0.011 (0.000 to 0.053)
PEECH total	Ward	0.027 (0.000 to 0.077)
JSE	Ward	0.000 (0.000 to 0.059)

Chapter summary

This chapter has focused on a comparison between intervention and control wards on the core outcomes at follow-up for the study, namely quality of staff–patient interaction, patient evaluation of emotional care and nursing self-reported empathy. The results suggest that CLECC may have a favourable effect in reducing negative interactions between staff and patients, and in reducing patients' experiences of lack of emotional connection with staff, but, as expected, we did not detect significant differences once other variables were accounted for. We found no evidence that nursing staff empathy improved because of CLECC, but these results in particular have to be viewed in the context of the low response rate to nursing surveys. Improving staff survey response rates in a future evaluation will improve confidence that bias is not skewing the results.

We reported between-ward differences, but differences at this level and on this small scale are as likely to be explained by random variation as by any other cause.

This chapter also reported on intracluster correlation for the three core outcome measures. All measures showed low variance at the ward level. However, there was a clear design effect apparent with QuIS at an observation session level, and this will need to be accounted for in the design of a future trial.

These results reported here will enable sample size calculation and inform outcome measure selection and use in a future trial, and are followed up further in *Chapter 11*.

Chapter 11 Discussion

This study aimed to assess the feasibility of implementing the CLECC intervention in acute hospital settings and to assess the feasibility of conducting a CRT with associated process and economic evaluations to measure and explain the effectiveness of CLECC.

The objectives were to:

1. determine the feasibility of implementing the CLECC intervention and sustaining the resulting work practices
2. inform the design of a definitive evaluation of the effectiveness of CLECC
3. inform the measurement of costs and benefits of CLECC in a definitive evaluation.

Findings showed that the CLECC intervention is feasible to implement in practice with medical and surgical nursing teams in acute care hospitals. We found strong evidence of good participation by nurses and HCAs, and staff reported benefits throughout CLECC's introductory period and beyond. Further impact and sustainability were limited by the focus on changing ward team behaviours rather than wider system restructuring. The pilot CRT proceeded as planned and randomisation was acceptable to teams. There was some evidence of contamination between wards in the same hospital but not between wards involved in the study. QuIS performed well, with a high recruitment rate and good inclusion of people with cognitive impairment. At follow-up there were higher total positive and lower total negative QuIS ratings for intervention wards than control wards. More control ward patients than intervention ward patients had lowest (i.e. more negative) scores on the PEECH connection subscale. These differences, although supported by the qualitative findings, are not significant. No significant differences in nursing empathy were observed, although response rates to the staff questionnaire were low. We also identified the costs associated with using CLECC and recommend that an impact inventory is used in any future study.

Each of the study objectives is addressed in more detail below.

Sections of this chapter are reproduced in Bridges et al.[121] This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See http://creativecommons.org/licenses/by/4.0/.

Feasibility of implementing and sustaining Creating Learning Environments for Compassionate Care[121]

The process evaluation aimed to identify and explain the extent to which the planned intervention was implemented into existing nursing practices and to draw conclusions about optimising the sustainability of future CLECC intervention. We found that, on all of the wards, many of the individual elements of CLECC were welcomed by staff and were possible to implement during the implementation period. We also found that sustaining this work beyond this time was difficult for some ward teams to achieve. Findings point to refinements needed to CLECC to improve the prospect of its impact and sustainability. The findings applied across all the ward contexts, in spite of some important differences between wards at the outset including specialty, staffing levels and ward leader experience.

Although CLECC had limited coherence for some staff, it was welcomed by teams and served as a broader stimulus to collective action. CLECC developed cultures in which reflection, learning, mutual support and innovation were legitimised within the work team, and in which expertise was seen to be distributed more widely between managers, RNs and HCAs. CLECC moved all the teams further along the continuum to becoming more expansive learning environments,[29] but implementation was mediated for all by the

context of working in an acute hospital environment. Staff highlighting what they valued about CLECC illuminated the stark realities of team working in such settings. The struggle to find the time to participate in CLECC reflects the pressure on staff to be constantly engaged in material patient care activities. Staff valued the cluster discussions because, otherwise, there was little opportunity to support each other's well-being. They appreciated the study days because they could get to know each other as people. They valued CLECC because otherwise they were lone workers, sharing working time and space with other team members, but not actually working as a team at all. The intensification of nursing work owing to rising patient complexity, in parallel with the application of increasingly stringent financial efficiency quotas, is well documented.[124–126] These findings paint a rich picture of the consequences for staff experiences at work and explain associations between hospital work team climate and staff well-being reported elsewhere.[26,27]

Staff at all levels were able to identify the benefits to patient care of ward staff engaging in CLECC activities, echoing other findings that the creation of unmanaged spaces for work team members to 'take shelter' provides the potential for valued learning and social support for difficult work with clients.[39,125] The findings confirm that intervening at the work team level can be successful, confirming an association conjectured from other research.[26–28,127] In spite of high workloads, CLECC empowered staff to reflect on local norms governing team practice, and on the relationships and resources that aligned with them, and to make some changes, confirming the original programme theory (set out earlier in *Chapter 1*, *Creating Learning Environments for Compassionate Care*) and indicating that collective agency can play a part in shaping relational capacity at the individual and work team levels.[125,128] However, we also found that implementation was uneven between teams, particularly over the longer term, reinforcing the value of paying attention to the sustainability of complex interventions beyond initial set-up.[129–131]

Factors outside the direct influence of the ward teams mediated the impact and sustainability of the intervention, in particular the institutional norms that legitimated staff's participation (or not) in CLECC activities and the interpretation by more senior figures (including PDNs, matrons and senior hospital managers) of what CLECC was and their role in supporting it. Although CLECC draws on the principles of democratic working, and we saw how HCAs in particular were enabled to take a lead in some CLECC activities, its longer-term success relied on cognitive participation from more senior members of the hierarchy. The authority that ward staff had to control how they spent their time, to innovate and to afford their own and colleagues' well-being some priority varied between teams and was dependent on the signals, or 'invitational qualities',[128] from these more senior figures as to what was legitimate or not. These findings that nurses do not control the conditions in which they work echo extensive research on the curtailment of professional autonomy in publicly funded health care, and the particular position that nursing as a profession occupies.[2,7,8,126,132–134] Matrons are the point at which organisational drivers, often business imperatives, must align with professional imperatives and the needs of frontline teams and their patients. The hybrid role and competing identities for nursing managers of this kind have been highlighted elsewhere[126,135,136] and it is unsurprising that we identified different approaches to managing this key role. Although the current CLECC activities relating to senior manager participation appear to have aided coherence, findings suggest that additional activities targeted at improving their cognitive participation may be needed. Contextual features that appeared to be relevant to CLECC's implementation journey included institutional norms regarding the legitimacy and nature of nursing work, staff learning and staff support; interpretation of key stakeholder roles (including nursing managers and PDNs); and ward-level characteristics, such as staffing levels in relation to patient workload, and stability of workforce and team leadership over time.

Our findings indicate that a higher and more sustained impact for interventions, such as CLECC, may be possible only through more substantial restructuring that reshapes the conditions in which people are able to act.[53,54] We support Parker's[39] assertion that caregiving organisations need to be designed to enable caregivers to access functional work teams within which they can interpret their experiences, and we have identified a number of concrete but modifiable barriers that merit attention in such design, including lack of time and institutional rules that undermine the value of staff-to-staff social support. They also include more clearly defining the role of nursing managers in signalling the legitimacy of staff providing each other with emotional support, supporting nursing teams to meet and learn together. Future versions of the CLECC intervention

should include new activities to engage nursing managers in the implementation period, involve them in the learning activities and create opportunities for them to engage and reflect with frontline staff.

These findings reflect the fact that relational work in caregiving organisations depends not just on individual caregiver agency but also on whether or not this work is adequately supported by features of the wider system. Relational capacity may thus be regarded not only as a property of individual practitioners, but also as a modifiable and situated property of work teams. The limitation that institutional norms of legitimate nursing work placed on staff finding time to meet together raises the prospect of lack of relational capacity at a wider system level, and suggests that wider restructuring beyond middle manager roles may in fact be needed to effect substantial and sustained change.

These findings offer possibilities for actively restructuring work team conditions to enable the relational aspects of caring and working. Complex interventions, such as CLECC, can be theorised as 'time-limited series of events, new activity settings and technologies' in systems, the focus of the intervention being to generate new structures of interaction and new shared meanings.[53,54] Adopting this perspective means that the focus of change efforts is not just on the behaviour of individual staff, but also on restructuring of relationships, norms and resources in the wider system that may play their part in the success or otherwise of the intervention. The findings point towards a number of organisational conditions in which high-quality care is most likely and in which interventions of this kind are most likely to succeed, and these are set out in *Chapter 12*. Further research across a wider range of organisational settings will enable these emerging theories to be refined to enhance their transferability. This further research will also enable us to more closely describe the links between context, implementation processes and outcomes associated with implementation of the CLECC intervention.

Informing future Creating Learning Environments for Compassionate Care evaluation

Our systematic review[16] found that any of the interventions we investigated might be deemed worthy of further investigation based on their positive outcomes but that none could be recommended for routine implementation, given the lack of theoretical basis and description for many interventions, the pervasive positive bias that is associated with weak study designs, and the lack of evidence for impact on patient outcomes in most studies. Adherence to recognised and emerging standards for developing and evaluating complex interventions, such as the UK MRC framework,[88] and fuller reporting of interventions and outcomes would address many of the issues noted in our review. We concluded that many researchers in this field have been unable or unwilling to use experimental designs within the context of mixed-methods approaches to evaluation.

The findings from this study indicate that the use of an experimental design to evaluate the effectiveness of compassionate care interventions within the context of a mixed-methods study is feasible, as is a focus on outcomes that are patient based. Ward teams were successfully randomised to intervention or control, and staff were amenable to the prospect of randomisation to either experimental condition. All wards recruited remained in the study throughout data collection and all clusters randomised to the intervention went on to receive it. Blinding of patients and visitors to ward allocation appeared successful, although strategies to blind researchers gathering data need further development in a future trial. Although we identified some differences between individual clusters and individual participants in the trial, none was sufficiently substantial to raise concerns of baseline imbalances between intervention and control conditions.[137,138]

We found evidence of pathways through which the CLECC intervention had the potential to influence practice in other wards in both of the participating NHS trusts. This may be a sign of CLECC's success in transforming and embedding in the wider system, but also indicated that future studies of effectiveness should not run intervention and control conditions in the same organisation over the same time period. This will be an important consideration in designing any future definitive experiments.[139,140]

Older patients with a range of cognitive abilities are a traditionally hard-to-reach group in research, especially when they are unwell and in hospital. Even though they are more prone to negative experiences of hospital care, they are often excluded from research.[6,141,142] Although cognitive deficits may limit some people's ability to share their experiences, our findings indicate that devising recruitment and data collection methods that maximise support and inclusion can be successful. It is estimated that at any one time, up to 25% of beds in acute hospitals are occupied by people with dementia, with the figure likely to be higher on specialist older people's care wards.[143,144] Overall, 25% of patients observed in this study had evidence of cognitive impairment, suggesting that our sample was representative of the wider hospital population. Of all the patient questionnaires returned, 12% were completed by patients with cognitive impairment and so we were less successful in achieving representativeness here.

Our findings that structured non-participant observation appears to be the most promising method to describe the experiences of older people with cognitive impairment in the general hospital setting echo those of Goldberg *et al.*[141] Participating in an observation does not require any particular state of health, abilities or performance from the patient in question, whereas, for instance, answering questions in an interview about one's care experiences requires as a minimum orientation to place, language skills and attention.[142] We did find that recruitment of people with cognitive impairment to the study took more time, as did completing questionnaires with them, and this has resource implications for future research with this patient group. In relation to patient involvement in general, this method overcomes the reluctance of patients to evaluate care while it is ongoing that was noted in other studies.[64]

Our development of the QuIS tool for use in acute settings worked well and our earlier work confirmed an association between patient-reported experience and observed staff–patient interactions.[99] However, our earlier work did not include people with a cognitive impairment and validation of QuIS ratings with this patient group may be a necessary next step in the tool's development. Although a clear design effect was apparent with QuIS at the observation session level that will need to be accounted for in the design of a future trial, benefits to the use of this tool were notable.[139] Acceptability of the tool to patients and staff was high, and reliability between observers was acceptable. We did not find any evidence that staff changed their behaviour as a result of being observed but this possibility cannot be eliminated. However, the findings across both assessment periods that a proportion of interactions observed were negative indicate that, even if staff planned to give consistently good care while being observed, they were not successful. Although we cannot eliminate the possibility of observer effects, the effect in a future trial will be present on all wards in all conditions and so differences between wards can still be attributed to the intervention. We know from other work that the quality of interactions with staff is very important to older people and shapes their experiences in hospital settings and so its successful measurement is a good indicator of compassionate care.[6] Our findings from this study confirm that observation-based measures are more inclusive of patient groups vulnerable to negative experiences in hospital. Overall, they support the selection of the quality of staff–patient interactions, measured by an observational tool (such as QuIS), as a candidate primary outcome in a future trial.

The response rate to nursing questionnaires (36%) was very low, particularly at follow-up stage on the intervention wards, perhaps owing to research fatigue. There was a perception that questionnaires were lengthy to complete and that staff were too busy. The low response rate means reduced certainty that the responses to the questionnaire represent the views of the nursing staff as a whole, although the findings of burnout and low empathy within the staff groups on every ward were important. The response rates are typical for surveys of this kind with this population. For instance, in a European study[145] of nurse staffing levels, an estimated 39% of 2917 registered nurses working on NHS medical and surgical units in England completed a questionnaire similar to the one used in this study. Deploying a shorter questionnaire and negotiating time for their completion with managers may enhance response rates in a future study.

Small but non-significant differences between experimental groups at follow-up in quality of interaction and patient evaluation of emotional connection with staff are promising findings, particularly in the context of qualitative findings that indicate benefits to patient care perceived by staff. Data were gathered

≥ 4 months after the end of the implementation period, indicating that if there is an effect that can be attributed to CLECC, it is sustainable beyond the period in which CLECC was being actively facilitated by PDNs. We found no evidence that nursing staff empathy may be improved because of CLECC but these results in particular have to be viewed in the context of the low response rate to nursing surveys.

These findings are an important contribution to a field in which the use of experiments is relatively rare, and the results reported here will inform study design, sample size calculation and outcome measure selection and use in a future trial.

Our qualitative findings illuminate the importance of context in shaping implementation and outcomes and strongly indicate that future measurement of benefit should be part of a mixed-methods evaluation.

Informing measurement of Creating Learning Environments for Compassionate Care costs and benefits

Our findings have established the feasibility of estimating the cost of a CLECC-type intervention. Intervention costs were calculated as training costs (PDN time and staff time attending study day) and ongoing implementation costs (cost of staff engaging in CLECC activities on the ward). The extent to which training costs would be an additional cost to existing training would need to clarified in a future study, including any further training costs incurred from shifting team membership and the need for refresher sessions for existing staff. Through qualitative interviews with staff, we were able to establish that, aside from initial CLECC training costs, the implementation of concrete CLECC activities by ward teams was not perceived to require additional resource use. However, given the size of our sample and its qualitative nature, we consider that staff time spent on CLECC activities would still need to be recorded and costed in any larger trial. Data should also be gathered on the amount of training delivered through a register of attendance at classroom training, action-learning sessions and cluster sessions. The findings also indicate candidate activities associated with supporting CLECC that merit attention in a future evaluation.

A patient-level outcome measure, such as EQ-5D-5L, was shown not to be feasible, mainly because different patients with different ailments and severity were involved at baseline and at follow-up. Use of a patient-specific outcome requires either that the same patients in both intervention and control groups are measured over time or that adequate adjustments for the casemix of severity can be made. Neither seems likely. Neither QuIS nor any of the secondary outcomes were promising candidates for cost-effectiveness analysis, owing mainly to lack of single summary measures. Although QuIS is used as an audit tool in Scotland and has been applied to some services in the English NHS, it does not at present lend itself to use in cost-effectiveness analyses. Provision of data on QuIS and other ward-level scores may be of use in other evaluations or in developments towards a summary measure.

Consequently, we consider that any proposed economic evaluation of a CLECC-type intervention should comprise two elements. The first is an impact inventory, including comprehensive data on the costs and benefits of the interventions, distinguishing training and implementation. The second element proposed is a series of cost-effectiveness analyses linking cost to each of the primary and secondary outcomes. This approach has value for the evaluation of novel complex interventions with uncertain resource implications.

In summary, an impact inventory would provide a comprehensive listing of the resources, cost and benefits of CLECC with a focus on those to do with providing the interventions, but set within a wider context that includes effects on staff and on patients. Cost per change in each of the primary and secondary outcomes could also be estimated and compared with other later studies.

Strengths and limitations

This detailed and thorough mixed-methods study makes an important contribution to the evidence base on the design and evaluation of compassionate care interventions. Based on data gathered from a range of English NHS ward contexts, the findings are relevant to those seeking to influence and evaluate nursing practice in acute hospital settings in similar contexts. The qualitative findings indicate that staff welcomed CLECC and perceived positive benefits to their own well-being, to improved team working and to more compassionate patient care. As detailed earlier, we have theorised that the impact and sustainability of CLECC can be enhanced by attention to wider system restructuring, and refining these theories using data gathered in a wider range of organisational contexts will be an important next step. In particular, closer attention to defining the contexts in which CLECC is most likely to succeed will be a necessary focus in the next stage of this programme of research. We found variations in intervention fidelity attributable to a variety of contextual features. These factors merit refinement of elements of the intervention for future use, but also deserve continuing investigation in future studies because there may be other important features, identifiable only through a larger-scale study of context. Continuing qualitative investigation, in the form of interviews with frontline staff and service managers, and more detailed observation of CLECC implementation in practice, will enhance an understanding of the influence of context on implementation. It will also provide explanations for findings regarding CLECC's efficacy.

The pilot RCT findings lend some support to the nurses' views as regards the benefits to patient care, but larger-scale evaluation is needed before definitive claims are merited. Insufficient information was available at the outset of this study to enable power calculations that informed sample size and so there is no certainty that any apparent positive effects are not produced by chance alone, rather than the impact of the CLECC intervention. Potential issues of lack of researcher blinding to experimental allocation and contamination pathways between intervention and control wards also mean that caution should be applied in drawing conclusions of efficacy from the findings presented.

We have generated useful findings about the performance of a range of outcome measures in relation to compassionate care and have demonstrated the feasibility of using patient-based outcome measures in this field. Our findings indicate the strengths of observation-based evaluations of care delivered, but further research to assess the validity of these evaluations in relation to the experiences of people with cognitive impairment is warranted.

The research to date has focused on nursing teams in hospital settings and no claims are made about the generalisability of these findings to other types of team or other settings. We propose that, with some modifications to account for different contexts, the CLECC intervention may be of value to other teams in other settings but research of the kind reported here will be an important foundation to its use and evaluation in new contexts.

Our findings indicate that further evaluation is merited and point the way to how such future evaluation should be designed and carried out. *Chapter 12* draws together the conclusions from the study and sets out implications for health care and recommendations for this future research.

Chapter 12 Conclusions

Our conclusions focus on two main areas: the implications of the findings for health care and recommendations for future research.

Implications for health care

Our in-depth analysis of the process of implementing a complex intervention targeted at compassionate care raises questions about the extent to which interventions of this kind should in fact target and seek to influence and restructure relationships, norms and resources in the wider system. They suggest that health-care leaders who interpret their role as mobilising structural capacity to support the relational work of frontline staff may well improve the relational capacity of teams and their individual members. We have defined elements of this mobilisation in relation to the CLECC intervention, and our planned enhancements to the original intervention are set out in *Box 2*. The enhancements clarify the role of leaders outside the ward team in supporting ward teams to implement CLECC by engaging leaders in the programme, involving them in the learning activities and creating opportunities for them to engage and reflect with frontline staff. We also propose tying in CLECC with wider staff education strategies in the organisation to help its wider integration, thus enabling the possibility that its goals become embedded and reflected across educational provision.

BOX 2 Recommendations for future CLECC intervention modifications

---

Implementation period

This can be reduced to 3 months.

Person specification and support for Creating Learning Environments for Compassionate Care facilitator

The individual PDN leading the implementation of CLECC should have solid educational experience and ideally be an existing member of the organisation's education team. They should also have regular supervision in relation to their CLECC role to enable support to be counter-cultural and to keep this up.

Appointment of Creating Learning Environments for Compassionate Care champions

Each team should appoint two CLECC champions with the authority to initiate and lead CLECC activities and act as a resource about CLECC for colleagues. These champions do not need to be senior members of the team but should have the ward leader's support. They should have training for their role and access to regular mentoring/supervision. They should change annually and take the lead in training the next champions.

Induction/information for staff new to Creating Learning Environments for Compassionate Care

A written summary of CLECC should be given to all team members.

New staff members appointed to the ward team should be given the opportunity to learn about CLECC and what it means. This induction should include being given a written summary about CLECC and hearing about how it works on the ward from the CLECC champion.

BOX 2 Recommendations for future CLECC intervention modifications (*continued*)

New ward leaders need a more in-depth briefing and ideally some mentoring over the first couple of months in the post, including checking that they have access to supervision/action-learning, and regular meetings with the matron.

Matrons and senior nursing managers

Matrons and senior nursing managers, including the director of nursing, need to learn about CLECC and be given specific responsibilities in relation to CLECC early on in the implementation period and at regular intervals throughout. They should be given a written summary about CLECC. Their role could include ward visits explicitly to learn about CLECC, participating in study days and clusters (and care on the ward when this is possible), helping to develop the sustainability plan, actively encouraging staff to innovate, and should have generally visible involvement in and support of CLECC.

Educational strategy tie-in

The CLECC intervention should be tied in with the organisation's education strategy and be seen as part of the educational offer to staff. One possibility is that the CLECC intervention is delivered by the organisation's educators (practice educators in the NHS). Other practice educators working with the team implementing CLECC should be given written information about CLECC and the opportunity to discuss their role in supporting it.

Action-learning sets

Sets should be facilitated by someone with training and preferably experience in action-learning, and preferably someone who is linked in with the organisation's education infrastructure. This would ideally be the PDN leading CLECC but could be someone else not involved in directly managing the ward team. Action-learning facilitation skills are more important than knowledge of CLECC.

Study days

A more comprehensive outline to guide the study-day programme is needed. This should include structure of the day, learning activities, materials and educational philosophy. The PDNs will need to be educated about how the study day fits in with the CLECC programme.

Sustainability plan

Teams should produce a sustainability plan at the end of the 3-month implementation period, which sets out how the team will take CLECC forward, measurable goals, and the identification of resources and support required to implement the plan. A structured outline is needed for the plan to prompt reflection (e.g. 'How could action-learning continue?').

Plans should include the development of an innovation plan to guide the development and implementation of innovations from an idea by an individual through to a change being realised.

Boost

Opportunities should be created for teams to revisit the CLECC principles, practices and sustainability plan after the end of the initial implementation period. It provides teams with activities 3–6 months after the initial implementation period that are designed to promote fidelity to CLECC values, refresh people's knowledge about CLECC, motivate them to continue and enable them to reflect on progress to date and to strengthen sustainability plans.

Proposed enhancements to the CLECC implementation period include appointing PDNs with solid experience as educators, and a more detailed specification for study days, to ensure consistency between sites in adhering to the CLECC principles. The appointment of team members as CLECC champions and inviting teams to develop sustainability plans will encourage staff to identify concrete contextually specific activities that will support CLECC going forward, along with an articulation of the roles, responsibilities and resources required to sustain CLECC.

Other proposed enhancements to CLECC focus on the period beyond the initial implementation period. Sustainability plans will encourage staff to agree on explicit expectations for discussing and developing an understanding of CLECC principles on an ongoing basis, and may embed this culture further. New team members will need induction to understand what CLECC is and what their role is in relation to CLECC. This particularly applies to ward leaders joining a ward where CLECC is in place, and careful attention and mentoring will be needed to support them as they develop their role.

Recommendations for research

The complexity of the intervention and the clear relationship between context and impact reflected in the findings, in addition to the continuing need to establish the efficacy of interventions of this kind, require a mixed-methods approach to future evaluation within the context of a programme of research to:

1. identify the organisational contexts in which optimal impact and sustainability of the CLECC intervention are most likely
2. further establish the feasibility and validity of the QuIS tool in relation to the experiences of patients in acute care settings
3. evaluate the effectiveness and cost-effectiveness of the CLECC intervention relative to usual care.

A programme of research would enable each of these objectives to be addressed. First, identifying the contexts in which optimal impact and sustainability are most likely is an important goal for the next stage of this research. Evaluating the processes of implementation in relation to contextual features in a wider range of acute care contexts will enable us to generate and test a typology of organisational types (at hospital and ward levels) that specifies their receptiveness to interventions of this kind. The nature of this type of evaluation is likely to be largely qualitative, drawing on observations of practice and interviews with key stakeholders over time to describe variations and identify relationships between implementation processes and context, and the resultant impact and sustainability. Qualitative and quantitative contextual data gathered from the feasibility study reported here (on, for instance, ward leadership, staff perceptions of care and staff well-being) could be added to equivalent data gathered from other organisations in a future study and combined to inform the development of the typology and the identification of contextual features relevant to implementation processes, impact and sustainability. The opportunity to investigate the wider dissemination and embedding of ideas and practices originating from the intervention but spreading beyond the target team to the wider system can also be exploited by qualitative exploration within clinical departments and within the wider health-care system.

Meeting this first objective is achievable through the study of context, implementation, impact and sustainability of the CLECC intervention in acute hospital nursing teams sampled to ensure heterogeneity. The findings regarding the influence of differences in contextual features at the institutional and team levels can be used to inform this sampling, namely institutional norms regarding the legitimacy and nature of nursing work, staff learning and staff support; interpretation of key stakeholder roles (including nursing managers and PDNs); and ward-level characteristics, such as staffing levels in relation to patient workload and stability of workforce and team leadership over time. Conducting this study over a 2-year period will enable us to capture the impact of variations over time and also to build a picture of longer-term sustainability.

Although our findings on the use of the QuIS tool are promising in relation to its inclusivity of hard-to-reach patient groups, further work is merited to inform its use as a primary outcome measure in future experiments in acute settings, in particular on its validity in relation to patient experience. Our work indicates that there is an association between QuIS ratings and patient experiences in acute hospital settings, but this relationship needs testing on a wider scale, in particular with patients who have a cognitive impairment. The second objective of the proposed programme of work can be met by a study that evaluates staff–patient interactions through QuIS rating, as used here, but also through patient ratings of the same interactions. These sets of ratings can then be compared. Further work will be needed to establish a means by which patients with cognitive impairment can rate interactions. If the proportion of negative interactions is the primary outcome measure in a future study, understanding which interactions are rated by observers (and, when possible, patients) as negative, and why, is an important next step, as is working with patient representatives to establish their views on the size of a meaningful reduction in negative interactions. Further study could also be used to develop more effective procedures to blind observers from experimental allocation in advance of an experimental study. In addition, the high intracluster correlation that we found at an observation session level merits the exploration of the cause of this variance and the feasibility of different approaches to data collection that might reduce its impact, for instance, shorter observation sessions. This further evaluation and testing of QuIS across these parameters would be a valuable foundation to its further use as an outcome measure in acute settings.

Drawing on our findings about the feasibility of experimental approaches to evaluating compassionate care interventions, our third objective for a programme of work focuses on the delivery of a definitive multicentre trial to establish CLECC's effectiveness and cost-effectiveness. Our findings indicate that, given variations in the implementation journey over time, outcomes should be captured over a long period of time, at least 12 months, to ensure that sustainability is tested. Evaluation in a range of organisational contexts will improve the generalisability of findings and so a multicentre trial is merited. If QuIS is selected as the primary outcome measure, the high ICC at the observation session level indicates that it would be more efficient to conduct more observation sessions of shorter duration. The length of time taken to recruit a patient group with more complex needs militates against conducting sessions of < 1 hour. We can explore this further in the QuIS study proposed above but, for the calculations that follow, we assume that observation sessions are 1 hour in length. Detecting a 50% reduction in the rate (odds) of a negative interaction (i.e. a reduction from 10% of all interactions to 5%) at 90% power would require observation of 582 interactions per group, that is 1164 interactions in a parallel group trial. Allowance for clustering is achieved through use of a multiplicative factor $[1 + (n \text{ cluster} - 1) \times \text{ICC}]$. Patients in our feasibility study had an average of six interactions with staff per hour. If the cluster (observation session) is six (interactions), that is, the observation sessions are 1-hour long, the factor is $[1 + (6 - 1) \times 0.411] = 3.055$. We would therefore need $1164 \times 3.055 = 3556$ observed interactions to detect a difference, that is, 593 observation periods of 1 hour each in total, rounded up to 300 per group.

If individual wards are observed for 20 hours each at each assessment period (the amount used in the feasibility study), 30 wards would need to be engaged in a trial (15 in each group). Our estimates of the costs and feasibility of implementing the CLECC intervention in each NHS organisation and the work required to set up and oversee the study in each organisation indicate that five wards each in six different NHS hospitals would enable this level of data collection to be achieved. These calculations assume that one patient is observed at a time but take no account of clustering within wards. In reality we were generally able to observe more than one patient at a time, which would provide additional data that would be more than sufficient to compensate for the relatively small increase in sample size required because of clustering at the ward level, given the low ICC at this level.

Our findings indicate that measuring patient views on care and staff self-rated empathy are useful and feasible as secondary outcome measures, although careful attention would need to be paid to maximise staff survey response rates. The study design would need to ensure that intervention and control conditions do not run in the same organisation at the same time. A waiting list control may be helpful

here, with all study wards eventually receiving the intervention, but this would double the intervention costs and funding constraints are very likely to reduce the feasibility of this option.

As outlined above, cost-effectiveness can be evaluated through an impact inventory, including comprehensive data on the costs and benefits of the interventions, distinguishing training and implementation. We also propose a series of exploratory cost-effectiveness analyses linking cost to each of the primary and secondary outcomes.

To summarise, our findings indicate that further intervention development and evaluation work of the CLECC intervention through a programme of research is now merited.

Acknowledgements

Contributors to the study

Project management team

- Professor Jackie Bridges.
- Professor Peter Griffiths.
- Dr Ruth M Pickering.
- Professor Avan Aihie Sayer.
- Professor Alison Fuller.
- Dr Lily Yao.
- Dr Greta Westwood.
- Mrs Rosemary Chable.
- Ms Emma Munro.
- Professor James Raftery.

Research team

- Ms Ines Mesa-Eguiagaray.
- Dr Lisa Gould.
- Ms Hannah Barker.
- Ms Paula Libberton.
- Dr Wendy Wigley.
- Dr Shihua Zhu.
- Dr Christopher McLean.
- Ms Lisa Shipway.

Systematic review team

- Dr Karin Blomberg, Örebro University.
- Professor Jackie Bridges, University of Southampton.
- Professor Peter Griffiths, University of Southampton.
- Professor Carl May, University of Southampton.
- Professor Yvonne Wengström, Karolinska Institutet.

Other Quality of Interactions Schedule observers/questionnaire administrators

- Emily Oliver.
- Chloe Dyer.
- Naomi Gallant.
- Robyn Gentle.
- Ruth Lutz Tracey.
- Lyndsey Rickman.
- Annabel Rule.
- Iola Williams.

Study steering committee

- Professor Jill Maben, King's College London (chairperson).
- Mrs Janet Gollop, study PPI Reference Group.
- Mrs Brenda Hadfield, study PPI Reference Group.
- Dr Jennifer Wilkinson, Newcastle Clinical Trials Unit.
- Dr Caroline Nicholson, King's College London.
- Dr Sarah Goldberg, University of Nottingham.

Study advisory group

- Dr Mary Flatley, St Joseph's Hospice.
- Ms Bev Fitzsimons, Point of Care Foundation.
- Ms Lesley Carter, Age UK.
- Ms Sarah Balchin, Portsmouth Hospitals NHS Trust.
- Professor Carl May, University of Southampton.

Patient and public involvement reference group

- Janet Gollop.
- Mary Carnegie.
- Benita Wilkinson.
- Pam Holloway.
- Debra Ridley.
- Brenda Hadfield

Administrator

- Sue Warren.

Contributions of authors

Jackie Bridges (Professor of Older People's Care) was chief investigator, leading the design and implementation of the study, the acquisition, analysis and interpretation of data, and the preparation and writing up of results.

Ruth M Pickering (Associate Professor of Medical Statistics) contributed to study design and oversaw the analysis of quantitative data and the presentation of quantitative results.

Hannah Barker (Senior Research Assistant, Ageing and Dementia) gathered qualitative and quantitative data, trained other research team members, prepared the quantitative data for analysis and contributed to analysis of qualitative data.

Rosemary Chable (Associate Director of Nursing) was based at a participating NHS site and oversaw intervention implementation and staff participation in the study.

Alison Fuller (Professor of Vocational Education and Work) contributed to the study design, the intervention implementation and the qualitative data analysis.

Lisa Gould (Research Fellow, Ageing and Dementia) managed the study day to day, gathered qualitative and quantitative data, trained other research team members, contributed to the analysis of qualitative data and helped to prepare the results for publication.

Associated publications and presentations

Barker H, Griffiths P, Gould L, Mesa-Eguiagaray I, Pickering R, Bridges J. Quantity and quality of interaction between staff and older patients in UK hospital wards: a descriptive study. *Int J Nurs Stud* 2016;**62**:100–7.

Barker H, Griffiths P, Mesa-Eguiagaray I, Gould L, Bridges J. *Quantity and Quality of Interaction Between Staff and Older Patients in UK Hospital Wards: A Descriptive Study*. Paper presented at the Royal College of Nursing International Research Conference, Edinburgh, UK, 7 April 2016.

Blomberg K, Griffiths P, Wengstrom Y, May CR, Bridges J. Interventions for compassionate nursing care: a systematic review. *Int J Nurs Stud* 2016;**62**:137–55.

Bridges J, May CR, Griffiths P, Fuller A, Wigley W, Gould L, *et al.* Optimising impact and sustainability: a qualitative process evaluation of a complex intervention targeted at compassionate care. *BMJ Qual Saf* 2017;**26**:970–7. http://qualitysafety.bmj.com/content/early/2017/09/15/bmjqs-2017-006702.full

Bridges J. *Creating Learning Environments for Compassionate Care (CLECC): Developing and Evaluating the Feasibility of a Complex Intervention*. Paper presented at the International Learning Collaborative on Fundamentals of Care, Oxford, UK, 13 June 2016.

Bridges J. *Culture of Care: Innovation in Practice*. Paper presented at the Older Person's Nurse Fellowship Conference, London, UK, 14 June 2016.

Bridges J. *How Clinical Academic, Research Nurses and Ward Nurses Contributed to Make the CLECC Study Work*. Paper presented at It's okay to ask about . . . research nursing, Southampton, UK, 20 May 2016.

Bridges J. *Intervention Studies in Relational Care in Acute Settings: Content, Context, and Consequences*. Symposium presented at the Royal College of Nursing International Research Conference, Edinburgh, UK, 7 April 2016.

Bridges J, Fuller A. Creating Learning Environments for Compassionate Care (CLECC): a programme to promote compassionate care by health and social care teams. *Int J Older People Nurs* 2015;**10**:48–58.

Bridges J, Griffiths P, Barker H, Pickering RM, Mesa-Eguiagaray I, Libberton P, *et al. Creating Learning Environments for Compassionate Care (CLECC): an Acute Care Feasibility Study*. Poster presented at European Nursing Congress, Rotterdam, the Netherlands, 5–7 October 2016.

Bridges J, Libberton P, Barker H, Gould L, Wigley W, Griffiths P. *Creating Learning Environments for Compassionate Care (CLECC): Developing and Evaluating the Feasibility of a Complex Intervention*. Paper presented at the Royal College of Nursing International Research Conference, Edinburgh, UK, 7 April 2016.

Gould L, Griffiths P, Barker H, Mesa-Eguiagaray I, Bridges J. *Creating Learning Environments for Compassionate Care (CLECC): Feasibility of Evaluating Impact on Patient Care*. Paper presented at the Royal College of Nursing International Research Conference, Edinburgh, UK, 7 April 2016.

Libberton P, Bridges J. *Older Adults in General Hospitals Want their Mental Health Needs Addressed as Part of Nursing Care: Why is this so Challenging*? Paper presented at the International Network for Psychiatric Nursing Research (NPNR) Conference, Nottingham, UK, 15 September 2016.

Libberton P, Griffiths P, Gould L, Barker H, Mesa-Eguiagaray I, Bridges J. *Relational Capacity of Nursing Teams: Exploring the Relationship Between Team Context, Relational Capacity and Caring Practices*. Paper presented at the Royal College of Nursing International Research Conference, Edinburgh, UK, 7 April 2016.

McLean C, Griffiths P, Mesa-Eguiagaray I, Pickering RM, Bridges J. Reliability, feasibility, and validity of the quality of interactions schedule (QuIS) in acute hospital care: an observational study. *BMC Health Serv Res* 2017;**17**:380. https://doi.org/10.1186/s12913-12017-12312-12912

Mesa Eguiagaray I, Böhning D, McLean C, Griffiths P, Bridges J, Pickering RM. Inter-rater reliability of the QuIS as an assessment of the quality of staff-patient interactions. *BMC Med Res Methodol* 2016;**16**:171. https://doi.org/10.1186/s12874-016-0266-4

Mesa Eguiagaray I, Pickering R, McLean C, Griffiths P, Bridges J. *An Application of Weighted Kappa in Nursing Research*. Paper presented at the International Biometric Conference, Victoria, BC, Canada, 12 July 2016.

Mesa Eguiagaray I, Oliver E, Griffiths P, Bridges J, Gould L, Pickering R. A multilevel logistic model of poor quality staff-inpatient interactions accounting for variability across patients and observation periods. Poster presented at the International Society for Clinical Biostatistics Annual Conference, Vigo, Spain, 12 July 2017.

Data-sharing statement

Data can be requested from the corresponding author.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: https://understandingpatientdata.org.uk/data-citation.

References

1. Francis R. *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. London: The Stationery Office; 2013. URL: http://webarchive.nationalarchives.gov.uk/20150407084003/ http://www.midstaffspublicinquiry.com/report (accessed 18 February 2017).

2. Francis R. *Independent Inquiry into Care Provided by Mid Staffordshire NHS Foundation Trust January 2005 – March 2009*. London: The Stationery Office; 2010. URL: www.gov.uk/ government/uploads/system/uploads/attachment_data/file/279109/0375_i.pdf (accessed 18 February 2017).

3. National Institute for Health Research. *National Institute for Health Research Health Services and Delivery Research Programme Commissioning Brief 13/07 – After Francis: Research to Strengthen Organisational Capacity to Deliver Compassionate Care in the NHS*. Southampton: National Institute for Health Research; 2013. URL: https://njl-admin.nihr.ac.uk/document/download/2010144 (accessed 17 July 2018).

4. Department of Health and Social Care. *Hard Truths: The Journey to Putting Patients First*. London: Department of Health and Social Care; 2014. URL: www.gov.uk/government/publications/mid-staffordshire-nhs-ft-public-inquiry-government-response (accessed 18 February 2017).

5. Care Quality Commission. *Dignity and Nutrition Inspection Programme: National Overview*. Newcastle upon Tyne: Care Quality Commission; 2011. URL: www.cqc.org.uk/sites/default/files/ documents/20111007_dignity_and_nutrition_inspection_report_final_update.pdf (accessed 18 February 2017).

6. Bridges J, Flatley M, Meyer J. Older people's and relatives' experiences in acute care settings: systematic review and synthesis of qualitative studies. *Int J Nurs Stud* 2010;**47**:89–107. https://doi.org/10.1016/j.ijnurstu.2009.09.009

7. Kagan SH. Compassion. *Geriatr Nurs* 2014;**35**:69–70. https://doi.org/10.1016/j.gerinurse. 2013.11.006

8. Care Quality Commission. *The State of Health Care and Adult Social Care in England 2015/16*. 2016. URL: www.gov.uk/government/publications/health-care-and-adult-social-care-in-england-2015-to-2016 (accessed 18 February 2017).

9. National Institute for Health and Care Excellence (NICE). *Safe Staffing for Nursing in Adult Inpatient Wards in Acute Hospitals*. NICE; 2014. URL: www.nice.org.uk/guidance/sg1/resources/ safe-staffing-for-nursing-in-adult-inpatient-wards-in-acute-hospitals-61918998469 (accessed 7 July 2017).

10. Health Education England. *Workforce Plan for England: Proposed Education and Training Commissions for 2016/17*. 2016. URL: http://qna.files.parliament.uk/qna-attachments/600490/ original/Workforce%20Plan%20for%20England%202016-17.pdf (accessed 18 February 2017).

11. Smith C, Baltruks D. *The Nursing Journey: Recruitment and Retention*. Sedlescombe: Good Governance Institute; 2015. URL: www.good-governance.org.uk/services/the-nursing-journey-recruitment-and-retention/ (accessed 24 August 2018).

12. Ackroyd S, Bolton S. It is not Taylorism: mechanisms of work intensification in the provision of gynaecological services in a NHS hospital. *Work Employ Soc* 1999;**13**:369–87. https://doi.org/ 10.1177/09500179922117980

13. Ackroyd S, Kirkpatrick I, Walker RM. Public management reform in the UK and its consequences for professional organization: a comparative analysis. *Public Admin* 2007;**85**:9–26. https://doi.org/ 10.1111/j.1467-9299.2007.00631.x

14. Bridges J, Hughes J, Farrington N, Richardson A. Cancer treatment decision-making processes for older patients with complex needs: a qualitative study. *BMJ Open* 2015;**5**:e009674. https://doi.org/10.1136/bmjopen-2015-009674

15. Bridges J, Nicholson C, Maben J, Pope C, Flatley M, Wilkinson C, *et al.* Capacity for care: meta-ethnography of acute care nurses' experiences of the nurse-patient relationship. *J Adv Nurs* 2013;**69**:760–72. https://doi.org/10.1111/jan.12050

16. Blomberg K, Griffiths P, Wengström Y, May C, Bridges J. Interventions for compassionate nursing care: a systematic review. *Int J Nurs Stud* 2016;**62**:137–55. https://doi.org/10.1016/j.ijnurstu.2016.07.009

17. Dewar B, Adamson E, Smith S, Surfleet J, King L. Clarifying misconceptions about compassionate care. *J Adv Nurs* 2014;**70**:1738–47. https://doi.org/10.1111/jan.12322

18. Maben J, Cornwell J, Sweeney K. In praise of compassion. *J Res Nurs* 2010;**15**:9–13. https://doi.org/10.1177/1744987109353689

19. Schantz ML. Compassion: a concept analysis. *Nurs Forum* 2007;**42**:48–55. https://doi.org/10.1111/j.1744-6198.2007.00067.x

20. Chochinov HM. Dignity and the essence of medicine: the A, B, C, and D of dignity conserving care. *BMJ* 2007;**335**:184–7. https://doi.org/10.1136/bmj.39244.650926.47

21. Cameron RA, Mazer BL, DeLuca JM, Mohile SG, Epstein RM. In search of compassion: a new taxonomy of compassionate physician behaviours. *Health Expect* 2015;**18**:1672–85. https://doi.org/10.1111/hex.12160

22. Dewar B, Cook F. Developing compassion through a relationship centred appreciative leadership programme. *Nurse Educ Today* 2014;**34**:1258–64. https://doi.org/10.1016/j.nedt.2013.12.012

23. Von Dietze E, Orb A. Compassionate care: a moral dimension of nursing. *Nurs Inq* 2000;**7**:166–74. https://doi.org/10.1046/j.1440-1800.2000.00065.x

24. Hartrick G. Relational capacity: the foundation for interpersonal nursing practice. *J Adv Nurs* 1997;**26**:523–8. https://doi.org/10.1046/j.1365-2648.1997.t01-12-00999.x

25. May C. Individual care? Power and subjectivity in therapeutic relationships. *Sociology* 1992;**26**:589–602. https://doi.org/10.1177/0038038592026004003

26. Patterson M, Nolan M, Rick J, Brown J, Adams R, Musson G. *From Metrics to Meaning: Culture Change and Quality of Acute Hospital Care for Older People*. Report for the National Institute for Health Research Service Delivery and Organisation programme. London: The Stationery Office; 2011.

27. Maben J, Peccei R, Adams M, Robert G, Richardson A, Murrells T, Morrow E. *Exploring the Relationship Between Patients' Experiences of Care and the Influence of Staff Motivation, Affect and Wellbeing. Health Serv Deliv Res* 2012. Available online: www.netscc.ac.uk/hsdr/files/project/SDO_FR_08-1819-213_V01.pdf (accessed 18 February 2017).

28. Bridges J, Fuller A. Creating learning environments for compassionate care: a programme to promote compassionate care by health and social care teams. *Int J Older People Nurs* 2015;**10**:48–58. https://doi.org/10.1111/opn.12055

29. Fuller A, Unwin L. Expansive learning environments: integrating organizational and personal development. In Rainbird H, Fuller A, Munro A, editors. *Workplace Learning in Context*. London: Routledge; 2004. pp. 126–44.

30. Bridges J, Tziggili M. Piloting discovery interview technique to explore its utility in improving dignity in acute care for older people. *Int Prac Dev J* 2011;**1**:Article 4.

31. Fuller A. Theories of Learning and Communities of Practice. In Hughes J, Jewson N, Unwin L, editors. *Learning in Communities of Practice*. London: Routledge; 2007. pp. 17–29.

32. McCormack B, Dewar B, Wright J, Garbett R, Harvey G, Ballantine K. *A Realist Synthesis of Evidence Relating to Practice Development: Final Report to NHS Education for Scotland and NHS Quality Improvement Scotland*. Belfast: University of Ulster; 2006.

33. Davies S, Nolan M, Brown J, Wilson F. *Dignity on the Ward: Promoting Excellence in Care. Good Practice in Acute Hospital Care for Older People*. London: Age UK; 1999.

34. Meyer J, Johnson B, Procter S, Bryar R, Rozmovits L. Practitioner research: exploring issues in relation to research capacity-building. *Nurs Times Res* 2003;**8**:407–17. https://doi.org/10.1177/136140960300800603

35. Young S, Nixon E, Hinge D, McFadyen J, Wright V, Lambert P, *et al.* Action learning: a tool for the development of strategic skills for Nurse Consultants? *J Nurs Manag* 2010;**18**:105–10. https://doi.org/10.1111/j.1365-2834.2009.01059.x

36. Nicholson C, Flatley M, Wilkinson C, Meyer J, Dale P, Wessel L. Everybody matters 2: promoting dignity in acute care through effective communication. *Nurs Times* 2010;**106**:12–14.

37. McGill I, Beaty L. *Action Learning: A Guide for Professional, Management and Educational Development*. London: Kogan Page Ltd; 1992.

38. Flatley M. *The Role of the Ward Leader in the Delivery of Dignified Care*. Paper presented at RCN 2013 Annual International Nursing Research Conference, 20–22 March 2013, Belfast.

39. Parker VA. Connecting relational work and workgroup context in caregiving organizations. *J Appl Behav Sci* 2002;**38**:276. https://doi.org/10.1177/0021886302038003002

40. Clarke N. Developing emotional intelligence through workplace learning: findings from a case study in healthcare. *Human Resource Development International* 2006;**9**:447–65. https://doi.org/10.1080/13678860601032585

41. Nolan M. Creating an Enriched Environment of Care for Older People, Staff and Family Carers: Relational Practice and Organizational Culture Change in Health and Social Care. In Keating MA, McDermott AM, Montgomery K, editors. *Patient-centred Health Care: Achieving Co-ordination, Communication and Innovation*. Houndmills: Palgrave Macmillan; 2013. pp. 78–89. https://doi.org/10.1057/9781137308931_7

42. Dean R, Proudfoot R, Lindesay J. The quality of interactions schedule (QUIS): development, reliability and use in the evaluation of two domus units. *Int J Geriatr Psychiatry* 1993;**8**:819–26. https://doi.org/10.1002/gps.930081004

43. Bridges J, Flatley M, Meyer J, Brown Wilson C. Best practice for older people in acute care settings (BPOP): guidance for nurses (2009). *Nurs Stand* 2009;**24**.

44. Warfield C, Manley K. Nursing development unit: developing a new philosophy in the NDU. *Nurs Stand* 1990;**4**:27–30. https://doi.org/10.7748/ns.4.41.27.s37

45. Dewar B, Mackay R. Appreciating and developing compassionate care in an acute hospital setting caring for older people. *Int J Older People Nurs* 2010;**5**:299–308. https://doi.org/10.1111/j.1748-3743.2010.00251.x

46. Bridges J, Flatley M, Nicholson C, Meyer J. Best practice for older people in acute care settings (BPOP): essential guide (2009). *Nurs Stand* 2009;**24**.

47. Nicholson C, Flatley M, Wilkinson C, Meyer J, Dale P, Wessel L. Everybody matters. 1: How getting to know your patients helps to promote dignified care. *Nurs Times* 2010;**106**:12–14.

48. Nicholson C, Flatley M, Wilkinson C. Everybody matters 3: engaging patients and relatives in decision making to promote dignity. *Nurs Times* 2010;**106**:10–12.

49. Dewar B, Nolan M. Caring about caring: developing a model to implement compassionate relationship centred care in an older people care setting. *Int J Nurs Stud* 2013;**50**:1247–58. https://doi.org/10.1016/j.ijnurstu.2013.01.008

50. Dewar B. *Caring about Caring: An Appreciative Inquiry about Compassionate Relationship Centred Care*. PhD thesis. Edinburgh: Edinburgh Napier University; 2011.

51. Tadd W, Hillman A, Calnan S, Calnan M, Bayer T, Read S. Right place-wrong person: dignity in the acute care of older people. *Qual Ageing Older Adults* 2011;**12**:33–43. https://doi.org/10.5042/qiaoa.2011.0143

52. Adamson E, Dewar B, Donaldson JH, Gentleman M, Gray MA, Horsburgh D, *et al. Enhancing Patient Care by Promoting Compassionate Practice*. Edinburgh: Edinburgh Napier University; 2012.

53. Hawe P. Lessons from complex interventions to improve health. *Annu Rev Public Health* 2015;**36**:307–23. https://doi.org/10.1146/annurev-publhealth-031912-114421

54. Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. *Am J Community Psychol* 2009;**43**:267–76. https://doi.org/10.1007/s10464-009-9229-9

55. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*: Wiley Online Library; 2008. https://doi.org/10.1002/9780470712184

56. Sackett D, Richardson W, Rosenberg W, Haynes R. *Evidence-Based Medicine: How to Practice and Teach EBM*. Edinburgh: Churchill Livingstone; 1997.

57. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;**336**:924–6. https://doi.org/10.1136/bmj.39489.470347.AD

58. Greenhalgh T. *How to Read a Paper: The Basics of Evidence-Based Medicine*. Chichester: John Wiley and Sons; 2014.

59. Borek AJ, Abraham C, Smith JR, Greaves CJ, Tarrant M. A checklist to improve reporting of group-based behaviour-change interventions. *BMC Public Health* 2015;**15**:963. https://doi.org/10.1186/s12889-015-2300-6

60. Boscart VM. A communication intervention for nursing staff in chronic care. *J Adv Nurs* 2009;**65**:1823–32. https://doi.org/10.1111/j.1365-2648.2009.05035.x

61. Glembocki MM, Dunn KS. Building an organizational culture of caring: caring perceptions enhanced with education. *J Contin Educ Nurs* 2010;**41**:565–70. https://doi.org/10.3928/00220124-20100701-05

62. Wasner M, Longaker C, Fegg MJ, Borasio GD. Effects of spiritual care training for palliative care professionals. *Palliat Med* 2005;**19**:99–104. https://doi.org/10.1191/0269216305pm995oa

63. Ançel G. Developing empathy in nurses: an inservice training program. *Arch Psychiatr Nurs* 2006;**20**:249–57. https://doi.org/10.1016/j.apnu.2006.05.002

64. La Monica EL, Wolf RM, Madea AR, Oberst MT. Empathy and nursing care outcomes. *Sch Inq Nurs Pract* 1987;**1**:197–213.

65. McGilton K, Sorin-Peters R, Sidani S, Rochon E, Boscart V, Fox M. Focus on communication: increasing the opportunity for successful staff–patient interactions. *Int J Older People Nurs* 2011;**6**:13–24. https://doi.org/10.1111/j.1748-3743.2010.00210.x

66. Chenoweth L, Forbes I, Fleming R, King MT, Stein-Parbury J, Luscombe G, *et al.* PerCEN: a cluster randomized controlled trial of person-centered residential care and environment for people with dementia. *Int Psychogeriatr* 2014;**26**:1147–60. https://doi.org/10.1017/S1041610214000398

67. McCance T, Slater P, McCormack B. Using the caring dimensions inventory as an indicator of person-centred nursing. *J Clin Nurs* 2009;**18**:409–17. https://doi.org/10.1111/j.1365-2702.2008.02466.x

68. Pipe TB, Mishark K, Hansen RP, Hentz JG, Hartsell Z. Rediscovering the art of healing connection by creating the Tree of Life poster. *J Gerontol Nurs* 2010;**36**:47–55. https://doi.org/10.3928/00989134-20100330-04

69. Brown Wilson C, Swarbrick C, Pilling M, Keady J. The senses in practice: enhancing the quality of care for residents with dementia in care homes. *J Adv Nurs* 2013;**69**:77–90. https://doi.org/10.1111/j.1365-2648.2012.05992.x

70. Finnema E, De Lange J, Dröes RM, Ribbe M, Van Tilburg W. The quality of nursing home care: do the opinions of family members change after implementation of emotion-oriented care? *J Adv Nurs* 2001;**35**:728–40. https://doi.org/10.1046/j.1365-2648.2001.01905.x

71. McGilton KS, O'Brien-Pallas LL, Darlington G, Evans M, Wynn F, Pringle DM. Effects of a relationship-enhancing program of care on outcomes. *J Nurs Scholarsh* 2003;**35**:151–6. https://doi.org/10.1111/j.1547-5069.2003.00151.x

72. Ho AH, Dai AA, Lam SH, Wong SW, Tsui AL, Tang JC, Lou VW. Development and pilot evaluation of a novel dignity-conserving end-of-life (EoL) care model for nursing homes in Chinese societies. *Gerontologist* 2016;**56**:578–89. https://doi.org/10.1093/geront/gnv037

73. Flarity K, Gentry JE, Mesnikoff N. The effectiveness of an educational program on preventing and treating compassion fatigue in emergency nurses. *Adv Emerg Nurs J* 2013;**35**:247–58. https://doi.org/10.1097/TME.0b013e31829b726f

74. Potter P, Deshields T, Rodriguez S. Developing a systemic program for compassion fatigue. *Nurs Adm Q* 2013;**37**:326–32. https://doi.org/10.1097/NAQ.0b013e3182a2f9dd

75. Pålsson MB, Hallberg IR, Norberg A, Björvell H. Burnout, empathy and sense of coherence among Swedish district nurses before and after systematic clinical supervision. *Scand J Caring Sci* 1996;**10**:19–26. https://doi.org/10.1111/j.1471-6712.1996.tb00305.x

76. Gauthier T, Meyer RM, Grefe D, Gold JI. An on-the-job mindfulness-based intervention for pediatric ICU nurses: a pilot. *J Pediatr Nurs* 2015;**30**:402–9. https://doi.org/10.1016/j.pedn.2014.10.005

77. Horner JK, Piercy BS, Eure L, Woodard EK. A pilot study to evaluate mindfulness as a strategy to improve inpatient nurse and patient experiences. *Appl Nurs Res* 2014;**27**:198–201. https://doi.org/10.1016/j.apnr.2014.01.003

78. Palmer E. Compassionate caring: an evaluated pilot of mindfulness-based cognitive therapy training for hospice at home nurses (conference abstract). *J Palliat Care* 2010;**26**:211–2.

79. Langewitz W, Heydrich L, Nübling M, Szirt L, Weber H, Grossman P. Swiss Cancer League communication skills training programme for oncology nurses: an evaluation. *J Adv Nurs* 2010;**66**:2266–77. https://doi.org/10.1111/j.1365-2648.2010.05386.x

80. Taylor EJ, Mamier I, Bahjri K, Anton T, Petersen F. Efficacy of a self-study programme to teach spiritual care. *J Clin Nurs* 2009;**18**:1131–40. https://doi.org/10.1111/j.1365-2702.2008.02526.x

81. Yeakel S, Maljanian R, Bohannon RW, Coulombe KH. Nurse caring behaviors and patient satisfaction: improvement after a multifaceted staff intervention. *J Nurs Adm* 2003;**33**:434–6. https://doi.org/10.1097/00005110-200309000-00002

82. Searcy LM. *The Effects of Nurse Empathy Training as Measured by Patient Satisfaction Levels*. PhD thesis. United States International University, USA; 1990.

83. Puentes WJ. *Effects of a Reminiscence Learning Experience on Registered Nurses' Attitudes Toward and Empathy with Older Adults*. PhD thesis. Chester, PA: Widener University; 1995.

84. World Health Organization (WHO). *International Classification of Functioning, Disability and Health*. Geneva: WHO; 2001.

85. La Monica E. Construct validity of an empathy instrument. *Res Nurs Health* 1981;**4**:389–400. https://doi.org/10.1002/nur.4770040406

86. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, *et al.* Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;**348**:g1687. https://doi.org/10.1136/bmj.g1687

87. Möhler R, Bartoszek G, Köpke S, Meyer G. Proposed criteria for reporting the development and evaluation of complex interventions in healthcare (CReDECI): guideline development. *Int J Nurs Stud* 2012;**49**:40–6. https://doi.org/10.1016/j.ijnurstu.2011.08.003

88. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *Int J Nurs Stud* 2013;**50**:587–92. https://doi.org/10.1016/j.ijnurstu.2012.09.010

89. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. *Developing and Evaluating Complex Interventions: New Guidance*. London: Medical Research Council; 2008. URL: www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/ (accessed 18 February 2017).

90. May C, Finch T, Mair F, Ballini L, Dowrick C, Eccles M, *et al.* Understanding the implementation of complex interventions in health care: the normalization process model. *BMC Health Serv Res* 2007;**7**:148. https://doi.org/10.1186/1472-6963-7-148

91. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, *et al.* Designing and evaluating complex interventions to improve health care. *BMJ* 2007;**334**:455–9. https://doi.org/10.1136/bmj.39108.379965.BE

92. Eldridge SM, Lancaster GA, Campbell MJ, Thabane L, Hopewell S, Coleman CL, Bond CM. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. *PLOS ONE* 2016;**11**:e0150205. https://doi.org/10.1371/journal.pone.0150205

93. EuroQoL Group. *EQ-5D-5L User Guide Version 2.1*. 2015. URL: www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-5L_UserGuide_2015.pdf (accessed 18 February 2017).

94. May C. Agency and implementation: understanding the embedding of healthcare innovations in practice. *Soc Sci Med* 2013;**78**:26–33. https://doi.org/10.1016/j.socscimed.2012.11.021

95. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociology* 2009;**43**:535–54. https://doi.org/10.1177/0038038509103208

96. Health Advisory Service. *Not Because They Are Old: An Independent Inquiry into the Care of Older People on Acute Wards in General Hospitals*. London: Health Advisory Service 2000; 1998.

97. Brooker D. Looking at them, looking at me. A review of observational studies into the quality of institutional care for elderly people with dementia. *J Ment Health* 1995;**4**:145–56. https://doi.org/10.1080/09638239550037686

98. Fritsch T, Kwak J, Grant S, Lang J, Montgomery RR, Basting AD. Impact of TimeSlips, a creative expression intervention program, on nursing home residents with dementia and their caregivers. *Gerontologist* 2009;**49**:117–27. https://doi.org/10.1093/geront/gnp008

99. McLean C, Griffiths P, Mesa-Eguiagaray I, Pickering RM, Bridges J. Reliability, feasibility, and validity of the quality of interactions schedule (QuIS) in acute hospital care: an observational study. *BMC Health Serv Res* 2017;**17**:380. https://doi.org/10.1186/s12913-017-2312-2

100. Murrells T, Robert G, Adams M, Morrow E, Maben J. 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. https://doi.org/10.1136/bmjopen-2012-002211

101. Williams AM, Kristjanson LJ. Emotional care experienced by hospitalised patients: development and testing of a measurement instrument. *J Clin Nurs* 2009;**18**:1069–77. https://doi.org/10.1111/j.1365-2702.2008.02586.x

102. Williams AM, Pienaar C, Toye C, Joske D, Lester L. Further psychometric testing of an instrument to measure emotional care in hospital. *J Clin Nurs* 2011;**20**:3472–82. https://doi.org/10.1111/j.1365-2702.2011.03846.x

103. Williams AM, Dawson SS, Kristjanson LJ. Translating theory into practice: using Action Research to introduce a coordinated approach to emotional care. *Patient Educ Couns* 2008;**73**:82–90. https://doi.org/10.1016/j.pec.2008.04.011

104. Hojat M, Mangione S, Nasca TJ, Cohen MJ, Gonnella JS, Erdmann JB, *et al.* The Jefferson Scale of Physician Empathy: development and preliminary psychometric data. *Educ Psychol Meas* 2001;**61**:349–65. https://doi.org/10.1177/00131640121971158

105. Hojat M, Vergare MJ, Maxwell K, Brainard G, Herrine SK, Isenberg GA, *et al.* The devil is in the third year: a longitudinal study of erosion of empathy in medical school. *Acad Med* 2009;**84**:1182–91. https://doi.org/10.1097/ACM.0b013e3181b17e55

106. Hojat M, Mangione S, Nasca TJ, Rattner S, Erdmann JB, Gonnella JS, Magee M. An empirical study of decline in empathy in medical school. *Med Educ* 2004;**38**:934–41. https://doi.org/10.1111/j.1365-2929.2004.01911.x

107. Aiken LH, Sermeus W, Van den Heede K, Sloane DM, Busse R, McKee M, *et al.* Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. *BMJ* 2012;**344**:e1717. https://doi.org/10.1136/bmj.e1717

108. Bruyneel L, Van den Heede K, Diya L, Aiken L, Sermeus W. Predictive validity of the International Hospital Outcomes Study questionnaire: an RN4CAST pilot study. *J Nurs Scholarsh* 2009;**41**:202–10. https://doi.org/10.1111/j.1547-5069.2009.01272.x

109. Maslach C, Jackson SE. The measurement of experienced burnout. *J Org Behav* 1981;**2**:99–113. https://doi.org/10.1002/job.4030020205

110. Department for Constitutional Affairs. *Mental Capacity Act 2005 Code of Practice*. London: The Stationery Office; 2007.

111. Dewing J. Participatory research: a method of process consent with persons who have dementia. *Dementia: Int J Soc Res Practice* 2007;**6**:11–25. https://doi.org/10.1177/1471301207075625

112. Dewing J. Process consent and research with older persons living with dementia. *Res Ethics Rev* 2008;**4**:59–64. https://doi.org/10.1177/174701610800400205

113. Watson H, Booth J, Whyte R. Observation. In Gerrish K, Lacey A, editors. *The Research Process in Nursing*. Oxford: Wiley-Blackwell; 2010:382–94.

114. Gold R. Roles in Sociological Field Observation. In McCall G, Simmons J, editors. *Issues in Participant Observation: A Text and Reader*. London: Addison Wesley; 1969.

115. 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 Care* 2002;**14**:353–8. https://doi.org/10.1093/intqhc/14.5.353

116. Pawson R. *The Science of Evaluation: A Realist Manifesto*. London: Sage; 2013. https://doi.org/10.4135/9781473913820

117. IBM Corp. *IBM SPSS Statistics for Windows, Version 22.0*. Armonk, NY: IBM Corp.; 2013.

118. Barker HR, Griffiths P, Mesa-Eguiagaray I, Pickering R, Gould L, Bridges J. Quantity and quality of interaction between staff and older patients in UK hospital wards: a descriptive study. *Int J Nurs Stud* 2016;**62**:100–7. https://doi.org/10.1016/j.ijnurstu.2016.07.018

119. Appleby J. Are friends and family tests useful: agree, disagree, neither, don't know? *BMJ* 2013;**346**:16–17. https://doi.org/10.1136/bmj.f2960

120. Gould LJ, Griffiths P, Barker HR, Libberton P, Mesa-Eguiagaray I, Pickering RM, *et al.* Compassionate care intervention for hospital nursing teams caring for older people: a pilot cluster randomised trial. *BMJ Open* 2018;**8**:e018563

121. Bridges J, May C, Fuller A, Griffiths P, Wigley W, Gould L, *et al.* Optimising impact and sustainability: a qualitative process evaluation of a complex intervention targeted at compassionate care. *BMJ Qual Saf* 2017;**26**:970–7. https://doi.org/10.1136/bmjqs-2017-006702

122. Mesa-Eguiagaray I, Böhning D, McLean C, Griffiths P, Bridges J, Pickering RM. Inter-rater reliability of the QuIS as an assessment of the quality of staff-inpatient interactions. *BMC Med Res Methodol* 2016;**16**:171. https://doi.org/10.1186/s12874-016-0266-4

123. Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG. *Cost-effectiveness in Health and Medicine*. Oxford: Oxford University Press; 2016. https://doi.org/10.1093/acprof:oso/9780190492939.001.0001

124. Adams A, Lugsden E, Chase J, Arber S, Bond S. Skill-mix changes and work intensification in nursing. *Work Employment Soc* 2000;**14**:541–55. https://doi.org/10.1177/09500170022118563

125. Bolton SC. *Emotion Management in the Workplace*. Basingstoke: Palgrave Macmillan; 2005.

126. Hyde P, Granetr E, Hassard J, McCann L. *Deconstructing the Welfare State: Managing Healthcare in the Age of Reform*. London and New York, NY: Routledge; 2016.

127. Mimura C, Griffiths P. The effectiveness of current approaches to workplace stress management in the nursing profession: an evidence based literature review. *Occup Environ Med* 2003;**60**:10–15. https://doi.org/10.1136/oem.60.1.10

128. Billett S. Learning Through Work: Workplace Participatory Practices. In Rainbird H, Fuller A, Munro A, editors. *Workplace Learning in Context*. London: Routledge; 2004. pp. 109–25

129. May C. Towards a general theory of implementation. *Implement Sci* 2013;**8**:18. https://doi.org/10.1186/1748-5908-8-18

130. Martin GP, Weaver S, Currie G, Finn R, McDonald R. Innovation sustainability in challenging health-care contexts: embedding clinically led change in routine practice. *Health Serv Manage Res* 2012;**25**:190–9. https://doi.org/10.1177/0951484812474246

131. Bridges J, Fitzgerald L, Meyer J. New workforce roles in health care: exploring the longer-term journey of organisational innovations. *J Health Organ Manag* 2007;**21**:381–92. https://doi.org/10.1108/jhom.2007.02521daa.001

132. Ferlie E, Ashburner L, Fitzgerald L, Pettigrew A. *The New Public Management in Action*. Oxford: Oxford University Press; 1997.

133. Dawson S, Dargie C. New Public Management: A Discussion with Special Reference to UK Health. In McLaughlin K, Osborne SP, Ferlie E, editors. *New Public Management: Current Trends and Future Prospects*. London: Routledge; 2002. pp. 34–56.

134. Exworthy M, Halford S. *Professionals and the New Managerialism in the Public Sector*. Buckingham: Open University Press; 1999.

135. Fitzgerald L, Lilley C, Ferlie E, Addicott R, McGivern G, Buchanan D. *Managing Change and Role Enactment in the Professionalized Organization*. London: NCCSDO; 2006. URL: www.nets.nihr.ac.uk/__data/assets/pdf_file/0005/64337/FR-08-1201-021.pdf (accessed 18 February 2017).

136. Currie G, Finn R, Martin G. Role transition and the interaction of relational and social identity: new nursing roles in the English NHS. *Org Stud* 2010;**31**:941–61. https://doi.org/10.1177/0170840610373199

137. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ* 2003;**327**:785–9. https://doi.org/10.1136/bmj.327.7418.785

138. Gilbody S, Bower P, Torgerson D, Richards D. Cluster randomized trials produced similar results to individually randomized trials in a meta-analysis of enhanced care for depression. *J Clin Epidemiol* 2008;**61**:160–8. https://doi.org/10.1016/j.jclinepi.2007.04.015

139. Campbell MJ, Walters SJ. *How to Design, Analyse and Report Cluster Randomised Trials In Medicine And Health Related Research*. Chichester: John Wiley & Sons; 2014.

140. Eldridge S, Kerry S. *A Practical Guide to Cluster Randomised Trials in Health Services Research*. Chichester: Wiley; 2012. https://doi.org/10.1002/9781119966241

141. Goldberg SE, Harwood RH. Experience of general hospital care in older patients with cognitive impairment: are we measuring the most vulnerable patients' experience? *BMJ Qual Safety* 2013;**22**:977–80. https://doi.org/10.1136/bmjqs-2013-001961

142. van Baalen A, Vingerhoets AJJM, Sixma HJ, de Lange J. How to evaluate quality of care from the perspective of people with dementia: an overview of the literature. *Dementia* 2010;**10**:112–37. https://doi.org/10.1177/1471301210369320

143. Alzheimer's Society. *Counting the Cost: Caring for People with Dementia on Hospital Wards*. London: Alzheimer's Society; 2009.

144. Mukadam N, Sampson EL. A systematic review of the prevalence, associations and outcomes of dementia in older general hospital inpatients. *Int Psychogeriatr* 2011;**23**:344–55. https://doi.org/10.1017/S1041610210001717

145. Ball JE, Murrells T, Rafferty AM, Morrow E, Griffiths P. 'Care left undone' during nursing shifts: associations with workload and perceived quality of care. *BMJ Qual Saf* 2014;**23**:116–25. https://doi.org/10.1136/bmjqs-2012-001767

146. Garbett R, McCormack B. A concept analysis of practice development. *Nursing Times Research* 2002;**7**:87–100. https://doi.org/10.1177/136140960200700203

147. Watson J. *Nursing: The Philosophy and Science of Caring*. Boulder, CO: University Press of Colorado; 2008.

Appendix 1 Systematic review MEDLINE and CINAHL searches

Database	Main search	Additional keywords	Limitations
MEDLINE	compassion* OR empath* OR Empathy[a] OR person centered care OR person centred care OR relationship centered care OR relationship centred care OR client centered care OR client centred care OR Patient-Centered Care[a] OR Patient centered care OR patient centred care OR dignity	AND randomized controlled trial OR randomized controlled trial OR evaluation OR Nursing Evaluation Research[a] OR quasi experiment OR controlled trial OR time series OR Controlled Before-After Studies[a] OR before and after OR Comparative Study[a] AND nurs* OR Occupational Groups[a]	English
CINAHL	compassion* OR empath* OR Empathy[b] OR person centered care OR person centred care OR relationship centered care OR relationship centred care OR client centered care OR client centred care OR Patient-Centered Care[b] OR Patient centered care OR patient centred care OR dignity OR Human Dignity[b]	AND randomized controlled trial OR Randomized Controlled Trial[b] OR Evaluation[b] OR evaluation OR quasi experiment OR controlled trial OR time series OR Time Series[b] OR Controlled Before-After Studies[b] OR before and after OR Comparative Studies[b] OR comparative study AND Nurses[b] OR nurs* OR occupational groups	English, excluded MEDLINE records
Cochrane	Same search terms as above	Same search terms as above	English

a MeSH term.
b Subject heading.

Appendix 2 Systematic review summary study tables

TABLE 41 Interventions focusing on training

Number	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/ measures	Other outcomes	Results
1	Ançel 2006[63] Uncontrolled before-and-after study	Low	Nurses, $n = 190$ Adult department, hospital setting, Turkey	C: no control group I: training programme in empathic skills communication	Empathic communication skills, ECS-B	Satisfaction with the programme Trainees' satisfaction forms	Significant increase in nurses' empathic skills after training (ECS-B score for intervention group = + 24.9; $p = 0.05$) Of the nurses: 98.9% found the trainers adequate, 99.2% found the materials and techniques adequate, 97.7% found the content and its relevance adequate (trainees' satisfaction form)
2	Boscart 2009[60] Uncontrolled before-and-after study	Low	Patients, $n = 27$ RNs and licensed practical nurses, $n = 27$ Hospital setting, Canada	C: no control group I: 3-hour educational intervention on verbal interactions between nursing staff and patients	Quality of verbal interactions (quantified content analysis)	None	Significant improvement in positive nurse–patient interactions ($p = 0.001$)
3	Glembocki and Dunn 2010[61] Uncontrolled before-and-after study	Low	RNs, $n = 39$ Hospital setting, USA	C: no control group I: educational intervention Reigniting the spirit of caring using a 3-day seminar, focusing on relationship with self, colleagues and patients	CAC	None	Significant difference in CAC score between pre- and post-test ($p < 0.05$)

Number	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/ measures	Other outcomes	Results
4	La Monica et al. 1987[64] Cluster randomised controlled study	Medium	Nurses, n = 115 Patients, n = 656 Hospital setting, USA	C: 16-hour course in physical assessment I: empathy training programme lasting 14–16 hours	Empathy outcomes, ECRS	Patient satisfaction, LOPSS Patient mood and satisfaction, MAACL	No significant difference in empathy outcomes in nurses' and patients' rating after the intervention (ECRS nurses 171.3 vs. 177.0; $p > 0.05$; ECRS patients 201.0 vs. 228.5; $p = 0.05$) No significant difference in patient satisfaction (LOPSS $p > 0.05$) and mood between the experimental and control groups after treatment, but a significant difference in anxiety and hostility among patients cared for by the intervention group (MAACL $p = 0.004$)
5	Langewitz et al. 2010[79] Uncontrolled before-and-after study	Low	Nurses, n = 70 Hospital setting, Switzerland	C: no control group I: workshop-based communication skills training 2.5-day seminar, including role-play, video and telephone supervision (5 × 30 min) and booster session after 6 months	Patient-centred communication style, RIAS	None	Significant difference in patient-centeredness after the intervention (RIAS $p < 0.003$)
6	Puentes 1995[83] Post-test only randomised, controlled study	Low	RNs, n = 98 Hospital setting, USA	C = usual practice I = 1-hour reminiscence learning experience educational programme for nurses focusing on the incorporation of reminiscence techniques into interactions with clients, plus request to participants to implement techniques during the subsequent 3 weeks	Empathy levels, HES	Attitudes towards older adults, KAOP	Significant difference in empathy levels between experimental and control groups (HES 19.12 vs. 17.84; $p < 0.05$) Significant difference in attitudes towards older adults between experimental and control groups (KAOP 153.27 vs. 140.96; $p < 0.000$)

TABLE 41 Interventions focusing on training (*continued*)

Number	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/ measures	Other outcomes	Results
7	Searcy 1990[82] Before-and-after study with separate intervention and control groups	Medium	Patients, $n = 298$ Hospital setting, USA	C = usual practice I = 2 × 1-hour classes over a 2-week period aimed at enhancing nurses' skills for perceiving and responding with empathy	Empathy levels LEP	Patient satisfaction, including dissatisfaction, perceptions of interpersonal support and good impression of nursing care, LOPSS	No significant difference after training on empathy (LEP 2.69 vs. 2.74; $p = 0.48$), total patient satisfaction (LOPSS 112.45 vs. 112.16; $p = 0.91$), dissatisfaction (2.65 vs. 2.71; $p = 0.39$), interpersonal support (2.75 vs. 2.73; $p = 0.75$), or good impression (2.83 vs. 2.78; $p = 0.4$) in the intervention group No significant differences from control ($p > 0.5$)
8	Taylor *et al.* 2009[80] Uncontrolled before-and-after study	Low	RNs and nursing students, $n = 201$ Religious university, non-religious university, religious health-care institution, non-religious health-care institution, USA	C = no control group I = mailed self-study programme including 100-page interactive workbook and DVD on talking with patients about spirituality	Ability to respond empathically to patient spiritual pain RES	Personal spiritual experience, DSE Attitude towards spiritual caregiving, SCPS-R Knowledge about how to communicate to provide spiritual care, CSCT	Significant post-intervention improvements in empathic response to patient spiritual pain (RES + 12.2; $p < 0.0001$), personal spiritual experience (DSE –3.2; $p < 0.0001$), attitude to spiritual caregiving SCPS-R + 3.0; $p < 0.0001$) and knowledge about communication for spiritual care (CSCT + 2.0; $p < 0.0001$)

Number	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/ measures	Other outcomes	Results
9	Wasner et al. 2005[62] Uncontrolled before-and-after study	Low	Palliative care professionals, $n = 63$ Range of medical and social-care settings, Germany	C = no control group I = 3.5-day training to teach active and compassionate listening, and recognition and addressing causes of emotional and spiritual suffering; includes practical exercises and introducing contemplation and meditation practices	Self-transcendence: sense of connectedness within the self and with one's environment STS Compassion with severely ill and dying persons[a] Compassion with oneself[a]	Spiritual well-being, FACIT-Sp Religiosity, IIR Quality of life[a] Attitude towards one's family[a] Fear of dying process and death[a] Contentment with job[a] Meaningfulness of job[a] Attitudes towards colleagues[a] Perception of work-related stress[a]	Significant improvement in compassion for the dying ($+ 0.5$; $p < 0.01$) and for oneself ($+ 0.9$; $p < 0.01$) after the training and sustained 6 months later ($+ 0.5$, $p < 0.05$; $+ 0.7$, $p < 0.05$). Self-transcendence significantly improved after the training (STS $+ 1.9$; $p < 0.01$) but no significant difference from baseline to 6 months later (STS $+ 0.8$; $p > 0.05$) Significant improvement in spiritual well-being after the training (FACIT-Sp $+ 2.0$; $p < 0.01$) and sustained 6 months later ($+ 0.8$; $p < 0.05$) Significant improvements after the training of quality of life ($+ 0.6$; $p < 0.05$), attitudes towards family ($+ 0.7$; $p < 0.01$), fear of dying ($+ 0.6$; $p < 0.05$), fear of death ($+ 0.7$; $p < 0.01$), work satisfaction ($+ 0.7$; $p < 0.01$), meaningfulness of work ($+ 0.4$; $p < 0.01$), attitude towards colleagues ($+ 0.4$; $p < 0.05$), and work-related stress ($+ 1.3$; $p < 0.01$). Significant differences from baseline sustained at 6 months in all measures using numeric rating (0–10) with exception of quality of life, fear of death and meaningfulness of work No significant difference in religiosity between baseline and 6 months (IIR $–0.4$; $p > 0.05$)

TABLE 41 Interventions focusing on training (*continued*)

Number	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/ measures	Other outcomes	Results
10	Yeakel *et al.* 2003[81] Uncontrolled before-and-after study	Low	Patients (*n* = 477) Hartford hospital general surgery unit, USA	C = no control group I = educational programme for RNs during 1 month (a formal education session, staff identification of goals, peer reinforcement, incorporation of goals into performance management, posting of examples of caring behaviours on the unit to serve as reminders for the staff)	Nurse caring Wolf's Caring Behaviours Inventory	Patient satisfaction Hartford Hospital Satisfaction survey	Patients admitted after the intervention rated nurses' caring higher than patients admitted before the intervention (Z = –2.14; *p* = 0.032) Patients admitted after the intervention provided higher ratings of satisfaction than patients admitted before the intervention (Z = –2.86, *p* = 0.004)

C, control group; CAC, Caring Asesssment for Caregiver tool; CSCT, Communicating for Spiritual Care Test; DSE, Daily Spiritual Experience Scale; DVD, digital versatile disc; ECRS, Empathy Construct Rating Scale; ECS-B, Empathic Communication Skill B; FACIT-Sp, Functional Assessment of Chronic Illness Therapy—Spiritual Well-Being Scale; HES, Hogan's Empathy Scale; I, intervention group; IIR, Idler Index of Religiosity; KAOP, Kogan's Attitudes Toward Old People scale; LEP, La Monica Empathy Profile; LOPSS, La Monica/Oberst Patient Satisfaction Scale; MAACL, Multiple Affect Adjective Check List; RES; Response Empathy Scale; RIAS, Roter Interaction Analysis System; SCPS-R, Spiritual Care Perspective Scale-Revised; STS, Self-Transcendence Scale.
a Numeric rating (1–10).

TABLE 42 Interventions focusing on care models

Number	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/measures	Other outcomes	Results
1	Brown Wilson *et al.* 2013[69] Uncontrolled before-and-after study	Low	Staff, *n* = 11 Residents, *n* = 6 Families, *n* = 4 Managers, *n* = 3 Care homes (*n* = 2), UK	C = no control group I = training programme based on the Senses Framework,[41] including eight workshops	Care profiles to assess how a service might enhance resident, staff and family's sense of continuity, significance, belonging, purpose, achievement or security		Improvements reported in staff sense of security and belonging, and in practices theorised to improve residents' sense of significance, continuity and purpose Statistical significance of changes not reported
2	Chenoweth *et al.* 2014[66] Cluster randomised controlled study	High	People with dementia, *n* = 601 Residential aged care homes (*n* = 38), Australia	C = usual practice I = implementation of PCC, PCE or PerCEN	Care interaction quality (QUIS) ERiC	Quality of life, DEMQoL Behavioural and psychological symptoms of dementia, CMAI	Care interaction quality: significant overall effect from group by time interaction, but significant improvement in PerCEN group only ($p = 0.006$) Resident emotional responses to care: no significant overall effect from group by time interaction. Significant improvement in PerCEN group only ($p = 0.01$) Quality of life: no significant overall effect from group by time interaction. Significant improvements in PCC ($p = 0.0003$) and PCE ($p = 0.02$) groups, but not in PerCEN group Agitation: significant overall effect from group by time interaction. Significant improvements in PCC ($p = 0.002$) and PCE ($p = 0.05$) groups, but not in PerCEN group

TABLE 42 Interventions focusing on care models (*continued*)

Number	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/measures	Other outcomes	Results
3	Finnema *et al.* 2001[70] Cluster randomised controlled study	High	Family members for residents, *n* = 194 Staff members, *n* = 230 Nursing homes (16 wards in 14 nursing homes), the Netherlands	C: usual practice with implementation of a model care plan I: implementing emotion-oriented care in combination with model care plan. Training and supervision in emotion-oriented care for 9 months	None	Quality of care (newly developed instrument, 18 questions)	An increase of quality of care regarding the question 'Has anyone asked you about your relative's life history after the initial intake meeting?' in the experimental group after emotion-oriented care implementation ($p = 0.05$)
4	Ho *et al.* 2016[72] Uncontrolled before-and-after study	Low	Residents, *n* = 17 Nursing homes, China	C: no control group I: implementation of dignity-conserving end-of-life care model (several components of education and supportive care, at both group and individual level, advance care planning, pain and symptom management, etc.)	None	MQoL NF-QoL	A significant deterioration in physical quality of life ($p < 0.05$), and improved support quality of life ($p < 0.05$) between pre- and post-test No significant difference in NF-QoL were found
5	McCance *et al.* 2009[67] Uncontrolled before-and-after study	Low	Nurses, *n* = 122 Patients, *n* = 107 Hospital setting, Ireland	C: no control group I: PCN intervention based on framework of PCN and a model by Garbett and McCormack[146]	PCN including CDI and NDI	None	Significant difference over time in nurses' perception of caring (CDI 0.38 vs. 0.45; $p < 0.05$) after intervention Significant difference over time in patients' perceptions of caring (NDI 0.41 vs. 0.45; $p < 0.05$)

Number	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/measures	Other outcomes	Results
6	McGilton *et al.* 2003[71] Before-and-after study with separate intervention and control groups	Medium	Residents, *n* = 50 Nursing staff, *n* = 34 Nursing homes, Canada	C: usual practice I: implementing relationship-enhancing programme of care	RC scale Close relationship with care providers (Relationship VAS) Care providers' empathic and reliable behaviour (an observational scale)	Continuity of care (the continuity index)	Significant difference in RC ($p = 0.014$) and care providers' relational behaviour ($p = 0.046$) between the experimental and control group Significant difference in continuity of care ($p < 0.001$)
7	McGilton *et al.* 2011[65] Uncontrolled before-and-after study	Low	Nurses, *n* = 18 Patients, *n* = 9 Stroke continuing care unit, Canada	C = no control group I = development of individualised patient communication plans by SLPs; nurse attendance at full-day workshop focused on communication and behavioural management strategies; implementation of nursing staff support system by SLPs: observing interactions, providing feedback and demonstrating strategies	Patient satisfaction with nurses' relational care RC scale Global perception of closeness of nurse–patient relationship Patient close VAS Provider close VAS	Patient quality of life, SAQoL Patient depression, GDS Attitude of nurses towards patients with communication impairments, CIQ	Significant post-intervention improvement in patient satisfaction with nurses' relational care (RCS + 3.1; $p = 0.024$), patient perceptions of closeness of relationship with nurses (VAS + 15.9; $p = 0.041$), patient perception of own communication abilities (SAQoL + 3.8; $p = 0.037$), and nurse attitudes towards patients with communication impairment (CIQ + 2.4; $p = 0.007$) No significant post-intervention differences in patient psychosocial well-being (SAQoL + 1.8; $p = 0.601$), patient depression (GDS + 0.3; $p = 0.848$), or nurse perceptions of closeness of relationship with patients (VAS + 3.4; $p = 0.657$)

TABLE 42 Interventions focusing on care models (*continued*)

Number	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/measures	Other outcomes	Results
8	Pipe *et al.* 2010[68] Uncontrolled before-and-after study	Low	Patients, $n = 19$ General medical ward, USA	C = no control group I = life-story intervention based on Watson's theory of human caring,[147] including trained volunteers completed life-story interviews and created a 'Tree of Life' poster for every patient	None	Quality of Life, LASA Instrument Emotional well-being, Social support, MOS Social Support Survey Hope, HHI FACIT-Sp-Ex	Quality of life: a significant improvement in physical well-being ($p = 0.02$), and emotional well-being ($p = 0.005$) after intervention No significant improvement in emotional well-being (MOS) or hope (HHI) A significant improvement in spiritual well-being (FACIT-Sp-Ex) ($p = 0.02$)

C, control group; CDI, Caring Dimensions Inventory; CIQ, Communication-Impairment Questionnaire; CMAI, Cohen-Mansfield Agitation Inventory; DEMQoL, dementia quality of life instrument; EriC; resident emotional responses in care assessment; FACIT-Sp-Ex, Functional Assessment of Chronic Illness Therapy—Spiritual Well-Being Scale Extended version; GDS, Geriatric Depression Scale; HHI, Hope Herth Hope Index; I, intervention group; LASA, linear analogue self-assessment; MOS, medical outcomes study; MQoL, McGill Quality of Life Questionnaire; NDI, Nursing Dimension Inventory; NF-QoL, Nursing facilities quality-of-life questionnaire; PCC, person-centred care; PCE, person-centred environment; PCN, person-centred nursing; PerCEN, a combination of both PCC and PCE; RC, relational care; RCS, Relational Care Scale; SAQoL, Stroke and Aphasia Quality of Life Scale; SLP, speech and language pathologist; VAS, visual analogue scale.

TABLE 43 Interventions focusing on nurse support

#	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/measures	Other outcomes	Results
1	Flarity *et al.* 2013[73] Uncontrolled before-and-after study	Low	Nurses, $n = 73$ Emergency care, USA	C: no control group I: multifaceted compassion fatigue resiliency intervention programme: 4-hour interactive seminar plus multimedia resources	Compassion satisfaction, ProQoL CS subscale Compassion fatigue, ProQoL BO subscale	Secondary traumatic stress, ProQoL STS subscale	Significant post-intervention increase in compassion satisfaction (ProQoL CS + 1.9; $p = 0.004$), and decrease in burnout (ProQoL BO –3.9; $p < 0.001$) and secondary traumatic stress (ProQoL STS –2.1; $p = 0.001$)
2	Gauthier *et al.* 2015[76] Uncontrolled before-and-after study	Low	Nurses, $n = 60$ Paediatric intensive care unit, USA	C = no control group I = 5-minute mindfulness meditation/instruction in workplace at the beginning of each shift for 30 days	Symptoms of burnout, MBI Self-compassion, SCS	Levels of stress, NSS Mindfulness, MAAS Job satisfaction	No significant differences in burnout, emotional exhaustion and depersonalisation (mean, *p* not reported). Burnout personal accomplishment increased post intervention but decreased at 1-month follow-up ($p = 0.03$) No significant increase in self-compassion (SCS difference not reported, $p = 0.26$) Significant decrease in stress from baseline (78.92) to post intervention (74.03; $p = 0.006$) and 1-month follow-up (*p* not reported) No significant differences in mindfulness (MAAS, difference not reported, $p = 0.37$), job satisfaction (positive change reported, $p = 0.15$)

TABLE 43 Interventions focusing on nurse support (*continued*)

#	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/measures	Other outcomes	Results
3	Horner *et al.* 2014[77] Before-and-after study	Low	Nurses, *n* = 43 Patients, *n* unknown Hospital setting, USA	C: usual practice I: mindfulness training programme for 10 weeks, 30 minutes once a week including education and practice	Compassion satisfaction score and burnout score, ProQoL	Level of mindfulness MAAS measure Individual and unit stress levels (VAS 1–10) HCAHPS – hospital patient survey	No significant difference in compassion satisfaction score before and after intervention (ProQoL 53.20 vs. 52.93; $p = 0.76$), or burnout score (ProQoL 46.20 vs. 45.71; $p = 0.55$) or level of mindfulness (MAAS 4.2 vs. 4.4; $p = 0.37$) in the intervention group Significant difference before and after the intervention in individual stress (individual stress level 5.0 vs. 4.2; $p = 0.10$) and unit stress (unit stress level 5.8 vs. 5.1) in the intervention group No significant difference in the control group Patient satisfaction (HCAHPS): improvement in overall scores in the intervention group (32 points) compared with the control group, and improvement in 'communication with nurses' (17 points)
4	Palmer 2010[78] Uncontrolled before-and-after study	Low	Nurses, *n* = 9 Hospice at home, UK	C = no control group I = 8-week mindfulness-based cognitive therapy training	Clinician empathy, JCES	Mindfulness, MAAS Well-being, WHO-5 EWWS	Improvements in scores across all scales reported post intervention compared with 'expected population averages' but no further details reported
5	Pålsson *et al.* 1996[75] Before-and-after study	Medium	RNs, *n* = 33 District nursing for women with newly diagnosed breast cancer, Sweden	C = 40-hour training programme on medical care and treatment for breast cancer, psychological reactions, coping strategies, crisis intervention and organisation of nursing care I = training programme (as above) + 1.5–2 hours' clinical supervision every 2–4 weeks for 15–19 sessions	Burnout, BM Empathy, ECRS	SOC	No significant difference ($p > 0.05$) after clinical supervision on burnout (BM 2.7 vs. 2.5), empathy (ECRS 419 vs. 427) or SOC (SOC 148 vs. 151) in intervention group. No significant differences from control

#	Study	Quality rating	Setting and sample	Intervention	Compassion outcomes/measures	Other outcomes	Results
6	Potter et al. 2013[74] Uncontrolled before-and-after study	Low	RNs, $n = 13$ Outpatient oncology infusion centre, USA	C = no control group I = 5-week programme involving five 90-minute sessions on compassion fatigue resiliency	Symptoms of burnout, MBI Compassion satisfaction, ProQOL IV CS subscale Compassion fatigue, ProQOL IV BO subscale	Subjective distress caused by traumatic events, including avoidance, intrusions, hyperarousal, IES-R Secondary traumatic stress, ProQOL STS subscale Nursing job satisfaction, NJSS	No significant difference in symptoms of burnout between baseline and immediate post intervention, 3 months later and 6 months later (MBI Emotional Exhaustion subscale: immediate −2.92, $p > 0.05$; 3 months −2.38, $p > 0.05$; 6 months −3.46, $p > 0.05$. MBI depersonalisation subscale: immediate −1.46, $p > 0.05$; 3 months −1.31, $p > 0.05$; 6 months −0.31, $p > 0.05$. MBI Personal Accomplishment subscale: immediate −0.92, $p > 0.05$; 3 months −1.15, $p > 0.05$; 6 months −2.15, $p > 0.05$) No significant difference in compassion satisfaction (ProQOL CS: immediate −0.38, $p > 0.05$; 3 months −1.0, $p > 0.05$; 6 months −1.23, $p > 0.05$) No significant difference in compassion fatigue (ProQOL BO: immediate −0.85, $p > 0.05$; 3 months −0.23, $p > 0.05$; 6 months −1.15, $p > 0.05$) No significant difference in job satisfaction (no further details reported) Significant improvement in subjective distress caused by traumatic events between baseline and immediate post-intervention, (IES-R + 1.24, $p = 0.04$) 3 months later (+2.4, $p < 0.001$) and 6 months later (+1.77, $p = 0.005$) Significant decline in secondary traumatic stress between baseline and 6 months (+3.54, $p = 0.044$)

BM, Burnout Measure; C, control group; ECRS, Empathy Construct Rating Scale; EWWS, Edinburgh and Warwick Wellbeing Scale; HCAHPS, Hospital Consumer Assessment of Healthcare Providers and Systems; I, intervention group; JCES, Jefferson Clinical Empathy Scale; MAAS, Mindfulness Attention Awareness Scale; NSS, Nursing Stress Scale; ProQoL BO, Professional Quality of Life test: Burnout subscale; ProQoL CS, Professional Quality of Life test: Compassion Satisfaction subscale; SCS, Self-Compassion Scale; SOC, sense of coherence; VAS, visual analogue scale; WHO-5, World Health Organization's Well-Being Index.

Appendix 3 Guidance for Quality of Interactions Schedule ratings in acute care settings

QuIS rating	Examples
Positive social **Interaction principally involving 'good, constructive, beneficial' conversation and companionship:** Polite, friendly and respectful interactions in which any element is: Casual/informal and relating to 'everyday' social topics (e.g. family; sport; weather; TV programmes) OR Responding to concerns/interests/topics introduced by the service user The service user may be expected to feel valued, cared about or respected as a **person**	General chat and conversation, on its own or during other care activities Allowing and responding to the expression of feelings and emotions Greetings that invite an response Giving time and attention to elicit people's concerns ('How are you today?')
Positive care **Interactions during the appropriate delivery of physical care:** Interactions which are polite, professional, respectful or good humoured in tone, but in which the topic is set by staff and restricted to issues of care delivery (e.g. 'your discharge'; 'your wash'; 'your medication'; 'your surgery') The service user may be expected to feel safe, secure, cared for or informed as a **patient**	Providing explanation and reassurance or encouragement while delivering care (e.g. providing encouragement to mobilise) Giving information, opportunities for questioning and checking for understanding Offering simple choices in regard to essential activities of living (e.g. do you want sugar in your tea?)
Neutral **Brief, indifferent interactions not meeting the definitions of the other categories:** Interactions which have no positive or negative aspects, and which would not be expected to impact on the feelings of the service user	Undirected greetings (if noted by service user) Putting plates down with cursory or no verbal/non-verbal contact
Negative protective **Providing care, keeping safe or removing from danger, but in a restrictive manner, without explanation or reassurance: in a way which disregards dignity or fails to demonstrate respect for the individual:** Interactions which fail to fully maintain dignity or demonstrate respect due to the focus of staff on doing their 'work'. Staff may appear rushed or task orientated The service user may be expected to feel rushed, misunderstood, frustrated or poorly informed	Failing to offer choices Incomplete/inadequate responses to a need for explanation or reassurance Keeping safe or removing from danger without explanation or reassurance (e.g. 'Don't eat that, it's been on the floor') Helping people to eat without giving them control over the speed of eating Asking people to wait for something (e.g. medication/treatment/food and drink) without a good reason or explanation

QuIS rating	Examples
Negative restrictive	Ignoring people (including not answering call bells)
Interactions that oppose or resist people's freedom of action without good reason, or which ignore them as a person:	Moving or examining people without warning or explanation
	Telling service users not to swear/show anger
Interactions which are rude/controlling or abusive and pay no regard to the perspective of the patient. Patients expressed needs/preferences are ignored or denied. Staff may be authoritative, controlling, rude or angry	Telling people to do something (e.g. button dress) without discussion, explanation or offer of help
	Telling people they cannot have something (e.g. medication/treatment/food and drink) without good reason or explanation
The service user may be expected to feel ignored or humiliated	Swearing at or physically assaulting people

Column 1 reproduced from Dean *et al.*,[42] with permission from John Wiley & Sons. Copyright © 1993 John Wiley & Sons, Ltd.
Note
Bold text indicates original guidance from Dean *et al.*[42]

Appendix 4 Example process evaluation staff interview schedules

Interview schedule for ward leaders during and following implementation period

To what extent is CLECC being made workable and integrated into everyday practice by the nursing team?

Can you tell me what you think of the CLECC intervention?

What does CLECC make you do differently?

How well does CLECC fit with other things you and the team do on the ward?

Do you think CLECC supports the delivery of compassionate care? Explain.

What does CLECC require nursing team members to do to put it into practice?

How committed is the team to CLECC? Explain why.

What has helped the team put CLECC into practice? Give examples.

What has got in the way of putting CLECC into practice? Give examples.

What factors are influencing the extent to which the nursing team can put CLECC into practice?

What resistance has there been to CLECC from the team? Give examples.

What will happen when the CLECC study finishes?

To what extent do you think CLECC is supported by your Trust?

Intervention ward nursing staff interview schedule, mid-implementation

How would you explain CLECC to a new member of staff on the ward?

What does putting CLECC into practice require people in your role (as a care assistant/registered nurse/ward manager) to do differently (if anything)?

How straightforward has it been to make the changes required by CLECC? Explain your answer.

How well does CLECC fit with other things you're (as an individual) supposed to do on the ward?

Allow answer, then use prompts to cover these questions if not already covered:

Workload 'does your workload allow the space to put CLECC into practice?'

Organisation of work 'does the way your work is organised enable you to?'

Skills 'have you been equipped with the knowledge and skills?'

Perceived work role identity 'is CLECC a relevant part of your job? Are other jobs/roles better suited?'

How well does CLECC fit with other things your colleagues in the team are supposed to do on the ward?

Prompts:

Workload 'does team workload allow the space (consider across the team)?'

Organisation of work 'does the way their work is organised enable them to (consider across the team)?'

Skills 'has everyone been equipped with the knowledge and skills needed for CLECC?'

Perceived work role identity 'is it a relevant part of each person's job, or does this vary across the team?'

How committed, if at all, is the team on your ward to CLECC? Can you give some examples of how you know this?

What resistance, if at all, has there been to CLECC among team members? Can you give some examples of how you know this?

What has helped your team put CLECC into practice?

What gets in the way of your team putting CLECC into practice?

What more needs to happen to put CLECC into practice?

Has anything significant happened on the ward or in the wider trust that has affected CLECC's use by the team, or its impact? If yes, can you tell me more?

Aside from the factors you've already talked about, what are the most important influences on whether or not the team can use CLECC in everyday practice?

Prompts:

Is it valued by managers in the organisation? Which managers?

How do you know this?

Is it supported or not by other organisational policies or priorities?

What's useful (if anything) about CLECC that you think needs to keep happening after [PDN name] has left?

Do you think CLECC is, or has the potential to, support the delivery of compassionate care on the ward? Can you explain your answer?

What needs to happen to keep CLECC going on the ward after [PDN name] has left?

Appendix 5 Patient and visitor qualitative interviews

We piloted qualitative interviews with a small number of patients ($n = 12$) and visitors ($n = 12$) about relational care on the wards to inform a future process evaluation. This appendix presents an overview of the method and findings.

The aim of this work was to assess the feasibility of interviewing hospital inpatients and visitors about their ward-based experiences of relational care.

Sampling and recruitment

Two visitors and two patients were purposively sampled and recruited from each of the six wards.

Patients with characteristics that put them at risk of a more negative experience were approached and invited to take part in an interview while still inpatients. These characteristics were high physical disability, high dependency on others, communication impairment and/or cognitive impairment. Eligible interviewees did, however, need to have capacity to decide about taking part in the research, to be oriented to their location and to have sufficient attention and ability to communicate to participate in an interview,[142] and so were excluded if cognitive or other impairments precluded their participation. Interviewers were trained to take time and use skill to maximise participation. Eligible candidates were identified with the help of the nurse in charge of the ward. Following the provision of written and verbal information about the study, eligible individuals were invited to take part and, if they agreed to take part, they signed a consent form. The interview then took place straight away, after checking with the patient that this was acceptable for them.

Visitors were invited to volunteer to be interviewed through written letters distributed by hand. Visitors who expressed an interest in taking part were then given further information about the study and, if they agreed to take part, signed a consent form. One or two visitors were interviewed at a time arranged in advance. Most were interviewed directly after agreeing to take part, providing this was convenient for them.

Data collection and analysis

Patient and visitor interviews took place on the ward in a single side-room or the ward day room. Patient interviews lasted on average 21 minutes (range 10–39 minutes). The interview schedules were designed to capture individual views and experiences and focused on relational aspects of care on that ward during that admission. Demographic information was gathered on interviewees including gender, age and patient cognitive status.

Interviews were audio-recorded, transcribed verbatim and transcripts checked for accuracy by the interviewer. Researchers kept field notes of the process of recruiting and interviewing patients and visitors.

Thematic analysis of interview text was used to examine what interviewees said and to assess the extent to which they were able to comment on relational care during admission to the ward in question. Preliminary analyses were then enhanced by the recruitment and data collection issues recorded in researcher field notes.

Findings

In total, 25 patients were approached to take part in an interview. Four were excluded after being approached because they did not have capacity. Nine declined to take part and the most common reason for declining was that patients felt too tired or unwell. Twelve patients consented to take part and all 12 went on to be interviewed. Two patients were interviewed from each ward. Nine were female and three were male. Two were aged < 60 years and 10 were aged ≥ 60 years, including five people in the 80–89 years age range, and two people in the 90–99 years age range. Two patients interviewed had evidence of cognitive impairment.

A total of 23 visitors were approached to take part in an interview. Eleven declined to take part and the most common reason for declining was that visitors were too busy. Twelve visitors consented to take part and all 12 went on to be interviewed. Two visitors were interviewed from each ward. Seven were female and five were male. Two were aged < 60 years and 10 were aged ≥ 60 years, including three people aged 70–79 years, and two aged 80–89 years. Records were not made of visitors' cognitive status.

Interviews took place while the patient was an inpatient, which required the interview process to be tailored to the particular ward environment. Interviews were conducted in a room with just the patient/visitor and interviewer present in order to offer privacy and an environment conducive to audio-recording. Visitors were generally available only during set visiting times and many prioritised spending time with the patient over being interviewed. Interviews with patients and visitors were significantly shorter than with nursing staff. Patient interviews lasted on average 21 minutes (range 10–39 minutes), visitor interviews lasted on average 20 minutes (range 10–41 minutes) and nursing interviews lasted on average 46 minutes (range 17–70 minutes). The ward routine affected when interviews could take place with patients (e.g. meal times, medication administration, medical consultations). The care provided on the ward was a priority for patients and as such took precedence over the interview starting or continuing. However, no interviews were cut short because of care needs taking priority.

The interviews focused on relational aspects of care during the current admission; however, both patients and visitors spontaneously talked about previous admissions to the same hospital but not necessarily to the same ward or for a similar issue. Patients had often experienced care in another setting in the hospital immediately prior to being admitted to the current ward. It was difficult to work out whether the experience being described was solely about the current ward. Patients were able to comment only on individual nurse–patient interactions with them as individuals, or with others observed across the bay, and could not place the interaction in the wider context of the whole ward. Patients were generally confined to bed or their bed space and usually remained in the same location for the duration of the admission. Visitors tended to be familiar only with the patient space and staff who entered the space. Patients were interviewed at one point during their admission, and although they participated only if they had mental capacity to consent, their ability to engage with the interview process was affected by their stage of recovery.

Conclusion

Recruitment to and conducting qualitative interviews with patients and visitors was feasible, but this pilot highlighted a number of issues that may have affected data quality and that indicate this may not be a successful method to explore patient and visitor views and experiences.

Interview questions for patients and visitors

What does compassionate care mean to you in hospital?

What does the term 'compassionate care' mean to you?

How important is it to you that you're (your relative is) looked after compassionately?

Do you think the nurses on this ward are compassionate? Explain your answer.

Can you tell me about a time on this ward that you felt (your relative was) well cared for?

Can you tell me about a time on this ward that you didn't feel (your relative was) well cared for?

What do nursing staff on this ward do to get to know who you are/your relative as a person?

How do the nurses on this ward make you feel when they are looking after you?/How do you feel about the way nurses on this ward are looking after your relative?

What do you do when you have concerns or worries (about your relative) in hospital?

Do you feel able to talk to a nurse on this ward about your concerns?

How do the nurses on this ward involve you in decisions about your (your relative's) care?

Do you feel that you understand what is happening to you (your relative) in hospital?

How well do you think the nurses on this ward work as a team?

Do you know who the ward manager or sister is on this ward?

Have you seen the ward manager or sister on this ward?

If yes, how often have you seen her/him? What do you see her/him doing?

If yes, do you think she/he supports the nurses in their work? Explain your answer

Appendix 6 Introduction to QI Tool software

Background

This document presents an overview of the development of the QI Tool software, developed for use in the CLECC study.

One of the aims of the study was to test outcome measures for use in a future definitive evaluation. At an early stage of proposal development, the team identified the existing QuIS as a promising candidate for primary outcome measure. QuIS is a time sampling tool that gives a measure of both the volume and quality of interactions (Dean et al.[42]). It is administered through researcher observations of health-care interactions in real-time. Previous uses of QuIS have involved manual data collection in which the QuIS rating and a small amount of contextual data are recorded using pen and paper. These data are then entered manually onto a database for analysis at a later date. In the research team's previous experience of using QuIS, these manual methods can take significant amounts of researcher time and mean that timely data analysis is not always possible. Translation errors between the manual and the database versions of the data also compromise validity.

When making the decision to pilot QuIS in the CLECC study, the team identified the opportunity to develop a software application to enable entry of QuIS ratings and a significantly larger amount of contextual data using a computer tablet in real-time during the observation for later wireless upload to a central database. We discussed these ideas with Dr Rudi Lutz, a freelance Android software developer, who confirmed their viability and worked with us to develop some early models for the work.

This software development and testing was funded by the NIHR. Dr Rudi Lutz developed the software through a consultancy agreement that included the specification that the University of Southampton retains 100% of the intellectual property for the software. David Pepper and Martin Chivers of the iSolutions team at the University of Southampton built the associated database for the software, which was funded by the university. The QI Tool was developed and then piloted successfully during the CLECC study 2015–16, and used successfully on two other projects, one of which was external to the university.

The Quality of Interactions Tool

The QI Tool is a tablet-based interface that enables users to enter data in real time for subsequent wireless upload to an encrypted central database. Data gathered include the quality, length and frequency of all interactions between participating patients and staff during the planned observation sessions.

It is best used to observe the care of one or more people who are an inpatient in a hospital setting. The terminology used is hospital-based but it may be possible to use the tool in other settings. It is designed for use by an observer who is located near to the patients under observation and who can then rate the quality of interactions with any staff that approach. It is not designed to rate the quality of interactions of individual members of staff, so local adaptations may be needed to achieve this purpose.

The QI Tool enables data to be gathered on up to six patients during a period of observation. These sessions can be planned in advance, can take place on an unscheduled basis or can be used for training purposes.

One patient is designated as the index patient so that, in the eventuality of a very busy observation session, observations of the index patient can take priority.

The QI Tool uses the QuIS as a framework for rating the quality of staff–patient interactions.[42] It also enables the collection of ward-based, patient-based and interaction-based contextual data. These data add valuable context to the quality of interactions that could help explain the ratings. Contextual data are gathered on the observation session (number of patients on the ward, staffing levels and skill mix), on the patients (age, gender, evidence of cognitive impairment) and on each interaction (including number of staff, staff type, and purpose of interaction).

Information gathered on each interaction includes:

- initial information – recording the patient's initial status
- QuIS category – recording the rated category for the quality of the interaction
- interaction content – recording the main purpose of the interaction
- initiated by – recording if the patient or staff member initiated the interaction
- one or two way – recording whether the interaction was one or two way
- number of staff – recording the number of staff involved in the interaction
- staff types – recording the job role of the member(s) of staff involved
- comments – recording any additional information about the interaction.

Once the observation session has been completed, data from planned and unplanned sessions (but not practice sessions) can be uploaded wirelessly to the central encrypted database for analysis. The database then generates a report in the form of a spreadsheet displaying all of the ward-based, patient-based and interaction-based data for each interaction. It also displays the date and time of the observation session, and the names of hospital, ward and observer.

Phase 2

The QI Tool was designed for use in the CLECC study. However, we recognised the potential for wider use within the NHS and other health services, and by other academic researchers nationally and internationally, and NIHR is keen for us to exploit these opportunities. In addition to the history of the original QuIS as a research instrument, it has also been used within the NHS as an improvement tool, enabling managers and frontline staff to directly measure the quality of staff–patient interactions. Given the high profile of compassionate care, our NHS partners in the CLECC study advise that there will be appetite for wider roll-out of the QI Tool. In addition to cutting out the translation errors mentioned earlier, there are two key advantages to the QI Tool over the original QuIS. One is the collection of a large number of highly relevant contextual data around the QuIS rating so, for instance, if the ward is short-staffed during the observation session, it would be possible to identify this during analysis. The second is the speed at which results can be made available, so for improvement purposes in particular, frontline staff and managers can view the results within a more meaningful timescale. For research teams and funders, there are clear cost and accuracy advantages to using the QI Tool over manual methods.

There are a number of technical reasons why the original version of the QI Tool was not suitable for wider roll-out and so phase 2 developed the software and associated database further to enable its use by a wider group of users.

Phase 2 included:

- removing the requirement that users have to have a University of Southampton user account
- rebranding (removing references to the original study)
- internationalisation to enable the QI Tool to be used by (say) a Swedish speaker, with all text displayed anywhere in the tool appearing in Swedish

The following tasks have also been completed:

- The preparation of documentation to support the QI Tool code – this will enable another software developer to take on development of the code if our current developer is not available.
- Software testing and adjustment to enable use of the QI Tool on a wider range of tablets than the one Android tablet it was developed to work on. It has been tested on the following:

  - Samsung Galaxy Tab S 10.5
  - Google Nexus 8.9
  - Samsung Galaxy s2 8
  - Lenovo Yoga 3 10.1.

Appendix 7 Quantity and quality of interaction between staff and older patients

Quantity and quality of interaction between staff and older patients in UK hospital wards: A descriptive study

Hannah Ruth Barker, M.Sc BN RN[a], Peter Griffiths, PhD BA RN FEANS FHEA[a], Ines Mesa-Eguiagaray, MSc BSc[b], Ruth Pickering, PhD MSc[b], Lisa Gould, MSc BSc[a], Jackie Bridges, PhD MSN B. Nurs (Hons) RN PG Cert[a,*]

[a] Faculty of Health Sciences, University of Southampton, United Kingdom
[b] Faculty of Medicine, University of Southampton, United Kingdom

ARTICLE INFO

ABSTRACT

Background: The quality of staff-patient interactions underpins the overall quality of patient experience and can affect other important outcomes. However no studies have been identified that comprehensively explore both the quality and quantity of interactions in general hospital settings.
Aims & objectives: To quantify and characterise the quality of staff-patient interactions and to identify factors associated with negative interaction ratings.
Setting: Data were gathered at two acute English NHS hospitals between March and April 2015. Six wards for adult patients participated including medicine for older people (n = 4), urology (n = 1) and orthopaedics (n = 1).
Methods: Eligible patients on participating wards were randomly selected for observation. Staff-patient interactions were observed using the Quality of Interactions Schedule. 120 h of care were observed with each 2 h observation session determined from a balanced random schedule (Monday-Friday, 08:00-22:00 h). Multilevel logistic regression models were used to determine factors associated with negative interactions.
Results: 1554 interactions involving 133 patients were observed. The median length of interaction was 36 s with a mean of 6 interactions per patient per hour. Seventy three percent of interactions were categorized as positive, 17% neutral and 10% negative. Forty percent of patients had at least one negative interaction (95% confidence interval 32% to 49%). Interactions initiated by the patient (adjusted Odds Ratio [OR] 5.30), one way communication (adjusted OR 10.70), involving two or more staff (adjusted OR 5.86 for 2 staff, 6.46 for 3+ staff), having a higher total number of interactions (adjusted OR 1.09 per unit increase), and specific types of interaction content were associated with increased odds of negative interaction (p < 0.05). In the full multivariable model there was no significant association with staff characteristics, skill mix or staffing levels. Patient agitation at the outset of interaction was associated with increased odds of negative interaction in a reduced model. There was no significant association with gender, age or cognitive impairment. There was substantially more variation at ward level (variance component 1.76) and observation session level (3.49) than at patient level (0.09).
Conclusion: These findings present a unique insight into the quality and quantity of staff-patient interactions in acute care. While a high proportion of interactions were positive, findings indicate that there is scope for improvement. Future research should focus on further exploring factors associated with negative interactions, such as workload and ward culture.
© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

* Corresponding author: Faculty of Health Sciences, University of Southampton, 02380597924, Building 67 Faculty of Health Sciences, University of Southampton, Highfield, SO17 1BJ, United Kingdom.
E-mail addresses: h.r.barker@soton.ac.uk (H.R. Barker), Peter.griffiths@soton.ac.uk (P. Griffiths), I.Mesa-Eguiagaray@soton.ac.uk (I. Mesa-Eguiagaray), rmp@soton.ac.uk (R. Pickering), l.j.gould@soton.ac.uk (L. Gould), Jackie.bridges@soton.ac.uk (J. Bridges).

1. Introduction

Considerable attention has been paid in recent years to the quantity and quality of interactions between staff and older patients in acute hospital settings. In the UK, retrospective analyses of care failures suggest that interactions between patients and

staff, particularly nurses, were of low quality and frequency, undermining quality of care and patient experiences (Care Quality Commission, 2015; Francis, 2013; Maben et al., 2012a). Findings from enquiries into these care failures have been accompanied by a crisis of public confidence in the ability of nurses in general to be compassionate (Maben and Griffiths, 2008; Prime Minister's Commission on the Future of Nursing and Midwifery in England, 2010; Report of the Willis Commission, 2012). A variety of reforms have resulted across the health service, such as changes to nursing education and recruitment (Department of Health, 2013a; Report of the Willis Commission, 2012). While the UK care failures have had particular prominence, evidence suggests that concerns about the frequency and quality of interactions between nursing staff and patients are shared internationally (Corbin, 2008; Kagan, 2014; Reader and Gillespie, 2013). However, surprisingly little is known about the quantity and quality of interactions between staff and patients outside of settings in which care failures have been identified and studied. This paper presents findings based on observations of staff-patient interactions in six hospital wards in two National Health Service (NHS) hospitals.

When people come into hospital, the quality of their interactions with staff is key to shaping experiences during their stay. For example, older people want nurses and others to use interactions to maintain identity ("see who I am"), to create community ("connect with me") and to share decision making ("involve me")' (Bridges et al., 2010). There may also be wider benefits to high quality interactions beyond patient experience. For instance, nurses aim to use their relationships with patients to provide tailored care, comfort and support, including supporting informed decision-making, and assessing responses to treatments, suggesting a clinically therapeutic potential to interactions (Bridges et al., 2013). Furthermore, the links that have recently been indicated between positive experience, patient safety and clinical effectiveness, suggest that quality of interactions may impact on a wider range of important outcomes such as adherence to recommended medication and treatments or technical quality of care delivery (Doyle et al., 2013).

Few studies offer a clear indication of how common the problems regarding staff-patient interactions are. Many that report on staff-patient interactions give retrospective global evaluations using questionnaires. For instance, The 2014 NHS inpatient survey involving 59,000 inpatients showed that 24% of inpatients could not find a member of the hospital staff to talk to about their worries and fears, and 13% did not get enough emotional support from hospital staff (Care Quality Commission, 2015). Measures such as the NHS survey offer a partial view because not everyone can participate, memories may be inaccurate and respondents cannot give a clear view of the frequency of negative experience.

Given the limitations of questionnaire methods, which tend to exclude some of the groups that may be most vulnerable to the impact of negative interactions such as those with cognitive impairment, observational methods may be a more appropriate method to measure the quantity and quality of interactions in general hospital care (Goldberg and Harwood, 2013). A review of the care of older people in 11 acute hospitals in Northern Ireland reported that 67% of 1836 interactions observed were rated as positive and 7% were rated negative (The Regulation and Quality Improvement Authority, 2015). While assessments of interaction quality were made using the validated Quality of Interactions Schedule (The Regulation and Quality Improvement Authority, 2015) the sampling method and context are unclear. A number of studies focusing on the nurse as the unit of analysis indicated that the amount of direct contact time was low, but no data were gathered on interaction quality (Westbrook et al., 2011). No studies have been identified that comprehensively explore both the quality and quantity of interactions with the patient as the unit of

analysis in general hospital settings, an important gap given the degree of attention this issue is attracting in the UK and beyond.

The study aims to address the important gap identified. The specific objectives were:

1. To identify the frequency and length of staff-patient interactions.
2. To characterise the quality of staff-patient interactions.
3. To identify associations between negative interactions and patient characteristics, staffing characteristics, interaction characteristics and observation session characteristics.

3. Methods

Data were collected as part of a feasibility study to develop and evaluate a compassionate care intervention for ward nursing teams (Bridges and Fuller, 2014). The data presented here were collected during the baseline phase of the study using a descriptive design.

3.1. Setting and sample

Data were collected in two acute NHS hospitals in England between March and April 2015. Managers of seven medical and surgical wards with high proportions of older in-patients were invited to include their ward in the study. Six wards participated: medicine for older people (n = 4), urology (n = 1) and orthopaedics (n = 1). Each ward had between 28 and 32 beds. We excluded patients identified by the nurse in charge as palliative, critically ill or reverse barrier nursed. All other patients were eligible for inclusion in the study.

Observations were undertaken in randomly generated time slots for ten two-hour sessions on each ward over a three week period (Monday–Friday, 08:00–22:00), there were 60 observation sessions in total. Observation sessions were balanced between wards and time of day. For each observation session, a random number generator was used to identify an index patient who was then approached and invited to take part in the study. If the patient agreed to take part, other patients in their vicinity were also approached and invited. If the index patient declined to take part, a new index patient was selected. This process continued until an index patient agreed to participate.

3.2. Data collection

The quality of interactions was measured using the Quality of Interactions Schedule (QuIS) (Dean et al., 1993), an observation-based tool that has been used in a number of studies in NHS acute care settings. Interactions between staff and patients are coded as positive social, positive care, neutral, negative protective and negative restrictive (Table 1). The QuIS has been shown to be sensitive to change in service quality (Algar et al., 2016; Brooker, 1995; Dean and Briggs, 1993; Health Advisory Service, 1998; Wewers et al., 1994). In long term residential settings QuIS has been shown to be reliable with kappa coefficients of above 0.75 (Dean et al., 1993). Concurrent validity has been demonstrated by the association of increased quantity and quality of interactions experienced by residents with improvements in ratings of residents' cognitive impairment, observed depression, and functional capacities (Dean and Briggs, 1993). QuIS was originally designed for long term settings, and so prior to the current study a protocol was developed for use by the research team to guide its application in acute settings, including a definition of what constituted the beginning and end of an interaction and how to decide between the different ratings (see Table 1) (McLean et al.,

Table 1
QuIS Category Definitions.

QuIS Category	QuIS Category Definitions	
	Dean et al. (1993)	McLean et al. (2014)
Positive Social	Interaction Principally involving 'good, constructive, beneficial' conversation and companionship	· Interactions, which may be expected to make the service user feel valued, cared about or respected as a person. This is achieved through: • Polite, friendly and respectful interactions in which any element is: Casual/informal and relating to 'everyday' social topics (e.g. family; sport; weather; TV programmes) or • Responding to concerns/interests/topics introduced by the service user
Positive Care	Interactions during the appropriate delivery of physical care.	Interactions, which may be expected to make the service user feel safe, secure, cared for or informed as a patient. This is achieved through polite, professional, respectful or good humoured interactions in which the topic is largely determined by staff and restricted to issues of care delivery (E.g. "your discharge"; "your wash"; "your medication"; "your surgery").
Neutral	Brief, indifferent interactions not meeting the definitions of the other categories.	Interactions and which would not be expected to impact on the feelings of the service user, which they would be indifferent to or which they may barely notice. Interactions with no positive or negative aspects
Negative Protective	Providing care, keeping safe or removing from danger, but in a restrictive manner, without explanation or reassurance: in a way, which disregards dignity or fails to demonstrate respect for the individual.	Interactions that may be expected to make the service user feel rushed, misunderstood, frustrated or poorly informed as a patient. Such interactions fail to fully maintain dignity or demonstrate respect due to the focus of staff on doing their 'work'. Staff may appear rushed or task orientated.
Negative Restrictive	Interactions that oppose or resist peoples' freedom of action without good reason, or which ignore them as a person.	Interactions which may be expected to leave the service user feel ignored, devalued or humiliated as a person. Such interactions may be rude, abusive or controlling and pay no regard to the perspective of the patient. Patients expressed needs/preferences are ignored or denied and staff may be authoritative, controlling, rude or angry.

2014). Inter-rater reliability testing was conducted on acute care patients recruited opportunistically. Kappa for QuIS rating was found to be 0.61, indicating good agreement.

Data gathered included the quality, length and frequency of all interactions between participating patients and staff during each observation session. Contextual data were also gathered on the session (number of patients on the ward, staffing levels and skill mix), on the patients (age, gender, evidence of cognitive impairment, agitation at outset of interaction) and on individual interactions (including number of staff, staff type, and content of interaction into seven types as detailed in Table 3). The platform used for data collection was the Quality of Interactions Tool (QI Tool), a tablet-based interface that enables users to enter data in real-time for subsequent wireless upload to an encrypted central database. Data were gathered through direct observation of interactions between patients and staff. Single researchers located themselves in a discrete location near enough to the patient(s) to be able to see and hear interactions. If curtains were drawn researchers stayed within hearing distance but did not enter in order to uphold the privacy and dignity of the patient. Five researchers were involved in collecting data. Each attended a seven-hour classroom training session and four hours of ward-based direct observation training.

3.3. Data analysis

Exploratory data analyses were performed to check the data and identify inconsistencies. Descriptive statistics for patient and interaction characteristics were calculated. Frequencies and percentages were computed to describe the type of interaction and QuIS ratings. Amongst patients with a full two hours of observation, the percentage with at least one negative (either protective or restrictive) was calculated and presented with a 95% confidence interval (CI).

A four level mixed-effects logistic regression model was fitted to investigate the effect of the predictive variables on the probability of an interaction being rated as negative (protective and restrictive combined). The individual interactions recorded between patients and staff were considered as the lowest level of the model. Patient, observation session and ward were included in

the model as random effects making up the higher three levels of the model. Predictive variables were included as fixed effects and presented as odds ratios (OR) with 95% CI. Models were fitted including each predictive variable as a fixed effect on its own (Model A), all predictive variables (Model B) and a selection of predictive variables (Model C). Terms were deemed statistically significant at the 5% level by virtue of the 95%CI around an OR including the value 1.00 or not. Models were estimated using the command xtmelogit in Stata 11.0 (StataCorp.2009.Stata Statistical Software: Release 11. College Station, TX: StataCorp LP). Agreement was assessed by calculating the Intra-class correlation (ICC) for the number of interactions observed. ICCs for agreement in the number of ratings recorded for each category, between the two observers, was also calculated. ICCs were calculated using the one way random model for a single measure in command reliability in SPSS.

3.4. Ethics

Procedures were in place to ensure that the principles of the Mental Capacity Act (2005)(Department for Constitutional Affairs, 2007) were adhered to. Personal consultees were consulted if an individual patient was assessed as lacking mental capacity to decide about whether or not to take part in the study. Any staff, including non-nursing staff, who interacted with recruited patients during the observation sessions were included, unless they declined to participate. Ethical approval for the study was granted by the Social Care Research Ethics Committee for England: study reference number 14/SC/1313.

4. Results

The care of 133 patients was observed over 120 h of planned observation. The intra-class correlation coefficient for the number of interactions recorded by paired observers was 0.94 (95% CI 0.67–0.99, P < 0.001). During this time there were 1554 interactions recorded.

Patients and interactions: The mean patient age was 83 years (range 18–101 years). Seventy nine per cent were female (n = 105). Thirty-one percent (n = 41) of patients had evidence of cognitive

Table 2
Patient and interaction characteristics for all observations. Values are number (%) unless stated otherwise.

Characteristic		
Patients		n = 133
Age in years	mean (SD)	83 (11)
	min to max	18 to 101
Gender	Male	28 (21%)
	Female	105 (79%)
Cognitive impairment	Yes	41 (31%)
	No	92 (69%)
Interactions/patient/observation session	mean (SD)	12 (6)
	min to max	1 to 41
Interactions/patient/hour	mean (SD)	6 (3)
	min to max	1 to 20
Number of index patients		60 (45%)
Observation session		n = 60
Number of patients per registered nurse + health care assistant	mean (SD)	3.5 (0.7)
	min to max	1 to 5.7
Number of registered nurse	mean (SD)	4.5 (1.2)
	min to max	2 to 7
Number of health care assistant	mean (SD)	3.8 (1.1)
	min to max	1–6
Skill mix	mean (SD)	0.5 (0.1)
	min to max	0 to 3–0.8
Interaction		n = 1554
Mood of the patient	Agitated	63 (4%)
	Non agitated	1491 (96%)
Visitors at start of interaction	Yes	100 (6%)
	No	1454 (94%)
Initiator of interaction	Staff	1262 (81%)
	Patient	292 (19%)
Type of communication	One way	232 (15%)
	Two way	1322 (85%)
Cognitive impairment	Yes	529 (34%)
	No	1025 (66%)
Length of interaction in seconds	mean (SD)	101 (190)
	median (LQ to UQ)	36 (14–106)
	min to max	0 to 2337
Number of staff involved	1	1420 (91%)
	2	104 (7%)
	3	30 (2%)

impairment. The proportion of positive interactions was the same for patients with/without cognitive impairment. The patient was not agitated at the outset of most interactions (n = 1491, 96%).

Length of interaction: There was a mean of six interactions per patient per hour (range 1–20) (Table 2). The mean length of interaction was 101 s with a median of 36 s (range 0–2337 s, or 0–39 min).

Interaction rating: Sixty percent of interactions (n = 927) were rated as positive care. In addition, 13% (n = 204) of interactions attracted the rating of positive social. Ten per cent (n = 156) of interactions were classified as negative, of which over half 6% (n = 97) were given the lowest rating of negative restrictive. Forty percent (47; 95% CI 32% to 49%) of patients with two hours of observation (n = 117) had at least one negative interaction.

Initiation of interactions: Eighty-one percent of interactions were initiated by staff rather than patients (n = 1262) and most were two-way interactions, that is the patient and staff member(s) were involved (n = 1322, 85%). Interactions typically occurred with no visitors present (n = 1454, 94%).

Type of interaction: Twenty-eight percent (n = 439) of interactions were classed as functional (including delivery of food and drink, bed-making, documentation and patient transfer), 25% (n = 383) were focused on communication, and 22% (n = 345) were focused on delivery of personal care (Table 3).

Staff and interactions: On average 4.5 registered nurses (RN) and 3.8 health care assistants (HCA) were present on the ward at

the start of an observation session, with a mean of 3.5 patients per RN + HCA (range 1–5.7) (Table 2). Registered nurses were involved

Table 3
Interaction Content: Total Frequencies and Percentages for All Observations in each Category.

Category	Subcategory	Frequency	Percentages%
Functional	Deliver food/drink	181	41
	Hotel services	95	22
	Bed making	89	20
	Documentation	70	16
	Transfer	4	1
Communication	Comfort	100	26
	Information	18	5
	Other conversation	265	69
Personal Care	Using the toilet	97	28
	Eating/drinking	86	25
	Moving/walking	71	21
	Personal hygiene	40	12
	Other Personal care	51	15
Planning	Nursing care	54	40
	Non-nursing care	46	34
	Medical treatment	35	26
Treatments	Medications	94	82
	Other treatments	20	18
Assessments	Vital signs	42	42
	Physical examination	11	11
	Other assessments	48	48
Other	Other	37	100

Table 4
Frequencies and Percentages of QuIS ratings and context for all observations.

	Positive Social	Positive Care	Neutral	Negative Protective	Negative Restrictive	Total
Wards						
A	46 (18%)	127 (49%)	44 (17%)	22 (9%)	18 (7%)	257
B	50 (15%)	196 (61%)	54 (17%)	8 (3%)	14 (4%)	322
C	21 (10%)	116 (55%)	35 (17%)	16 (8%)	22 (10%)	210
D	49 (16%)	205 (66%)	34 (11%)	5 (2%)	20 (6%)	313
E	22 (9%)	144 (57%)	58 (23%)	7 (3%)	20 (8%)	251
F	16 (8%)	139 (63%)	42 (21%)	1 (1%)	3 (2%)	201
Total	204 (13%)	927 (60%)	267 (17%)	59 (4%)	97 (6%)	1554
Individual Staff Group						
Registered nurse	85 (14%)	371(62%)	77 (13%)	26(4%)	37(6%)	596
Student nurse	2 (33%)	1(17%)	0	2(33%)	1(17%)	6
Health care assistant	89 (16%)	363 (64%)	67 (12%)	24 (4%)	29 (5%)	572
Doctor	4 (6%)	43 (63%)	14 (21%)	2 (3%)	5 (7%)	68
Allied health professional	12 (25%)	29 (59%)	5 (10%)	3 (6%)	0 (0%)	49
Other staff	39 (12%)	156 (47%)	105 (31%)	4 (1%)	32 (10%)	336
Cognitive Impairment						
No	133 (13%)	612 (60%)	193 (19%)	42 (4%)	45 (4%)	1025
Yes	71 (13%)	315 (60%)	74 (14%)	17 (3%)	52 (10%)	529
Total	204 (13%)	927 (60%)	267 (17%)	59 (4%)	97 (6%)	1554
Length of interaction						
Less than 5 s	14 (10%)	46 (33%)	74 (53%)	1 (1%)	5 (4%)	140
5 s to <1 min	64 (8%)	517 (61%)	172 (20%)	32 (4%)	69 (8%)	854
1 to <5 min	75 (17%)	300 (69%)	21 (5%)	19 (4%)	21 (5%)	436
5 to <10 min	33 (38%)	48 (55%)	0 (0%)	5 (6%)	2 (2%)	88
10 to <30 min	17 (50%)	15 (44%)	0 (0%)	2 (6%)	0 (0%)	34
30 min or more	1 (50%)	1 (50%)	0 (0%)	0 (0%)	0 (0%)	2
Total	204 (13%)	927 (60%)	267 (17%)	59 (4%)	97 (6%)	1554

in the largest proportion of interactions (n = 596, 38%) followed by HCAs (n = 572, 37%) (Table 4). Quality of interactions between patients and registered nurses and HCAs appeared to be similar; Seventy six percent of interactions involving registered nurses were rated as positive compared to 80% for HCAs (Table 4). Ninety-one percent of interactions involved just one member of staff (n = 1420).

Wards and session level: Wide variation between wards is evident, with negative restrictive ratings ranging from 3 to 18% of interactions on individual wards. Positive ratings (positive social and positive care combined) on individual wards ranged from 65 to 82%.

In the initial multilevel model (Model A) logistic model of predictors of a negative (combined protective and restrictive) interaction (Table 5) increasing age, agitation at outset of interaction, interactions initiated by the patient, one way communication, having two staff involved and a higher total number of interactions as well as some specific types of interaction content were associated with statistically significantly increased odds of negative interaction. For example, as patient age increases by 1 year, the odds of having a negative interaction will increase by 5%. The presence of a visitor reduced the odds of a negative interaction. There was a trend to marginally reduced odds for interactions involving a HCA in comparison to those involving registered nurses, amongst interactions with only one member of staff present. Although a higher number of patients per nurse was associated with increased odds of negative interaction this was not statistically significant. Similarly while a skill mix with a higher proportion of registered nurses was associated with reduced odds of negative interaction this was not statistically significant.

In Model B (Table 5) only interactions initiated by the patient (adjusted OR[1] 5.30), one way communication (adjusted OR 10.70 [5.64, 20.28]) having two or more staff involved (2 staff: adjusted OR 5.86 [2.33, 14.74]; 3+ staff adjusted OR 6.46 [1.45, 28.80]) and a higher total number of interactions (adjusted OR 1.09 per unit increase [1.03, 1.15]) plus specific types of interaction content were associated with increased odds of negative interaction. Interaction activity classified as communication (adjusted OR 2.56[1.18, 5.54]), personal care (adjusted OR 4.10 [1.84, 9.14]) or "other" (adjusted OR 8.36 [2.42, 28.91]) were more likely to be rated negatively. Associations with staffing levels or skill mix remained non-significant although the magnitude of the non-significant relationship for staffing was increased.

As a secondary analysis we estimated regression models including only patient, nurse and session level characteristics as predictors of quality of interaction (shown in Table 5 as Model C). In this model, patient agitation at the outset of interaction (adjusted OR 5.97 [2.23, 16.01]) was associated with increased odds of negative interaction and having visitors present was associated with reduced odds of negative interaction (adjusted OR 0.24 [0.07, 0.84]).

The estimate of variation from the multi-level model was lowest at the patient level (variance component 0.09) with substantially more variation at the ward level (variance component 1.76), and highest variation at the observation session level (variance component 3.49). In the reduced multi-level model, the variance components at the ward and observation session levels

---
[1] Adjusted odds ratio [95% confidence interval].

Table 5
Multilevel[*] logistic regression results: odds ratios (OR) of a negative (combined over protective and restrictive) interaction (n = 1554).

Variables	Model A (each predictive variable on its own) OR [95%CI]	Model B (for all predictive variables) Adjusted OR [95%CI]	Model C (for patient and staff contextual variables) adjusted OR [95% CI]
Patient's characteristics			
Age (per year increase)	**1.05 [1.01, 1.10]**	1.02 [0.98, 1.07]	1.05 [0.99, 1.10]
Male (vs female)	1.12 [0.27, 4.65]	2.65 [0.39, 18.01]	1.89 [0.42, 8.47]
With cognitive Impairment (vs without)	1.69 [0.73, 3.91]	1.15 [0.53, 2.50]	1.37 [0.59, 3.18]
Patient was agitated (vs non agitated)	**5.82 [2.21, 15.31]**	2.44 [0.82, 7.26]	**5.97 [2.23, 16.01]**
Type of staff (when only one member of staff is present)			
Registered nurse (n = 517)	1.00	1.00	1.00
Student nurse (n = 4)	9.51 [0.75, 120.62]	7.56 [0.46, 125.09]	10.99 [0.87, 138.95]
Health care assistant (n = 494)	0.91[0.54,1.52]	0.81 [0.44, 1.49]	0.90 [0.53, 1.53]
Doctor (n = 46)	0.97 [0.31, 3.05]	1.33 [0.37, 4.76]	1.07 [0.34, 3.43]
Allied health professional (n = 41)	0.89 [0.22, 3.59]	1.34 [0.26, 6.82]	1.03 [0.25, 4.27]
Other type of staff (n = 318)	1.20 [1.03, 4.98]	0.96 [0.45, 2.03]	1.30 [0.75, 2.27]
Interaction characteristics			
Length of interaction (in s)	1.00 [0.99, 1.00]	1.00 [0.99, 1.00]	
Visitors present (vs no visitors)	**0.17 [0.05, 0.59]**	0.31 [0.09, 1.09]	**0.24 [0.07, 0.84]**
Interaction initiated by patient (vs interaction initiated by staff)	**7.43 [4.67, 11.85]**	**5.30 [3.07, 9.16]**	
One way communication (vs two way communication)	**7.75 [4.64, 12.94]**	**10.70 [5.64, 20.28]**	
Interaction content			
Assessment	1.17 [0.34, 4.07]	1.48 [0.34, 6.45]	
Communication	**3.67 [1.93, 6.99]**	**2.56[1.18, 5.54]**	
Functional	1.00	1.00	
Personal care	**3.67 [1.90, 7.11]**	**4.10 [1.84, 9.14]**	
Planning	1.37 [0.56, 3.35]	1.67 [0.60, 4.71]	
Treatments	2.17 [0.81, 5.83]	2.76 [0.85, 8.94]	
Other	**20.93 [6.89, 63.59]**	**8.36 [2.42, 28.91]**	
Number of staff in interaction	1.00	1.00	1.00
1	**2.14 [1.02, 4.49]**	**5.86 [2.33, 14.74]**	**2.65[1.17, 5.96]**
2	2.76 [0.79, 9.62]	**6.46 [1.45, 28.80]**	2.26[0.60, 8.46]
3+			
Observation session characteristics			
Number of patients per nurse (per unit increase)	1.26 [0.60, 2.66]	1.60 [0.50, 5.09]	1.02 [0.41, 2.51]
Skill mix (per unit increase)	0.00 [0.00, 4.26]	0.00 [0.00, 45.03]	0.00 [0.00, 5.53]
Total number of interactions for that patient (per unit increase)	**1.11 [1.05, 1.18]**	**1.09 [1.03, 1.15]**	
Variance component (SE) [95% CI]			
Ward (n = 6)		1.76 (1.70) [0.27, 11.58]	1.02 (1.00) [0.15, 6.93]
Observation session level (n = 60)		3.49 (1.29) [1.68, 7.22]	1.60 (0.78) [0.61, 4.19]
Patient level (n = 133)		0.09 (0.28) [0.00, 35.59]	0.95 (0.47) [0.37, 2.49]

[*] All models include ward, observation session and patient as random effects. Bold results indicate statistically significant at 5% level.

were lower.

5. Discussion

Our findings provide a detailed description of the length, frequency and quality of staff-patient interactions in six UK hospital wards. Staff-patient interactions were frequent but tended to be relatively short. Interactions were mostly positive, but a small yet significant proportion were negative. Forty percent of patients had at least one negative interaction. A number of factors were independently associated with more negative interactions, including being patient-initiated, type of care and number of staff involved. Patient agitation at the outset of interaction was associated with increased odds of negative interaction and having visitors present was associated with reduced odds of negative interaction but only in Model C where no account was taken of the interaction content. Variability was highest at ward and session level, compared to patient level.

Findings on the length and quality of interactions are broadly similar to those reported in other studies, although frequency of interaction at a patient level has not been previously reported (The

Regulation and Quality Improvement Authority, 2015; Westbrook et al., 2011). Findings indicate that length of interaction is not associated with quality of interaction, suggesting that both positive and negative interactions can be accomplished in spite of short contact time. However more frequent interactions were associated with a more negative experience. More research is needed to better understand these patterns.

In contrast to other research, this study did not find that age or cognitive impairment increased the likelihood of a negative interaction (Bridges et al., 2013, 2010; Eriksson and Saveman, 2002; Fulmer et al., 2005). Previous studies have not controlled for multiple patient and staff factors and these findings clarify that age and cognitive impairment do not appear to be independently associated with interaction quality. In Model C, patient agitation at the outset of interaction was associated with higher odds of a negative interaction which adds to our understanding of which patients may be more vulnerable to negative interactions.

The finding that interactions focused on communication and personal care were associated with a higher likelihood of a negative interaction is noteworthy given that communication and personal care are aspects of nursing care in which positive

interactions are regarded as essential for establishing therapeutic relationships (Deparment of Health, 2010; Department of Health, 2012; Francis, 2013). Qualitative inquiry into interactions during these aspects of care would enable greater understanding of the association found here.

In Model C, the presence of visitors reduced the odds of a negative interaction, whereas having more than one member of staff present increased it. It might be anticipated that staff perform better when under the observation of others (Strauss, 2002), but these findings do not consistently support this claim. It may be that staff become so used to having an audience of other staff that any effect of observation is muted, but they become more conscious of their behavior when visitors are present. Conversation that occurs between staff when more than one is involved in an interaction may result in their being less focused on the patient. Results from the National Inpatient Survey in England suggest that staff talking in front of patients as if they were not there is a relatively common complaint, for example 19% of patients reported this. For doctors the percentage talking as if the patient wasn't there rose to 24% (Care Quality Commission, 2015). Alternatively it may be that with more members of staff involved, the opportunity for at least one of them to act negatively is increased. Visitors appear to have a protective effect, perhaps because staff become more conscious of their behavior when visitors are present. An alternative explanation is that visitors play an active part in improving the quality of an interaction. However, as this finding is only significant when the interaction content is not included in the model it may simply be that visitors are less likely to be present during encounters more likely to be associated with negative interactions, such as personal care.

The high variability associated with sessions and wards compared to patients suggests that common factors operating at these levels are leading to similar quality of interactions. This suggests that there may be factors associated with wards, for example ward cultures, that may influence the quality of interactions. The variability associated with an individual session suggest that factors present in the immediate situation, for example particular stressors or groups of staff, may have the biggest influence. The effects of staffing levels, skill mix or particular staff groups could operate at either of these levels, although we found no evidence of significant associations between these variables and the quality of interactions. However, given the large body of evidence linking staffing levels and skill mix to quality of care (Griffiths et al., 2014) the non-significant trends observed here warrant further exploration. Nonetheless these findings give no support to identify a specific staff group, and in particular health care assistants, as the source of negative interactions and these findings contradict previous largely anecdotal evidence (Cavendish, 2013). The importance of ward and organisational level cultural factors such as the nurse practice environment has also been established previously (Aiken et al., 2012; Maben et al., 2012b; Patterson et al., 2011) and warrants further investigation in future studies with a larger sample of wards.

5.1. Limitations

This study was undertaken on only six wards in two hospitals. The patient mix and ward characteristics suggest that the wards included in this study are typical of NHS general acute care wards (Department of Health, 2013b, 2010, 2000). In addition to the inclusion of a representative proportion of patients with cognitive impairment, the strengths of the study lie in the collection and analysis of data across a wide range of factors and on a large number of interactions, adding valuable understanding about the contextual influences on the quality of interactions between staff

and patients in acute hospital settings. However, caution must be exercised in generalising to other settings. The observational tool used is reliable and is supported by evidence of validity. Using observational methods means that patients who might otherwise be excluded are represented in these findings. However, further validation of the tool to explore its relationship with patient experience is warranted. We focused on the quality and quantity of interactions which occurred. Our study gives no indication of the appropriateness of the content or the extent to which necessary care was delivered. Other studies have shown that nurses frequently omit necessary elements of interpersonal care in the face of low staffing levels (Ball et al., 2014).

6. Conclusion

The findings from this study present a unique insight into the quality and quantity of interactions between staff and patients in acute hospital settings, and into the factors associated with the quality of an interaction. While a high proportion of interactions were positive, the findings also indicate that there is scope for further improving the quality of interactions: even over the short observation sessions a large number of patients experienced one or more interactions characterised as negative by an independent observer. It is unclear how these negative interactions affect overall patient experience of care or whether the overall number, or the presence of any negative interaction is most important. Further work is warranted to explore factors associated with a negative interaction, including workload and ward culture. Ward and session level factors are more important influences on the quality of interactions than patient factors, although our findings do not support attributing negative staff patient interactions to any particular staff group.

Acknowledgements

References

Aiken, L., Sermeus, W., Van den Heede, K., Sloanes, D., Busses, R., McKee, M., Kutney-Lee, A., 2012. Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. BMJ 344 doi:http://dx.doi.org/10.1136/bmj.e1717.

Algar, K., Woods, R.T., Windle, G., 2016. Measuring the quality of life and well-being of people with dementia: a review of observational measures. Dementia 15, 832–857.

Ball, J.E., Murrells, T., Rafferty, A.M., Morrow, E., Griffiths, P., 2014. Care left undone during nursing shifts: associations with workload and perceived quality of care. BMJ Qual. Saf. 23, 116–125. doi:http://dx.doi.org/10.1136/bmjqs-2012-001767.

Bridges, J., Fuller, A., 2014. Creating learning environments for compassionate care: a programme to promote compassionate care by health and social teams. Int. J. Older People Nurs. 10, 48–58. doi:http://dx.doi.org/10.1111/opn.12055.

Bridges, J., Flatley, M., Meyer, J., 2010. Older people's and relatives' experiences in acute care settings: systematic review and synthesis of qualitative studies. Int. J. Nurs. Stud. 47, 89–107. doi:http://dx.doi.org/10.1016/j.ijnurstu.2009.09.009.

Bridges, J., Nicholson, C., Maben, J., Pope, C., Flatley, M., Wilkinson, C., Meyer, J., Tziggili, M., 2013. Capacity for care: meta-ethnography of acute care nurses' experiences of the nurse-patient relationship. J. Adv. Nurs. 69, 760–772. doi: http://dx.doi.org/10.1111/jan.12050.

Brooker, D., 1995. Looking at them, looking at me: a review of observational studies into the quality of institutional care for elderly people with dementia. J. Ment. Heal. 4, 145–156.

Care Quality Commission, 2015. National NHS Patient Survey Programme National Results from the 2014 Inpatient Survey. Care Quality Commission and NHS.

Cavendish, C., 2013. The Cavendish Review: An Independent Review into Healthcare Assistants and Support Workers in the Nhs and Social Care Settings. (Retrieved from London: https://www.gov.uk/government/publications/review-of-healthcare-assistants-and-support-workers-in-nhs-and-social-care).

Corbin, J., 2008. Is caring a lost art in nursing? Int. J. Nurs. Stud. 45, 163–165. doi: http://dx.doi.org/10.1016/j.ijnurstu.2007.09.013.

Dean, R., Briggs, K., 1993. The domus philosophy: a prospective evaluation of two residential units for the elderly mentally ill. Int. J. Geriatr. Psychiatry 8, 807–817.

Dean, R., Proudfoot, R., Lindesay, J., 1993. The quality of interactions schedule (QUIS): development, reliability and use in the evaluation of two domus units. Int. J. Geriatr. Psychiatry 8, 819–826.

Department of Health, 2010. Essence of Care. Department of Health, London.

Department for Constitutional Affairs, 2007. Mental Capacity Act 2005 Code of Practice. The Stationery Office, London.

Department of Health, 2010. Quality Outcomes for People with Dementia: Building on the Work of the National Dementia Strategy. The Stationery Office, London.

Department of Health, 2012. Compassion in Practice. Department of Health, London.

Department of Health, 2013a. Delivering High Quality, Effective, Compassionate Care: Developing the Right People with the Right Skills and the Right Values: A Mandate from the Government to Health Education England: April 2013 to March 2015. Department of Health, London.

Department of Health, 2013b. In-patient Care Health Building Note 04-01: Adult In-patient Facilities. The Stationery Office, London.

Department of Health, 2000. Shaping the future NHS: Long term planning for hospitals and related services: Consultation document on the findings of the National Beds Inquiry. London : The Stationery Office.

Doyle, C., Lennox, L., Bell, D., 2013. A systematic review of evidence on the links between patient experience and clinical safety and effectiveness. BMJ Open 3, 1–18. doi:http://dx.doi.org/10.1136/bmjopen-2012-001570.

Eriksson, C., Saveman, B.-I., 2002. Nurses' experiences of abusive/non-abusive caring for demented patients in acute care settings. Scand. J. Caring Sci. 16, 79–85.

Francis, R., 2013. Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry. The Stationery Office, London.

Fulmer, T., Paveza, G., VandeWeerd, C., 2005. Dyadic vulnerability and risk profiling for elder neglect. Gerontologist 45, 525–534.

Goldberg, S.E., Harwood, R.H., 2013. Experience of general hospital care in older patients with cognitive impairment: are we measuring the most vulnerable patients' experience? BMJ Qual. Saf. 22, 977–980. doi:http://dx.doi.org/10.1136/bmjqs-2013-001961.

Griffiths, P., Ball, J., Drennan, J., James, L., Jones, J., Recio, A., Simon, M., 2014. The Association Between Patient Safety Outcomes and Nurse/healthcare Assistant Skill Mix and Staffing Levels and Factors That May Influence Staffing Requirements. University of Southampton.

Health Advisory Service, 1998. Standards of NHS Hospital Care for Older People Not Because They Are Old – an Independent Inquiry into the Care of Older People on Acute Wards in General Hospitals. Health Advisory Service 2000, London.

Kagan, S.H., 2014. Compassion. Geriatr. Nurs. (Minneap). 35, 69–70. doi:http://dx.doi.org/10.1016/j.gerinurse.2013.11.006.

Maben, J., Griffiths, P., 2008. Nurses in Society: Starting the Debate. King's College London, London.

Maben, J., Adams, M., Peccei, R., Murrells, T., Robert, G., 2012a. Poppets and parcels: the links between staff experience of work and acutely ill older peoples' experience of hospital care. Int. J. Older People Nurs. 7, 83–94. doi:http://dx.doi.org/10.1111/j.1748-3743.2012.00326.x.

Maben, J., Peccei, R., Adams, M., Robert, G., 2012b. Exploring the relationship between patients' experiences of care and the influence of staff motivation, affect and wellbeing. Natl. Inst. Health 1–444.

McLean, C., Griffiths, P., Bridges, J., 2014. A Feasibility Study to Develop the Quality of Interactions Schedule (QuIS) for Use as an Outcome Measure in Acute Hospital Care. University of Southampton, Southampton.

Patterson, M., Nolan, M., Rick, J., Brown, J., Adams, R., Musson, G., 2011. From metrics to meaning: culture change and quality of acute hospital care for older people. National Institute for Health Research .

Reader, T.W., Gillespie, A., 2013. Patient neglect in healthcare institutions: a systematic review and conceptual model. BMC Health Serv. Res. 13, 1–15. doi: http://dx.doi.org/10.1186/1472-6963-13-156.

Report of the Willis Commission, 2012. Quality with Compassion: the Future of Nursing Education. Royal College of Nursing, London.

Strauss, B., 2002. Social facilitation in motor tasks: a review of research and theory. Psychol. Sport Exerc. 3, 237–256. doi:http://dx.doi.org/10.1016/S1469-0292(01)00019-x.

The Regulation and Quality Improvement Authority, 2015. The Regulation and Quality Improvement Authority Review of the Care of Older People in Acute Hospitals Overview Report. The Regulation and Quality Improvement Authority, Belfast.

Westbrook, J., Duffield, C., Li, L., Creswick, N., 2011. How much time do nurses have for patients? a longitudinal study quantifying hospital nurses' patterns of task time distribution and interactions with health professionals. BMC Health Serv. Res. 11, 1–12.

Wewers, M., Muijen, M., Dean, R., Cooney, M., 1994. A nurse-delivered smoking cessation intervention among hospitalized postperative patients – influence of a smoking related diagnosis: a pilot study. Hear. Lung 23, 151–156.

Appendix 8 Questionnaire results

Nursing questionnaire

Values in tables are frequencies (%).

Baseline: how satisfied are you with your current job in this hospital?	Ward						Total (n = 91)
	A (n = 12)	B (n = 5)	C (n = 18)	D (n = 13)	E (n = 22)	F (n = 21)	
Very dissatisfied	0	0	2 (11)	1 (8)	0	1 (5)	4 (5)
A little dissatisfied	2 (18)	2 (40)	2 (11)	1(8)	2 (10)	3 (14)	12 (13)
Moderately satisfied	5 (46)	1 (20)	11 (61)	6 (46)	11 (52)	14 (67)	48 (54)
Very satisfied	4 (36)	2 (40)	3 (17)	5 (39)	8 (38)	3 (14)	25 (28)

Follow-up: how satisfied are you with your current job in this hospital?	Ward						Total (n = 87)
	A (n = 10)	B (n = 10)	C (n = 16)	D (n = 15)	E (n = 18)	F (n = 18)	
Very dissatisfied	0	0	1 (6)	0	1 (6)	2 (11)	4 (5)
A little dissatisfied	0	2 (20)	4 (25)	3 (20)	0	1 (6)	10 (12)
Moderately satisfied	5 (50)	5 (50)	5 (31)	9 (60)	11 (61)	10 (56)	45 (52)
Very satisfied	5 (50)	3 (30)	6 (31)	3 (20)	6 (33)	5 (28)	28 (32)

Baseline: how would you rate the work environment at your job in this hospital?	Ward						Total (n = 91)
	A (n = 12)	B (n = 5)	C (n = 18)	D (n = 13)	E (n = 22)	F (n = 21)	
Poor	0	0	3 (17)	2 (15)	2 (10)	1 (5)	8 (9)
Fair	4 (36)	2 (40)	7 (39)	2 (15)	1 (5)	6 (29)	22 (25)
Good	5 (46)	1 (20)	8 (44)	8 (62)	13 (62)	13 (62)	48 (54)
Excellent	2 (18)	2 (40)	0	1 (8)	5 (24)	1 (5)	11 (12)

Follow-up: how would you rate the work environment at your job in this hospital?	Ward						Total (n = 87)
	A (n = 10)	B (n = 10)	C (n = 16)	D (n = 15)	E (n = 18)	F (n = 18)	
Poor	0	1 (10)	1 (6)	3 (20)	1 (6)	1 (6)	7 (8)
Fair	1 (10)	3 (30)	7 (44)	6 (40)	5 (28)	4 (22)	26 (30)
Good	5 (50)	5 (50)	6 (38)	6 (40)	9 (50)	9 (50)	40 (46)
Excellent	4 (40)	1 (10)	2 (12)	0	3 (17)	4 (22)	14 (16)

Baseline: if possible, would you leave your current hospital within the next year as a result of job dissatisfaction?	Ward						Total
	A (n = 12)	B (n = 5)	C (n = 18)	D (n = 13)	E (n = 22)	F (n = 21)	(n = 91)
Yes	2 (18)	2 (40)	10 (56)	6 (46)	3 (14)	4 (20)	27 (31)
No	9 (82)	3 (60)	8 (44)	7 (54)	18 (86)	16 (80)	61 (69)

Follow-up: if possible, would you leave your current hospital within the next year as a result of job dissatisfaction?	Ward						Total
	A (n = 10)	B (n = 10)	C (n = 16)	D (n = 15)	E (n = 18)	F (n = 18)	(n = 87)
Yes	1 (10)	2 (20)	4 (25)	3 (21)	3 (17)	4 (22)	17 (20)
No	9 (90)	8 (80)	12 (75)	11 (79)	15 (83)	14 (78)	69 (80)

Baseline: if yes, what type of work would you seek?	Ward						Total
	A (n = 12)	B (n = 5)	C (n = 18)	D (n = 13)	E (n = 22)	F (n = 21)	(n = 91)
Nursing in another hospital	2 (100)	3 (100)	7 (78)	2 (33)	3 (75)	3 (50)	20 (61)
Nursing, but not in a hospital	0	0	0	1 (17)	0	2 (33)	3 (9)
Non-nursing	0	0	2 (22)	3 (50)	1 (25)	1 (17)	7 (21)

Follow-up: if yes, what type of work would you seek?	Ward						Total
	A (n = 10)	B (n = 10)	C (n = 16)	D (n = 15)	E (n = 18)	F (n = 18)	(n = 87)
Nursing in another hospital	0	3 (100)	3 (60)	3 (60)	4 (80)	2 (40)	15 (63)
Nursing, but not in a hospital	0	0	1 (20)	0	0	0	1 (4)
Non-nursing	1 (100)	0	1 (20)	2 (40)	1 (20)	3 (60)	8 (33)

Baseline: would you recommend your hospital to a nurse colleague as a good place to work?	Ward						Total
	A (n = 12)	B (n = 5)	C (n = 18)	D (n = 13)	E (n = 22)	F (n = 21)	(n = 91)
Definitely no	0	0	0	1 (8)	0	0	1 (1)
Probably no	1 (9)	1 (20)	3 (17)	1 (8)	2 (9)	4 (19)	12 (14)
Probably yes	6 (55)	2 (40)	8 (44)	4 (31)	13 (62)	11 (52)	44 (49)
Definitely yes	4 (36)	2 (40)	7 (39)	7 (54)	6 (29)	6 (29)	32 (36)

Follow-up: would you recommend your hospital to a nurse colleague as a good place to work?	Ward						Total
	A (n = 10)	B (n = 10)	C (n = 16)	D (n = 15)	E (n = 18)	F (n = 18)	(n = 87)
Definitely no	0	0	0	0	0	1 (6)	1 (1)
Probably no	0	1 (10)	3 (20)	1 (7)	0	0	5 (6)
Probably yes	3 (30)	6 (60)	9 (60)	9 (60)	11 (61)	9 (50)	47 (54)
Definitely yes	7 (70)	3 (30)	3 (20)	5 (33)	7 (39)	8 (44)	33 (38)

Baseline: would you recommend your hospital to your friends and family if they needed hospital care?	Ward						Total
	A (n = 12)	B (n = 5)	C (n = 18)	D (n = 13)	E (n = 22)	F (n = 21)	(n = 91)
Definitely no	0	0	0	2 (15)	0	0	2 (2)
Probably no	0	0	0	0	2 (10)	2 (9)	4 (4)
Probably yes	6 (54)	3 (60)	10 (56)	4 (31)	12 (57)	10 (48)	45 (51)
Definitely yes	5 (46)	2 (40)	8 (44)	7 (54)	7 (33)	9 (43)	38 (43)

Follow-up: would you recommend your hospital to your friends and family if they needed hospital care?	Ward						Total
	A (n = 10)	B (n = 10)	C (n = 16)	D (n = 15)	E (n = 18)	F (n = 18)	(n = 87)
Definitely no	0	0	1 (6)	0	0	1 (6)	2 (2)
Probably no	0	2 (20)	2 (13)	1 (7)	2 (11)	0	7 (8)
Probably yes	2 (20)	4 (40)	9 (56)	9 (60)	10 (56)	9 (50)	43 (49)
Definitely yes	8 (80)	4 (40)	4 (25)	5 (33)	6 (33)	8 (44)	35 (40)

Maslach Burnout Inventory

Subscale scores per ward: baseline	Ward						Total
	A (n = 12)	B (n = 5)	C (n = 18)	D (n = 13)	E (n = 22)	F (n = 21)	(n = 91)
Emotional exhaustion							
Mean (SD)	20 (11)	24 (11)	26 (11)	24 (13)	19 (11)	23 (13)	22 (12)
Minimum to maximum	0 to 35	9 to 37	3 to 43	12 to 52	2 to 38	3 to 47	0 to 52
Depersonalisation							
Mean (SD)	5 (3)	9 (7)	6 (5)	5 (7)	6 (5)	5 (5)	6 (5)
Minimum to maximum	0 to 11	0 to 17	0 to 15	0 to 20	0 to 14	0 to 17	0 to 20
Personal accomplishment							
Mean (SD)	39 (8)	38 (6)	38 (7)	41 (8)	38 (8)	37 (6)	38 (7)
Minimum to maximum	25 to 48	29 to 44	25 to 48	24 to 48	11 to 48	23 to 47	11 to 48

Subscale scores per ward: follow-up	Ward						Total (n = 87)
	A (n = 10)	B (n = 10)	C (n = 16)	D (n = 15)	E (n = 18)	F (n = 18)	
Emotional exhaustion							
Mean (SD)	14 (13)	21 (10)	22 (13)	24 (10)	19 (11)	19 (12)	20 (11)
Minimum to maximum	1 to 38	6 to 35	3 to 42	7 to 38	0 to 39	3 to 41	0 to 42
Depersonalisation							
Mean (SD)	6 (5)	6 (5)	6 (4)	5 (5)	5 (5)	4 (5)	5 (5)
Minimum to maximum	1 to 19	0 to 15	0 to 13	0 to 16	0 to 16	0 to 15	0 to 19
Personal accomplishment							
Mean (SD)	41 (8)	39 (9)	37 (6)	40 (6)	37 (7)	36 (11)	38 (8)
Minimum to maximum	21 to 47	22 to 47	24 to 47	28 to 48	24 to 46	14 to 48	14 to 48

Experiencing burnout: baseline	Ward, n (%)						Total (n = 91)
	A (n = 12)	B (n = 5)	C (n = 18)	D (n = 13)	E (n = 22)	F (n = 21)	
Emotional exhaustion	4 (33)	2 (40)	9 (50)	4 (31)	6 (27)	7 (35)	32 (36)
Depersonalisation	0	2 (40)	3 (17)	2 (15)	4 (18)	2 (10)	13 (14)
Personal accomplishment	2 (17)	1 (20)	4 (22)	2 (15)	3 (14%)	2 (10)	14 (16)

Experiencing burnout: follow-up	Ward, n (%)						Total (n = 87)
	A (n = 10)	B (n = 10)	C (n = 16)	D (n = 15)	E (n = 18)	F (n = 18)	
Emotional exhaustion	2 (20)	4 (40)	6 (38)	8 (53)	3 (17)	6 (33)	29 (33)
Depersonalisation	1 (10)	1 (10)	1 (6)	1 (7)	2 (11)	3 (17)	9 (11)
Personal accomplishment	1 (10)	2 (20)	3 (19)	2 (14)	4 (24)	3 (18)	15 (18)

Jefferson Scale of Empathy

Empathy score: at baseline	Ward						Total (n = 91)
	A (n = 12)	B (n = 5)	C (n = 18)	D (n = 13)	E (n = 22)	F (n = 21)	
Mean (SD)	113 (13)	112 (18)	115 (10)	120 (13)	115 (14)	107 (17)	113 (14)
Median	117	110	115	126	118	110	115
LQ, UQ	103, 121	98, 128	107, 122	109, 130	102, 127	91, 121	103, 124
Minimum to maximum	84 to 128	89 to 139	96 to 133	98 to 138	81 to 134	77 to 130	77 to 139

LQ, lower quartile; UQ, upper quartile.

| Empathy score: at follow-up | Ward | | | | | | Total |
	A (*n* = 10)	B (*n* = 10)	C (*n* = 16)	D (*n* = 15)	E (*n* = 18)	F (*n* = 18)	(*n* = 87)
Mean (SD)	108 (12)	113 (16)	113 (9)	114 (20)	112 (18)	113 (16)	112 (16)
Median	109	115	114	120	116	116	115
LQ, UQ	99, 117	99, 128	105, 120	102, 129	105, 126	104, 126	103, 124
Minimum to maximum	86 to 127	88 to 131	94 to 124	57 to 133	60 to 133	79 to 135	57 to 135

LQ, lower quartile; UQ, upper quartile.

Factors that enable Climate for Care

| Subscale scores per ward: at baseline | Ward | | | | | | Total |
	A (*n* = 12)	B (*n* = 5)	C (*n* = 18)	D (*n* = 13)	E (*n* = 22)	F (*n* = 21)	(*n* = 91)
Leading by example							
Mean (SD)	44 (7)	46 (5)	36 (9)	36 (11)	48 (6)	42 (9)	42 (9)
Minimum to maximum	34 to 52	39 to 52	21 to 50	18 to 53	38 to 55	22 to 55	18 to 55
Support from the top							
Mean (SD)	27 (4)	27 (6)	26 (4)	28 (5)	28 (6)	27 (6)	27 (5)
Minimum to maximum	20 to 34	22 to 35	18 to 32	22 to 39	18 to 40	14 to 40	14 to 40

| Subscale scores per ward: at follow-up | Ward | | | | | | Total |
	A (*n* = 10)	B (*n* = 10)	C (*n* = 16)	D (*n* = 15)	E (*n* = 18)	F (*n* = 18)	(*n* = 87)
Leading by example							
Mean (SD)	48 (7)	42 (9)	41 (10)	36 (11)	43 (7)	45 (11)	42 (10)
Minimum to maximum	35 to 55	22 to 55	23 to 55	18 to 53	31 to 54	17 to 55	17 to 55
Support from the top							
Mean (SD)	30 (5)	31 (5)	27 (5)	28 (5)	29 (5)	30 (6)	29 (5)
Minimum to maximum	22 to 36	24 to 40	13 to 34	22 to 39	19 to 39	21 to 40	13 to 40

Climate for Care

| Subscale scores per ward: at baseline | Ward | | | | | | Total |
	A (*n* = 12)	B (*n* = 5)	C (*n* = 18)	D (*n* = 13)	E (*n* = 22)	F (*n* = 21)	(*n* = 91)
Shared philosophy of care (5–25)							
Mean (SD)	21 (3)	22 (2)	22 (3)	21 (3)	20 (2)	19 (3)	20 (3)
Minimum to maximum	16–25	20–25	16–25	14–25	14–25	12–25	12–25
Having resources (3–15)							
Mean (SD)	10 (2)	10 (2)	9 (3)	8 (3)	8 (2)	10 (2)	9 (3)
Minimum to maximum	4–13	7–12	3–15	3–14	3–11	5–15	3–15

Subscale scores per ward: at baseline	Ward						Total (n = 91)
	A (n = 12)	B (n = 5)	C (n = 18)	D (n = 13)	E (n = 22)	F (n = 21)	
Supporting each other (6–30)							
Mean (SD)	23 (6)	22 (8)	22 (5)	21 (5)	24 (4)	23 (4)	23 (5)
Minimum to maximum	12–30	8–29	8–30	13–30	17–30	16–30	8–30
Feeling safe (4–20)							
Mean (SD)	15 (4)	13 (5)	15 (3)	14 (3)	15 (3)	14 (3)	15 (3)
Minimum to maximum	6–19	4–17	7–20	8–20	10–20	5–20	4–20
Improving practice (6–30)							
Mean (SD)	22 (5)	19 (9)	20 (5)	20 (4)	22 (4)	21 (5)	21 (5)
Minimum to maximum	14–30	6–28	9–27	12–25	14–30	9–30	6–30
Having a say (6–30)							
Mean (SD)	22 (5)	21 (5)	22 (5)	20 (3)	22 (4)	20 (6)	21 (5)
Minimum to maximum	15–29	15–28	10–30	14–26	16–30	7–30	7–30
Developing our skills (3–15)							
Mean (SD)	10 (2)	11 (2)	10 (2)	10 (3)	11 (2)	11 (3)	11 (2)
Minimum to maximum	6–13	7–13	6–14	3–15	6–15	3–15	3–15
Too much to do (4–20)							
Mean (SD)	14 (2)	13 (3)	15 (3)	14 (3)	14 (3)	14 (3)	14 (3)
Minimum to maximum	12–19	9–16	9–20	11–20	6–20	11–20	6–20
MDT working (2–10)							
Mean (SD)	8 (1)	7 (1)	8 (2)	8 (1)	8 (2)	8 (1)	8 (1)
Minimum to maximum	6–10	6–9	4–10	6–10	6–10	6–10	4–10

Subscale scores per ward: at follow-up	Ward						Total (n = 87)
	A (n = 10)	B (n = 10)	C (n = 16)	D (n = 15)	E (n = 18)	F (n = 18)	
Shared philosophy of care (5–25)							
Mean (SD)	23 (3)	20 (3)	19 (3)	20 (2)	20 (3)	20 (4)	20 (3)
Minimum to maximum	17–25	15–25	13–25	15–25	15–25	8–25	8–25
Having resources (3–15)							
Mean (SD)	9 (3)	10 (3)	9 (3)	8 (2)	10 (2)	11 (3)	10 (3)
Minimum to maximum	3–13	6–14	3–15	4–10	7–15	5–15	3–15
Supporting each other (6–30)							
Mean (SD)	26 (5)	24 (4)	23 (5)	23 (3)	24 (4)	24 (4)	24 (4)
Minimum to maximum	14–30	17–29	14–30	17–30	16–30	16–30	14–30
Feeling safe (4–20)							
Mean (SD)	17 (4)	16 (3)	15 (4)	13 (2)	15 (3)	15 (4)	15 (3)
Minimum to maximum	9–20	12–20	9–20	8–16	8–20	5–20	5–20

Subscale scores per ward: at follow-up	Ward						Total (*n* = 87)
	A (*n* = 10)	B (*n* = 10)	C (*n* = 16)	D (*n* = 15)	E (*n* = 18)	F (*n* = 18)	
Improving practice (6–30)							
Mean (SD)	24 (4)	21 (5)	19 (4)	19 (5)	21 (4)	23 (5)	21 (5)
Minimum to maximum	14–28	12–29	12–24	12–30	14–29	12–30	12–30
Having a say (6–30)							
Mean (SD)	26 (5)	21 (5)	21 (4)	19 (4)	22 (3)	24 (4)	22 (5)
Minimum to maximum	18–30	10–30	11–27	12–30	16–26	18–30	10–30
Developing our skills (3–15)							
Mean (SD)	12 (3)	12 (2)	10 (2)	10 (2)	11 (2)	11 (3)	11 (3)
Minimum to maximum	7–15	7–15	6–15	6–15	6–15	4–15	4–15
Too much to do (4–20)							
Mean (SD)	12 (3)	11 (3)	14 (3)	14 (3)	13 (3)	12 (3)	13 (3)
Minimum to maximum	7–16	6–16	10–19	10–20	7–20	4–19	4–20
MDT working (2–10)							
Mean (SD)	9 (1)	8 (2)	7 (2)	8 (1)	8 (2)	7 (2)	8 (2)
Minimum to maximum	6–10	6–10	3–10	6–10	4–10	4–10	3–10

Patients' questionnaire

Mean (SD) scores for each subscale (per ward and total).

Patient Evaluation of Emotional Care during Hospitalisation

Subscale: at baseline	Ward, mean score (SD)						Total (*n* = 168)
	A (*n* = 25)	B (*n* = 19)	C (*n* = 20)	D (*n* = 31)	E (*n* = 30)	F (*n* = 43)	
Security	2.53 (0.45)	2.29 (0.58)	2.21 (0.59)	2.30 (0.69)	2.58 (0.42)	2.43 (0.47)	2.41 (0.54)
Knowledge	2.42 (0.47)	1.78 (1.02)	1.96 (0.75)	2.05 (0.97)	2.35 (0.69)	2.45 (0.67)	2.22 (0.79)
Personal value	2.42 (0.54)	2.08 (0.67)	2.18 (0.65)	2.25 (0.63)	2.53 (0.40)	2.43 (0.54)	2.35 (0.57)
Connection	1.71 (0.61)	1.43 (0.79)	1.70 (0.85)	1.63 (0.83)	1.89 (0.71)	1.56 (0.84)	1.66 (0.78)
Total PEECH	51.5 (9.5)	44.9 (13.9)	44.5 (13.9)	45.8 (13.9)	52.9 (7.8)	50.2 (10.7)	48.9 (11.7)

Subscale: at follow-up	Ward, mean score (SD)						Total (*n* = 186)
	A (*n* = 32)	B (*n* = 33)	C (*n* = 31)	D (*n* = 29)	E (*n* = 29)	F (*n* = 32)	
Security	2.37 (0.57)	2.59 (0.41)	2.48 (0.52)	2.45 (0.54)	2.54 (0.48)	2.44 (0.46)	2.47 (0.50)
Knowledge	2.23 (0.85)	2.40 (0.70)	2.33 (0.67)	2.08 (0.99)	2.05 (0.97)	2.21 (0.66)	2.22 (0.82)
Personal value	2.32 (0.68)	2.63 (0.42)	2.25 (0.66)	2.43 (0.46)	2.32 (0.65)	2.36 (0.50)	2.39 (0.57)
Connection	1.74 (0.85)	2.11 (0.82)	1.82 (0.55)	1.64 (0.85)	1.75 (0.72)	1.63 (0.68)	1.78 (0.77)
Total PEECH	48.6 (13.5)	53.8 (8.5)	47.6 (10.8)	50.1 (10.2)	49.6 (12.3)	49.2 (9.1)	49.9 (10.8)

Visitors' questionnaire

Carer experiences of care

Subscale scores per ward: at baseline	Ward						Total (*n* = 89)
	A (*n* = 19)	B (*n* = 9)	C (*n* = 16)	D (*n* = 20)	E (*n* = 17)	F (*n* = 19)	
Giving my relative the best							
Mean (SD)	22.0 (4.3)	21.7 (6.7)	22.5 (3.9)	24.1 (4.6)	23.6 (4.2)	22.3 (4.3)	22.9 (4.5)
Minimum to maximum	16 to 28	6 to 30	15 to 28	14 to 30	16 to 30	13 to 30	6 to 30
Could do better							
Mean (SD)	6.1 (2.5)	6.0 (3.3)	5.6 (1.7)	5.7 (2.3)	5.6 (2.1)	5.6 (2.1)	5.7 (2.2)
Minimum to maximum	3 to 9	3 to 13	3 to 9	3 to 10	3 to 10	3 to 10	3 to 13
Feeling significant							
Mean (SD)	34.6 (9.6)	34.9 (10)	31.7 (5.1)	36.2 (6.5)	35.2 (6.3)	33.4 (7.6)	34.3 (7.2)
Minimum to maximum	22 to 46	13 to 48	23 to 41	18 to 46	20 to 44	12 to 45	12 to 48

Subscale scores per ward: at follow-up	Ward						Total (*n* = 87)
	A (*n* = 12)	B (*n* = 16)	C (*n* = 15)	D (*n* = 12)	E (*n* = 14)	F (*n* = 18)	
Giving my relative the best							
Mean (SD)	22.3 (3.7)	21.6 (4.9)	20.6 (4.9)	22.8 (4.9)	23.2 (3.9)	22.2 (3.5)	22.1 (4.3)
Minimum to maximum	17 to 27	12 to 28	14 to 30	14 to 30	16 to 30	17 to 28	12 to 30
Could do better							
Mean (SD)	5.9 (2.3)	6.1 (2.3)	6.0 (2.0)	6.1 (1.8)	6.2 (1.5)	5.9 (2.0)	6.0 (2.0)
Minimum to maximum	3 to 10	3 to 11	3 to 10	3 to 9	3 to 8	3 to 10	3 to 11
Feeling significant							
Mean (SD)	30.1 (4.4)	31.9 (8.0)	30.2 (6.9)	33.2 (9.0)	34.9 (14.6)	30.8 (7.1)	31.7 (6.9)
Minimum to maximum	23 to 36	18 to 46	17 to 39	19 to 47	27 to 43	17 to 41	17 to 47

Matron's assessment of quality of care

Subscale scores per ward	Ward						Total
	A	B	C	D	E	F	
Meeting patient's needs score							
Baseline	23	24	22	21	22	12	20.7 (4.4)
Follow-up	23	22	24	25	27	24	24.2 (2.4)
Looking to improve score							
Baseline	15	17	14	15	16	10	14.5 (2.4)
Follow-up	16	14	13	16	16	18	15.5 (1.8)

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library