UNIVERSITY of York

This is a repository copy of *Patient involvement in improving the evidence base on mental health inpatient care: the PERCEIVE programme*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/158417/

Version: Published Version

Article:
Wykes, Til HM, Csipke, Emese, Rose, Diana Susan et al. (5 more authors) (2018) Patient involvement in improving the evidence base on mental health inpatient care: the PERCEIVE programme. Programme Grants for Applied Research. pp. 1-211. ISSN 2050-4322

https://doi.org/10.3310/pgfar06070

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

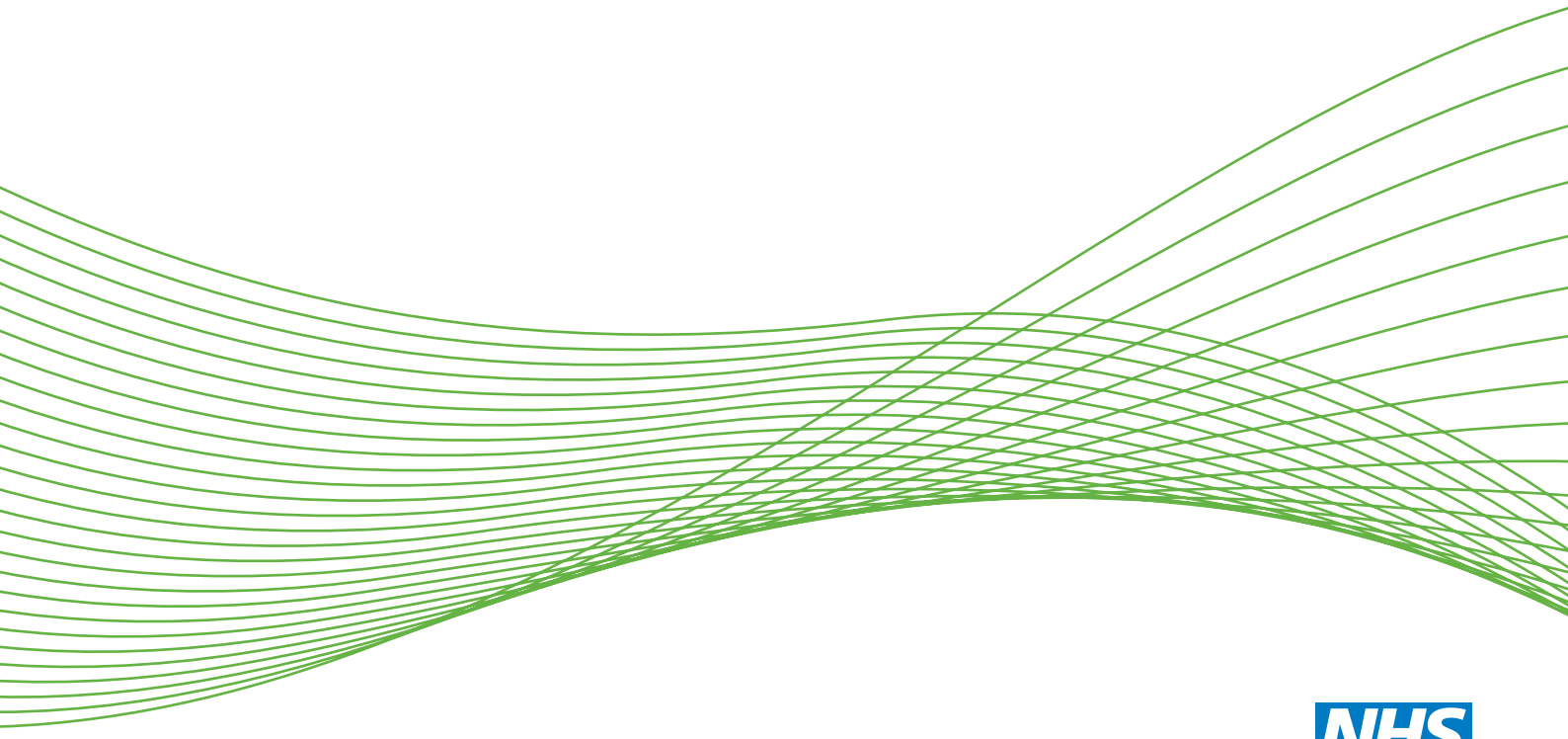
eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

PROGRAMME GRANTS FOR APPLIED RESEARCH

VOLUME 6 ISSUE 7 DECEMBER 2018
ISSN 2050-4322

Patient involvement in improving the evidence base on mental health inpatient care: the PERCEIVE programme

Til Wykes, Emese Csipke, Diana Rose, Thomas Craig, Paul McCrone, Paul Williams, Leonardo Koeser and Stephen Nash



NHS
National Institute for Health Research

DOI 10.3310/pgfar06070

Patient involvement in improving the evidence base on mental health inpatient care: the PERCEIVE programme

Til Wykes,[1]* Emese Csipke,[1] Diana Rose,[2] Thomas Craig,[2] Paul McCrone,[2] Paul Williams,[2] Leonardo Koeser[2] and Stephen Nash[3]

[1]Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[2]Health Services and Population Research, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[3]Department of Biostatistics, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

*Corresponding author

Declared competing interests of authors: none

Published December 2018
DOI: 10.3310/pgfar06070

This report should be referenced as follows:

Wykes T, Csipke E, Rose D, Craig T, McCrone P, Williams P, *et al*. Patient involvement in improving the evidence base on mental health inpatient care: the PERCEIVE programme. *Programme Grants Appl Res* 2018;**6**(7).

Programme Grants for Applied Research

ISSN 2050-4322 (Print)

ISSN 2050-4330 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full PGfAR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/pgfar. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Programme Grants for Applied Research* journal

Reports are published in *Programme Grants for Applied Research* (PGfAR) if (1) they have resulted from work for the PGfAR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Programme Grants for Applied Research programme

The Programme Grants for Applied Research (PGfAR) programme, part of the National Institute for Health Research (NIHR), was set up in 2006 to produce independent research findings that will have practical application for the benefit of patients and the NHS in the relatively near future. The Programme is managed by the NIHR Central Commissioning Facility (CCF) with strategic input from the Programme Director.

The programme is a national response mode funding scheme that aims to provide evidence to improve health outcomes in England through promotion of health, prevention of ill health, and optimal disease management (including safety and quality), with particular emphasis on conditions causing significant disease burden.

For more information about the PGfAR programme please visit the website: http://www.nihr.ac.uk/funding/programme-grants-for-applied-research.htm

This report

The research reported in this issue of the journal was funded by PGfAR as project number RP-PG-0606-1050. The contractual start date was in August 2007. The final report began editorial review in February 2014 and was accepted for publication in January 2017. As the funder, the PGfAR programme agreed the research questions and study designs in advance with the investigators. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The PGfAR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, CCF, NETSCC, PGfAR or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the PGfAR programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

NIHR Journals Library Editor-in-Chief

Professor Ken Stein Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals)

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Consultant Advisor, Wessex Institute, University of Southampton, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk

Abstract

Patient involvement in improving the evidence base on mental health inpatient care: the PERCEIVE programme

Til Wykes,[1]* Emese Csipke,[1] Diana Rose,[2] Thomas Craig,[2] Paul McCrone,[2] Paul Williams,[2] Leonardo Koeser[2] and Stephen Nash[3]

[1]Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[2]Health Services and Population Research, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
[3]Department of Biostatistics, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

*Corresponding author til.wykes@kcl.ac.uk

Background: Despite the movement towards care in the community, 40% of the NHS budget on mental health care is still attributed to inpatient services. However, long before the Francis Report highlighted grave shortcomings in inpatient care, there were reports by service user groups on the poor quality of these services in mental health. The programme provides a particular focus on the inclusion of the patient's perspective in the development and evaluation of evidence.

Objectives: To understand how changes to inpatient care affect the perceptions of the ward by service users and staff by using stakeholder participatory methods.

Design: The programme consisted of four work packages (WPs). (1) Lasting Improvements for Acute Inpatient SEttings (LIAISE): using participatory methods we developed two new scales [Views On Therapeutic Environment (VOTE) for staff and Views On Inpatient CarE (VOICE) for service users]. (2) Client Services Receipt Inventory – Inpatient (CITRINE): working with nurses and service users we developed a health economic measure of the amount of contact service users have with staff. The self-report measure records interactions with staff as well as the number of therapeutic activities attended. (3) Delivering Opportunities for Recovery (DOORWAYS): a stepped-wedge randomised controlled trial to test if training ward nurses to deliver therapeutic group activities would improve the perception of the ward by service users and staff. A total of 16 wards were progressively randomised and we compared the VOICE, VOTE and CITRINE measures before and after the intervention. A total of 1108 service users and 539 staff participated in this trial. (4) Bringing Emergency TreatmenT to Early Resolution (BETTER PATHWAYS) was an observational study comparing two service systems. The first was a 'triage' system in which service users were admitted to the triage ward and then either transferred to their locality wards or discharged back into the community within 7 days. The second system was routine care. We collected data from 454 service users and 284 nurses on their perceptions of the wards.

Main outcome measures: The main outcomes for the DOORWAYS and BETTER project were service user and staff perceptions of the ward (VOICE and VOTE, respectively) and the health economic measure was CITRINE. All were developed in WPs 1 and 2.

Results: We developed reliable and valid measures of (1) the perceptions of inpatient care from the perspectives of service users and nurses (VOICE and VOTE) and (2) costs of interactions that were valued by service users (CITRINE). In the DOORWAYS project, after adjusting for legal status, we found weak evidence

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

for benefit (standardised effect of –0.18, 95% CI 0.38 improvement to 0.01 deterioration; $p = 0.062$). There was only a significant benefit for involuntary patients following the staff training (N582, standardised effect of –0.35, 95% CI –0.57 to –0.12; $p = 0.002$; interaction p-value 0.006). VOTE scores did not change over time (standardised effect size of 0.04, 95% CI –0.09 to 0.18; $p = 0.54$). We found no evidence of an improvement in cost-effectiveness (estimated effect of £33, 95% CI –£91 to £146; $p = 0.602$), but resource allocation did change towards patient-perceived meaningful contacts by an average of £12 (95% CI –£76 to £98; $p = 0{\cdot}774$). There were no significant differences between the triage and routine models of admission in terms of better perceptions by service users (estimated effect 0.77-point improvement in VOICE score on the triage ward; $p = 0.68$) or nurses (estimated effect of 1.68-point deterioration in VOTE on the triage ward; $p = 0.38$) or in terms of the cost of the length of care provided (£391 higher on triage; $p = 0.77$).

Strengths and limitations: We have developed measures using methods involving both service users and staff from mental health services. The measures were developed specifically for acute inpatient services and, therefore, cannot be assumed to be useful for other services. For instance, extensions of the measures are under construction for use in mother and baby units. The strength of the BETTER PATHWAYS and DOORWAYS projects is the large-scale data collection. However, we were testing specific services based in inner city areas and stretching to inner urban areas. It may be that different effects would be found in more rural communities or in different types of inpatient care.

Future work: Our database will be used to develop an understanding of the mediating and moderating factors for improving care quality.

Trial registration: Current Controlled Trials ISRCTN06545047.

Funding: This project was funded by the NIHR Programme Grants for Applied Research programme and will be published in full in *Programme Grants for Applied Research*; Vol. 6, No. 7. See the NIHR Journals Library website for further project information.

Contents

List of tables xi

List of figures xiii

Glossary xv

List of abbreviations xvii

Plain English summary xix

Scientific summary xxi

Chapter 1 Overview of the programme 1
Background 1
Pressure on the number of inpatient beds available 1
Therapeutic activities and positive interactions on wards 2
Stakeholder involvement 3
Health economic considerations 3
Work package 1 3
Work package 2 4
Work package 3 4
Work package 4 4
Objectives 5

Chapter 2 Work package 1: LIAISE 7
Participatory research methods 7
Background 7
Methods 9
Phase 1: mapping out the dimensions of inpatient care 10
Phase 2: measure generation 10
Psychometric assessment 10
Samples 10
Phase 3: feasibility and acceptability 11
Phase 4: psychometric assessment 11
Results 12
Phase 1: mapping out the dimensions of inpatient care 12
Phase 2: measure generation 13
Psychometric assessment 17
Phase 4: psychometric assessment 18
Discussion 23
Can VOICE and VOTE distinguish differences in views? 23
Are VOICE and VOTE different from other measures? 24
Further developments 24
Strengths and limitations 24
Conclusion 25
Clinical implications 25
Research implications 25

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 3 Work package 2: CITRINE **27**
Systematic review of studies measuring staff–patient contact time on psychiatric inpatient wards 27
Methods 28
Results 28
Discussion 33
Systematic review of economic evaluations of therapeutic activities on psychiatric inpatient wards 35
Methods 35
Results 35
Discussion 35
Modelling the cost-effectiveness of inpatient interventions: social cognition training case study 35
Methods 36
Results 37
Discussion 37
Development of a measure of service use 39
Methods 40
Results 41
Discussion 41
Analysis of baseline costs 43
Methods 43
Results 43
Discussion 46
Conclusion 47

Chapter 4 Work package 3: DOORWAYS **49**
Introduction 49
Methods 51
Procedures 51
The setting 52
Participants 52
Main outcomes 52
Secondary outcomes 53
Background information 54
Intervention 54
Statistical methods 54
Results 56
Baseline data 56
Outcomes and estimation 59
Adverse events 67
Discussion 68
Summary 68
Impact on wards 68
Impact on service users 69
Impact on nursing staff 70
Costs and cost-effectiveness 71
Strengths and limitations of this study 71
Research implications 71
Clinical implications 72
Conclusions 72

Chapter 5 Work package 4: BETTER PATHWAYS 73
Introduction 73
Method 74
The setting 74
Study 1: is a triage system more efficient than traditional care? 75
Study 2: service user and staff perceptions of care 76
Results 78
Study 1: is the triage system more efficient than traditional care? 78
Study 2: the experience of service users and staff 81
Discussion 93
Strengths and limitations 93
Clinical implications 96
Implications for future research 96
Conclusion 97

Chapter 6 Discussion and conclusions from the PERCEIVE programme 99
Summary of results 99
Patient and public involvement in the PERCEIVE study 100
Service user involvement in the governance structure 101
Design of the PERCEIVE programme 101
Service user involvement in measure development 101
How does VOICE perform? 102
Extending service user involvement to other stakeholders 102
CITRINE measure development 102
Service user involvement in data collection 102
Service user impact on analyses 103
Service user impact on dissemination 103
Service user impact on protocol developments (add-on studies) 103
Strengths and limitations 104
Research implications 105
Clinical and organisational implications for mental health inpatient units 105
Conclusions 106

Acknowledgements 107

References 109

Appendix 1 Design in mind 123

Appendix 2 Perinatal VOICE 127

Appendix 3 VOCALISE: barriers to change 131

Appendix 4 Social cognition and interaction training 137

Appendix 5 The stepped-wedge design 141

Appendix 6 Costings 143

Appendix 7 VOICE 145

Appendix 8 VOTE 151

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 9 CITRINE 159

Appendix 10 VOCALISE 163

Appendix 11 PERCEIVE analysis plans 169

Appendix 12 CONSORT 2010 checklist of information to include when reporting a randomised trial 179

Appendix 13 Publications directly arising from the programme grant 181

List of tables

TABLE 1 Demographic data of service users in WP1 13

TABLE 2 Samples of individuals in the VOTE development in WP1 14

TABLE 3 Factor analysis of VOICE in WP1 19

TABLE 4 Factor analysis of VOTE in WP1 20

TABLE 5 Differences in mean VOICE scores by demographic and clinical group in WP1 21

TABLE 6 Mean VOTE scores by demographic and clinical group in WP1 22

TABLE 7 VOTE scores adjusted for country of birth (WP1) 22

TABLE 8 Methods of reviewed time and motion studies (WP2) 29

TABLE 9 Mean number and cost (£) of activities and nursing contacts reported by service users and from OTs/case notes (WP2) 41

TABLE 10 Sample characteristics: binary variables ($n = 334$) (WP2) 44

TABLE 11 Sample characteristics: continuous variables ($n = 334$) (WP2) 44

TABLE 12 Therapeutic activities (WP2) 44

TABLE 13 One-to-one contacts (WP2) 45

TABLE 14 Cost of services and activities (2008/9 £) (WP2) 45

TABLE 15 Regression of one-to-one contact costs and activity costs on demographic and clinical characteristics (clustering for ward) (WP2) 46

TABLE 16 Baseline characteristics of service user participants (WP3) 57

TABLE 17 Baseline characteristics of staff participants (WP3) 58

TABLE 18 Number of data collection periods to which staff contributed (WP3) 59

TABLE 19 Baseline characteristics of wards (WP3) 60

TABLE 20 Adjusted estimate of intervention effect on VOICE standardised scores ($N = 1051$) (WP3) 61

TABLE 21 Adjusted results for VOTE and VOTE subscales (standardised scales) (WP3) 62

TABLE 22 Cost of the intervention (WP3) 63

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 23 Results for cost of perceived service use (WP3) 63

TABLE 24 Number of violent incidents per service user (WP3) 66

TABLE 25 Unadjusted results for the rate of violence ($n = 1108$ from one NHS trust) (WP3) 66

TABLE 26 Adjusted results for IWS and MBI standardised scales (WP3) 66

TABLE 27 Ward reported activities (adjusting for ward effects) (WP3) 67

TABLE 28 Results for self-reported (CITRINE) number of activities attended (WP3) 67

TABLE 29 Results for self-reported (CITRINE) number of sessions attended (WP3) 67

TABLE 30 Description of admission characteristics to the routine and triage systems 79

TABLE 31 Length of inpatient stay (WP4) 80

TABLE 32 Re-admission to inpatient care (WP4) 83

TABLE 33 Service user demographic characteristics by care system (WP4) 84

TABLE 34 Staff demographic characteristics by care system (WP4) 85

TABLE 35 Results for primary service user outcomes (WP4) 88

TABLE 36 VOICE exploratory model (WP4) 89

TABLE 37 Full results for the regression models of the staff primary outcomes (WP4) 91

TABLE 38 Service use and costs (WP4) 92

TABLE 39 VOCALISE measure generation group 133

TABLE 40 Mean VOCALISE scores 134

TABLE 41 Demographic and clinical characteristics 139

TABLE 42 The ANCOVA controlling for baseline and PANSS negative $*p < 0.05$; $**p < 0.001$; $***p < 0.0001$ 139

TABLE 43 Unit costs of services and activities: care professional 143

TABLE 44 Unit costs of services and activities: activities 143

TABLE 45 Stepped-wedge randomisation procedure 169

TABLE 46 No cost extension additional wards and time points 169

TABLE 47 Stepped-wedge randomisation procedure 172

TABLE 48 No cost extension additional wards and time points 172

List of figures

FIGURE 1 Decision tree structure 36

FIGURE 2 Forest plot of meta-analysis: reduction in general symptoms PANSS 38

FIGURE 3 Cost-effectiveness plane: mean costs and effects of social cognition relative to TAU with 95% confidence region 39

FIGURE 4 Cost-effectiveness acceptability curve 39

FIGURE 5 Intervention schedule for DOORWAYS 51

FIGURE 6 Cost-effectiveness plane using VOICE with 95% confidence ellipse (WP3) 63

FIGURE 7 Cost-effectiveness acceptability curve using VOICE (WP3) 64

FIGURE 8 Cost-effectiveness plane using VOTE with 95% confidence ellipse (WP3) 64

FIGURE 9 Cost-effectiveness acceptability curve using VOTE (WP3) 65

FIGURE 10 Model of potential relationships between intervention and outcome (WP3) 69

FIGURE 11 Length of care components (WP4) 82

FIGURE 12 Service user VOICE for each system by phase (WP4) 86

FIGURE 13 Service user SSS-RES for each system by phase (WP4) 87

FIGURE 14 Staff outcomes for each system by phase (WP4) 90

FIGURE 15 Cost of perceived service use (WP4) 93

FIGURE 16 Cost-effectiveness plane using VOICE with 95% confidence ellipse (WP4) 94

FIGURE 17 Cost-effectiveness acceptability curve using VOICE (WP4) 94

FIGURE 18 Cost-effectiveness acceptability curve using VOTE (WP4) 95

FIGURE 19 Cost-effectiveness plane using VOTE with 95% confidence ellipse (WP4) 95

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Glossary

MIND MIND mental health charity.

VOCALISE Barriers to change. Measure developed to assess nurses' barriers to change.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

List of abbreviations

AIC	Akaike information criterion
AIHQ	Ambiguous Intentions Hostility Questionnaire
ANCOVA	analysis of covariance
ANOVA	analysis of variance
BETTER PATHWAYS	Bringing Emergency TreatmenT to Early Resolution
BME	black or minority ethnicity
CBT	cognitive–behavioural therapy
CI	confidence interval
CITRINE	Client Services Receipt Inventory – Inpatient
CMHT	Community Mental Health Team
CQC	Care Quality Commission
CRT	cognitive–remediation therapy
DOORWAYS	Delivering Opportunities for Recovery
EQ-5D	EuroQol-5 Dimensions
GAF	Global Assessment of Functioning
HONOS	Health of the Nation Outcome Scales
HTT	home treatment team
ICC	intraclass correlation
IWS	Index of Work Satisfaction
LIAISE	Lasting Improvements for Acute Inpatient SEttings
MBI	Maslach Burnout Inventory
NICE	National Institute for Health and Care Excellence
NOSIE	Nurses Observational Scale for Inpatient Evaluation
OT	occupational therapist
PANSS	Positive And Negative Syndrome Scale
PERCEIVE	Patient Involvement in Improving Patient Care
PICU	psychiatric intensive care unit
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
PROM	patient-reported outcome measure
QALY	quality-adjusted life-year
RCN	Royal College of Nursing
SCIT	social cognition and interaction training
SCT	social cognition training
SD	standard deviation
SE	standard error
SSS-30	Service Satisfaction Scale-30
SSS-RES	Service Satisfaction Scale – Residential Services Evaluation
SURE	Service User Research Enterprise
TAU	treatment as usual
ToM	theory of mind
VOICE	Views On Inpatient CarE
VOTE	Views On Therapeutic Environment
WAS	Ward Atmosphere Scale
WP	work package

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Plain English summary

This project set out to measure the therapeutic environment of NHS wards in mental health trusts and to see if this is affected by increasing ward activities or changing the way service users are admitted, and whether or not these changes affect costs. We created new ways of measuring the perceptions of service users and staff as well as a new economic scale to measure costs for individual service users.

We worked closely with service users and staff to design the questions that made up these new surveys and tested the questionnaires to ensure their reliability.

We hypothesised that the therapeutic environment of wards would improve by training nurses to deliver group activities. To test this, we ran a 4-year randomised controlled trial on 16 wards, comprising 1108 service users and 539 staff. Although activities did increase, we discovered only a significant improvement in the views of those service users who had been admitted under a legal section. There was no evidence that the introduction of training had any effect on staff even though their workload increased.

We also compared a model of admission called triage with traditional hospital admission to discover if it could affect the admission length or ward environment. In the triage system, patients remain on the ward for a maximum of 7 days and then are discharged or transferred to a longer-stay ward. We found no evidence that service user or staff perceptions differed between the two systems (triage and regular admission) or that cost-effectiveness was improved.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Scientific summary

Background

This programme had three aims: (1) to develop two measures of stakeholder perceptions using stakeholder methodology, (2) to evaluate stakeholder perceptions and the cost-effectiveness of two differing admission systems to inpatient acute wards and (3) to evaluate the stakeholder perceptions and the cost-effectiveness of increasing the number of evidence-based therapeutic activities on inpatient wards. Recently, the spotlight has been put on two aspects of this programme, patient and staff perceptions of inpatient care and improving the quality of inpatient care, that were central at its inception and have remained so throughout. The report into the failings of Mid Staffordshire Hospital highlighted the patient experience as of equal importance as quality and safety. The Patient Involvement in Improving Patient Care (PERCEIVE) programme sought to understand the experience of service users and nurses on mental health acute wards and whether or not a therapeutic intervention could improve what had, for many years, been seen as poor-quality care. We examined whether a new service configuration could improve experience. Mental health services are receiving more attention and parity between physical and mental health care has been agreed by all the major political parties and is in the 5-year forward plan. The public are also now aware of the problems in mental health care, for instance, the *Guardian* ran a feature by an ex-mental health nurse detailing the substandard inpatient care received by a friend. Equally, the most recent report by the Care Quality Commission painted a bleak picture of inpatient care, including reporting that the numbers of people detained and compulsorily treated was now at its highest point ever. Therefore, the results from the PERCEIVE study are timely.

Objectives

We had six principal objectives:

1. to use stakeholder participatory methods throughout the project
2. to develop and validate a measure of service user views of inpatient care
3. to develop and validate a measure of nurse views of inpatient care
4. to develop and validate a patient-reported measure of staff contacts and uptake of activities while on inpatient wards
5. to develop a training programme for nurses to increase their skills in running therapeutic activities and evaluate its effects on nurses and service users
6. to evaluate and compare a novel admission system to a traditional one on the perceptions of nurses and service users.

Work package 1: LIAISE

Methods

Lasting Improvements for Acute Inpatient SEttings (LIAISE) was designed to develop and test two self-report measures through a process of stakeholder (service user and nursing staff) involvement. This approach directly involved users of mental health services and nursing staff in order to ensure that measures were produced that captured an accurate picture of an acute care ward from both perspectives. The research was conducted in six stages:

1. Three stakeholder groups mapped the dimensions of inpatient care.
2. Four focus groups each for service users and nurses met twice to discuss what issues mattered to them.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

3. The researchers constructed a mixed-methods measure.
4. Expert panels of service users and nurses refined the measures.
5. A total of 146 inpatients and 40 staff took part in feasibility exercises including completing the measure and answering questions assessing whether or not the measure was easy to complete, understand and acceptable.
6. Test–retest reliability and validity were assessed.

Results

1. The service user measure, Views On Inpatient CarE (VOICE), is a 19-item self-report measure that was found to be easy to complete and understand. It has good test–retest reliability and good criterion validity. It had a two-factor structure and the factors were termed 'care' and 'security'.
2. Views On Therapeutic Environment (VOTE), the nurse measure, is a 22-item self-report tool that was found to be easy to complete and understand by nurses working on inpatient wards. It has good test–rest reliability and a three-factor structure, the factors being named 'workload intensity', 'team dynamics' and 'interaction anxiety'.
3. The test–retest reliability of the service user measure ($n = 40$) was high [ρ 0.88, 95% confidence interval (CI) 0.81 to 0.95] and there was no difference in score between the two assessments. After removing items with poor reliability, this left a 19-item scale with high internal consistency ($\alpha = 0.92$). The results for the staff measure showed a mixture of moderate and substantial reliability. Concordance between the total scores ($n = 34$) was good (total score, ρ 0.77, 95% CI 0.65 to 0.89). The internal consistency of the measure was good with the overall alpha at 0.82.

Secondary qualitative analysis showed that service users and nurses had different perceptions of life on an acute ward with nurses emphasising management tasks, pressure of work and violence, and service users emphasising the inaccessibility of nurses and unnecessary coercion.

Work package 2: CITRINE

Methods

1. Systematic review of studies measuring staff–patient contact time on psychiatric inpatient wards.

    i. We searched electronic databases with the criteria for inclusion that (1) the study was based in an inpatient setting and (2) that it measured or observed and recorded time spent by patients and/or staff on different activities.

2. Systematic review of economic evaluations of therapeutic activities on psychiatric inpatient wards.

    i. We searched the same databases for economic evaluations of interventions of therapeutic interventions based in an inpatient setting.

3. Modelling the cost-effectiveness of inpatient interventions.

    i. Modelling to assess cost-effectiveness was demonstrated using social cognition training as a case study. Treatment lasting 8 weeks with two 45-minute sessions per week led by a psychotherapist was assumed and the general psychopathology subscale of the Positive and Negative Syndrome Scale (PANSS) was chosen as an outcome. Cost-effectiveness was assessed using the net monetary benefit approach.

4. Development of a measure of service use.

    i. To develop a tool to collect data on care contacts and activities of inpatients we conducted interviews with inpatient staff to discuss content and structure, collected data from service users and assessed reliability. Contacts with staff members were combined with unit costs. Activity costs were calculated based on session duration, preparation required, staff involved and materials for each activity provided on each ward.

Results

1. The first search identified 86 papers of potential relevance. After reviewing abstracts, two studies were initially included with 11 others included after examining reference lists.
2. The second search identified 236 papers but none met the search criteria.
3. For the model, three relevant studies provided a weighted mean reduction in PANSS score owing to the intervention (vs. usual care) of 3.49 points. To be at least as equally cost-effective as usual care, the intervention needed to reduce hospitalisation by at least 8.5–10.5 percentage points.
4. The Client Services Receipt Inventory–Inpatient (CITRINE) typically took < 10 minutes to complete. Service users reported more activities than shown in case notes and the average cost difference was £10. Case notes reported more one-to-one nursing contacts, resulting in an average cost difference of £4.

Work package 3: DOORWAYS

Methods

Delivering Opportunities for Recovery (DOORWAYS) was a cluster randomised trial, using a stepped-wedge design, to evaluate the impact of training nurses to deliver therapeutic group activities on (1) service users (2) nursing staff and (3) cost-effectiveness. Initially, 16 wards (clusters) were randomised to receive the training at 6-month intervals, two wards in each period. Eight were included initially and a further eight in an extension. This was a pragmatic design that produced different measurement periods for the control and experimental periods in the two sets of randomisations (from three to five measurement periods). The perceptions of service users and staff were collected from wards every 6 months using the VOICE and VOTE designed in work package 1. Service users were excluded if they:

1. had not been on the ward for at least 7 days
2. could not understand English well enough, or were too unwell, to give informed consent
3. had previously participated in the study.

Nurses were excluded only if they were temporary staff who had completed fewer than seven shifts in the preceding month.

Results

A total of 1108 service users participated in the study: 670 on pre-intervention wards and 438 on post-intervention wards. There were 539 staff participants, who contributed between one and five measurements. Using regression analysis accounting for time and ward as fixed effects we found no evidence overall that service user VOICE scores were improved by the intervention (standardised intervention benefit 0.19: mean VOICE score pre intervention = 56.5, SD = 19.1, $n = 644$; mean post intervention = 54.2, SD = 17.2, $n = 414$). However, we found an improvement for those who had been admitted under a legal section [standardised improvement of –0.35 (95% CI –0.12 to –0.57; $p = 0.002$)]. However, we found no evidence of any change in VOTE scores from nursing staff (standardised effect size = 0.04, 95% CI –0.09 to 0.18; $p = 0.54$).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Work package 4: BETTER PATHWAYS

Methods

Study 1: is a triage system more efficient and cost-effective than traditional care?

Data on all admissions and discharges were extracted from the comprehensive electronic patient record. For each service user, dates of inpatient admission and discharge, data on input from home treatment teams (HTTs) and demographic and diagnostic characteristics were extracted. Two data sets were created to contain:

1. all admissions to the two systems between 1 January and 31 December 2009 for analyses of length of stay and costs
2. all re-admissions in the year following discharge from their first admission (i.e. index admission) in 2009.

We multiplied the number of inpatient days and the number of HTT visits by appropriate unit cost figures to estimate the cost of service use.

Study 2: service user and staff perceptions of care

We collected data at four time points on staff and service user perceptions of care in the triage and routine care system. We compared differences between systems and differences between different types of wards within the triage system.

Results

Study 1

There were few differences between the characteristics of people admitted in the two systems. We found no evidence of a difference in length of inpatient stay (adjusted estimate 0.5 days shorter in triage; $p = 0.90$). The mean cost of an acute care episode (inpatient length of stay and number of home treatment visits) was £15,233 in the triage system compared with £15,476 in the routine system. Adjusting for potential confounders, we did not find any evidence of a difference between the two systems in the cost of acute care for service users treated within their catchment area (estimated difference: triage system £391 higher, 95% CI –£2535 to £2748; $p = 0.77$).

Study 2

There was no difference in the VOICE scores of service users between the two systems (adjusted estimate: 0.77 better in triage, 95% CI –2.90 to 4.44; $p = 0.68$). There was no significant difference ($p = 0.38$) between the two systems in staff VOTE scores at baseline, but staff experience in both systems changed over the period of the study and these changes differed between systems.

Conclusion

Our studies showed that it is possible to produce psychometrically robust measures of stakeholder experience using a participatory methodology and to produce a meaningful measure of staff and service user interaction by involving service users and nurses. The measure of service use is innovative in that it focuses on what service users find meaningful. The training intervention (DOORWAYS) resulted in improvements to the patient experience only in those involuntarily admitted as measured by VOICE but, contrary to our expectations, our intervention had no effect on nursing staff. Post-intervention focus groups indicated that the DOORWAYS groups were often not implemented as planned owing to staff being needed for other things, incidents on the ward or their lack of appeal to service users. The DOORWAYS intervention resulted in a modest shift of inpatient resources towards direct patient care. Whether triage systems are considered in terms of a large database or the service users who participated in the Bringing Emergency TreatmenT to Early Resolution (BETTER PATHWAYS) research, they had no effect on length of

stay, re-admissions or cost-effectiveness. VOICE and VOTE scores did not differ by service system but time did have an effect. Our database will be used to develop an understanding of the mediating and moderating factors for improving care quality.

Trial registration

This trial is registered as ISRCTN06545047.

Funding

Funding for this study was provided by the Programme Grants for Applied Research programme of the National Institute for Health Research.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 1 Overview of the programme

Mental health, including inpatient mental health, is in the news. In 2011, the government released a strategy entitled *No Health Without Mental Health* to close the gap between physical and mental health services.[1] It stated 'We are clear that we expect parity of esteem between mental and physical health services' (contains public sector information licensed under the Open Government Licence v1.0).[1] More recently, the 5-year forward plan now includes parity between mental and physical health[2] and a policy document now promises to remedy the crisis in mental health care.[3] The *Guardian*[4] also ran a feature in the same week describing the deleterious effects of bed closures on mental health inpatient services. These problems, although highlighted in 2014, were on the horizon when this research programme was developed and its results are now even more relevant with the recognition of further problems and the political support for parity between mental and physical health services.

In this chapter we describe the background, rationale and aims of our programme, which investigates the inpatient experience of mental health service users. We do this from a stakeholder perspective by involving both mental health services users and clinicians in all phases of the study. Throughout the report we refer to the programme as PERCEIVE which refers to Patient Involvement in Improving Patient Care. PERCEIVE consists of four work packages (WPs), two methodological and two investigating innovations into inpatient mental health care. PERCEIVE's overarching aim is to investigate and improve the delivery of inpatient mental health services. This descriptive information can also be found on www.perceive.iop.kcl.ac.uk/about.html.[5]

Background

In the UK, as well as internationally, there has been a gradual move to treating mental health service users in the community and a reduction in the number of mental health hospital beds. Alongside this there has been an unceasing pressure on mental health services to reduce costs. Inpatient care is the costliest component and so has been at the forefront of these cost-saving exercises. A great deal of effort, time and money has gone into improving mental health services over the past 50 years, often without any rigorous evaluation. This project seeks to examine the costs, the effects of innovative organisational changes made to service delivery, what activities service users engage in while inpatients and, most importantly, the experience of inpatient care from a stakeholder perspective.

Pressure on the number of inpatient beds available

Advances in psychopharmacology, changing attitudes towards mental illness and institutionalisation and the growing concerns about the quality of life in asylums have led to the gradual transition from the long-term hospitalisation of those with mental health problems to an increase in community treatment. However, hospitalisation for working-age adults is still an indispensable feature of mental health care and in 2011/12 accounted for around 40% of the NHS budget on mental health care for this group.[6] Bed numbers have continued to fall over the past 50 years, although the need for them to provide respite for the acutely ill remains the same. This reduction in the number of beds means that wards are now generally reserved for the most acutely ill. Consequently, there is an increase in the proportion of inpatients who are compulsorily detained and in the levels of behavioural disturbance.[7–9] In turn, this fraught environment can lead to difficulties in staff recruitment and retention. Taken together, this often brings about ward environments that appear to be more custodial than therapeutic.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Therapeutic activities and positive interactions on wards

Despite heavy investment, service users and front-line staff continue to complain about the quality of inpatient care.[10] More than 50 years ago, when institutionalism was being re-evaluated, Wing and Brown carried out their well-known *Three Hospitals* study which confirmed that the hospitals that provided richer social environments and social opportunities had markedly less disturbance in verbal and social behaviours.[11] Over the 5 years of the Wing and Brown study, the social environment improved in all participating hospitals and the clinical condition of patients improved alongside it. Those patients with the least access to the outside world, the least amount of social interaction and with the fewest interesting activities to take part in and who spent the majority of their time doing nothing were the most unwell. However, the study found that as their environments became more stimulating, they too showed some improvements.

All inpatient wards strive to provide an environment that is conducive to recovery and the importance of the social environment and therapeutic activities is recognised not just by those providing services, but also by those who use the services. For instance, a Mind report[10] examining acute wards across England and Wales found that service users often complained about the intense boredom they experienced. The lack of therapeutic activities (or the cancellation of those on offer) is often as a result of staff shortages and/or too many demands being placed on staff. Service users are aware of this dilemma.[10] Nurses are often blamed for being immersed in the crisis management for a minority of patients or for spending time on administrative duties; both of these problems have increased as the acuity on wards increases. Nurses are therefore much less able to spend time on therapeutic activities or in direct patient contact.[12,13] Another reason often given by staff for not providing or running groups on acute wards is that service users are too ill to appreciate or take part in structured activities. However, it is not obvious that this is the case and it is the subject of some of our studies. This is not borne out in empirical studies. For instance, Kavanagh *et al.*[14] showed that education groups for individuals on a psychiatric intensive care ward were beneficial in improving knowledge of medication and the side effects of medication, and were appreciated by service users despite high levels of acuity in this group. A further barrier to the regular provision of therapeutic activities might be that the nurses lack the skills or confidence to run them. Hence, we included training and increasing skills for all staff on the ward. Although the majority of research in this area has originated in the UK, a recent review of the literature on patient activities and interaction encompassed international perspectives and found that this lack of service user engagement seems to hold true across North America, Australia and Europe.[15]

A Sainsbury Centre for Mental Health report[16] provided a more optimistic view of inpatient care, finding that nearly three-quarters of ward managers across England reported that both practical activities (e.g. cooking skills) and talking therapies occurred regularly on their wards, and 64% reported that leisure activities (e.g. bingo) occurred regularly. However, therapies with a strong evidence base, such as cognitive–behavioural therapy (CBT), occurred only on 20% of wards. The recommendations included increasing the availability of training in therapeutic activities that have an evidence base.[16] What remains unknown from the report is how often therapeutic activities happen because, although activities may be something the wards aspire to provide and schedule into the ward timetables, how often they happen in practice may be a different matter.

Innovations in the service pathway have also been suggested to improve the therapeutic atmosphere of wards by reducing the patient flux where recovery-oriented approaches might be made. Hence, we included a comparison of two different service systems that might show how changes in patient flow could affect service user views. The innovative change was the introduction into one system of a triage ward where service users were assessed for longer-term care or a return to the community with the support of the home treatment team (HTT).

Stakeholder involvement

Our programme of research sought to involve relevant stakeholders in all aspects of the programme, not only in the WPs themselves but also in the analysis and publications from the programme. PERCEIVE was proposed by a group of clinicians, NHS managers, clinical academics and service users. Our stakeholders were represented on the steering and management committees. A senior service user researcher (DR) who spent time on inpatient wards contributed to the overall design of the programme as a whole. WP1 was wholly dedicated to stakeholder involvement and was co-ordinated and run from the Service User Research Enterprise (SURE). This is a research unit which is co-directed by a service user researcher and employs people who are skilled researchers and who also have experience of using mental health services. WP1 [Lasting Improvements for Acute Inpatient SEttings (LIAISE)] researchers were situated in SURE, supervised by DR, making LIAISE a user-led programme of work.

We involved service users and nurses in developments in each WP. The largest input was to WP1, which involved producing both a service user and nurse measure of perceptions of inpatient wards. To do this, a participatory model was used,[17] such that service user researchers facilitated the production of the service user measure and nurse researchers the generation of the nurse measure. It was made clear to participants that the researchers themselves were like them and that it was their experience that counted, something that is not always the case in health services research. At all stages of the measure development, including the measure developed in WP2, researchers met with service users for feedback at various stages and changes were made based on their suggestions. In all WPs, we also employed service user and nurse researchers to collect data and to develop the programme of activity. At the completion of WP3, we returned to the wards involved and sought feedback from service users and nurses on the impact the project had and what we could have done better.

Additional studies that flowed from the programme were nearly all the result of service user suggestions (see *Appendices 1–4* for details).

Involving service users as researchers and seeking to focus on the patient experience is a vital component of this project. The Francis report[18] has highlighted patient experience as important in terms of both clinical outcomes and safety and quality, something which has been a key part of our project from the start.

Health economic considerations

Given the high cost of inpatient care and, in particular, the context of limited health-care resources, it is essential to consider issues of cost-effectiveness in evaluations of new and existing interventions. However, there is a paucity of cost-effectiveness evidence pertaining to interventions focused on psychiatric inpatient wards such as increased therapeutic activities or triage models. Indeed, the methods used in most economic evaluations may actually be deficient in evaluating inpatient care because these frequently just record the number of days an individual spends as an inpatient and combines this with a standard unit cost. This ignores the fact that there will be heterogeneity in the number of staff who inpatients are in contact with and the uptake they have of activities irrespective of length of stay. The health economic component of the programme addressed these shortcomings by (1) developing an inpatient-focused resource use questionnaire and (2) using this questionnaire in the evaluations described in WP3 and WP4.

Work package 1

The first WP has the acronym LIAISE standing for Lasting Improvements for Acute Inpatient Settings. This was the WP that constituted service user-led research and extended that model to nurse research. The aim of the WP was to produce two measures capturing the experience of living and working in acute inpatient

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

settings. The measures were generated using focus groups, expert panels and a feasibility study. The researchers for the service user measure were service user researchers [Views On Inpatient CarE (VOICE)] and the researchers for the nurse measure were nurse researchers [Views On Therapeutic Environment (VOTE)]. Both measures underwent psychometric testing and showed sound properties.

Work package 2

This WP was called Client Services Receipt Inventory – Inpatient (CITRINE) and had as the main aim the development of a measure of staff contacts and use of therapeutic activities on inpatient wards. As a precursor to this we conducted a systematic review of the literature to establish how inpatient activities had been measured in previous studies. This was then followed by the development of a draft questionnaire to record contacts and activities over a 1-week retrospective period. Consultation with clinical staff was crucial in this process and the questionnaire was piloted to determine its validity, reliability and acceptance from service users. The final version was then used in WP3 to evaluate the introduction of therapeutic activities to the wards and WP4 to evaluate the triage model.

Alongside the development and use of the questionnaire, we also reviewed the literature to identify economic evaluations of inpatient-based interventions. Our expectation was that these would be limited in number and, as an alternative to trial or observational methods, provide a demonstration of a decision modelling approach to conducting such evaluations.

Work package 3

The third WP [Delivering Opportunities for Recovery (DOORWAYS)] aimed to increase staff training and thereby increase the number of therapeutic activities occurring on the wards. Nurses working on the participating wards were offered training in a number of evidence-based therapeutic activities. After extensive discussion with nursing staff and their managers, a menu of therapies was chosen, from which each ward could choose the activities best suited to their needs. A stepped-wedge design was used to investigate the effect of the intervention on staff and service user perceptions of the wards. Nurses on all wards were offered the training, with two wards receiving training in each 6-month period. The order in which the wards were offered the training was determined by randomisation.

The wards randomised to the intervention would then receive the training. This consisted of a series of workshops involving a service user trainer and a clinical psychologist. For psychological activities, the clinical psychologist then demonstrated the group on the ward to the nurses, who then co-facilitated the group with the psychologist until ward staff were able to run them independently. Data collection on service user and nurse perceptions of the wards were completed before and after each randomisation and are the main outcomes in this WP.

Work package 4

The final WP [Bringing Emergency TreatmenT to Early Resolution (BETTER PATHWAYS)] involved evaluating a unique admission system with a traditional admission system. One service within our catchment area developed a novel approach to solving the overcrowding and long admissions problems usually associated with inpatient care. They designated a specific ‘triage’ ward to accept all acute admissions in the catchment areas with immediate discharge planning beginning at admission, a guarantee to be assessed by senior members of the team and a tight working relationship with their HTTs. At the end of 7 days, service users were either discharged back into the community or transferred to one of three locality wards providing longer-term, more rehabilitative care. Initial reports suggested that it was working well but whether or not this has been maintained has not been evaluated. We undertook this WP in order to investigate whether or

not the triage system does indeed reduce lengths of stay in inpatient care and whether or not it leads to a greater number of re-admissions than a more traditional system of inpatient care. Furthermore, we also investigated service users' and nurses' perception of these two systems to see if there were any clear preferences.

Objectives

We had six principal objectives:

1. To use methodology throughout the project which involved stakeholder participants.
2. To use stakeholder participatory methodology to develop and validate a measure of service user views of inpatient care.
3. To use stakeholder participatory methodology to develop and validate a measure of nurse views of inpatient care.
4. To develop and validate a patient-reported measure of staff contacts and uptake of activities while on inpatient wards.
5. To develop a training programme for nurses to increase their skills in running therapeutic activities and evaluate its effects on nurses' and service users' views of the wards.
6. To evaluate and compare a novel admission system to a traditional one on the perceptions of nurses and service users.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 2 Work package 1: LIAISE

Participatory research methods

The VOICE and VOTE measures were generated using a model of stakeholder involvement which holds at its core the concept of empowering under-represented groups in order to create well-balanced, responsive services. This method, pioneered in service user research,[17,19–21] is novel because a researcher from the same group as the participants guides the process and the data are allowed to emerge by following the themes that the participants feel are the most important. This was also adapted for a nursing setting, with the emphasis being on a shared professional understanding, which fostered a non-judgemental research forum on which staff could express their experiences safely. A more balanced power dynamic between the researcher and the researched led to data and, subsequently, items on the scales that more accurately represented the views of services users and staff.

Background

In the UK and across the world, mental health hospitals have seen nursing staff shortages, bed shortages and an increase in complex presentations, especially in urban areas of high demand.[22–24] Indeed, mental health nurses perceive these issues as barriers to change in acute ward settings as well as poor/indeterminate leadership and incidents.[25,26] The effects of current working practices and the ward environment on the well-being of both staff and service users should be explored in more detail because an untherapeutic environment may negatively affect service user satisfaction and staff morale, may have a direct impact on the quality of the interactions between service users and staff and, moreover, may be detrimental to the overall quality of care.[27]

Dissatisfaction with adult acute inpatient care is not a new issue and is well documented both in Britain and internationally. Negative service user reports of acute inpatient settings have emerged, describing the experience of hospital care as non-therapeutic and coercive.[28–30] Limited interaction between staff and service users is commonly reported and users express a need for good interpersonal relationships and support that is sensitive to individual needs.[31–33] Poor levels of involvement and a lack of information associated with medication, care and treatment have also been identified.[34] On many wards, there is little organised activity and service users experience intense boredom.[10] Security is of particular concern as many service users feel that they are not treated with respect or dignity, have significant safety concerns and report high levels of verbal and physical violence.[10]

There is a rather sparse literature written by service user researchers themselves on acute inpatient settings, but it is important because information given to academic or clinical researchers may be influenced by the differential status of the two parties. Both Walsh and Boyle[34] and Rose[19] report that service users experience low levels of involvement and a lack of information about their care and treatment. This mirrors the nursing literature. Rose's[19] participants reported a lack of activities leading to crushing boredom and, like Mind's Ward Watch campaign, patients reported feeling unsafe.[10]

Alison Faulkner, a service user-researcher who spent time in acute wards, combines her systemic knowledge of mental health services with her own experience and that she has witnessed of others.[35,36] Drawing on the grey literature,[30] she confirms the 'shabbiness' of wards, the intense boredom felt by the inpatients and the 'petty' rules and regulations.[35] She further concludes that in the inpatient wards she has frequented she was treated as 'less than human' and witnessed other patients also being treated as if they were a different species.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Studies in which the researcher was a user[37,38] found that what was most important to service users was the quality of the therapeutic relationship. This is entirely in line with both UK policy and some of the nursing literature. Interestingly, good therapeutic relationships were prized whether the user was in a traditional acute ward or an alternative style of service, such as a crisis house.

Turning to the work environment for nurses, a growing body of literature links the stress of working on a ward to low morale, often measured as 'burnout' and poor job satisfaction.[39–41] Stress, burnout and lack of social support have been described as symptoms of a negative work environment.[26,42–44] Stressors such as high caseloads of patients, high volume of work, violence on wards[45] and management issues such as poor leadership and low staffing levels[42,46,47] have been linked to low morale. Different staff groups also seem to respond differently to work-related stressors. Jenkins and Elliott[43] reported differences in perceptions of work-related stressors according to the occupational status of nursing staff. In their study, qualified staff cited poor staffing levels as the main stressor, while nursing assistants reported the main stressor as difficult interactions with distressed clients. Burnout was common to both groups. Cushway *et al.*[48] noted that occupational stress led to poorer mental health outcomes in male nursing staff than in females, perhaps because male staff are more likely to respond to violent situations. Burnout has also been described as affecting job performance and contributing to poor-quality care.[46,47,49] Both low morale and low job satisfaction result in poor staff retention, which again affects care quality.[50]

The methods through which organisational outcomes are decided and the effects of these methods on the perceptions of the workforce may also be implicated in this negative trend. Organisational change in the NHS is often imposed via a top-down approach, which may not take into consideration the views of nursing staff. Indeed, change process issues, such as poor involvement in planning, implementation and control of a project, have been highlighted as potential barriers to success.[51] An inequitable stance to decision-making, perhaps as a result of poor staff involvement, may shape attitudes to planned changes if staff feel such changes are unfair to them or do not meet the needs of their client group.[52] This argument is supported by evidence from both the nursing and wider literature, which shows that staff who are in leadership roles may view changes more optimistically than staff who are involved in direct care with service users because they form part of the consultation process for the proposed changes.[53,54] The emotional response of staff to proposed changes may be an important barrier to the success of implementation.

Therefore, we propose a model where negative staff perceptions of the daily pressures of working on an acute ward and negative staff perceptions of barriers to change lead to burnout and low job satisfaction, and this affects the quality of care delivered. However, in order to test this model, the specific pressures for mental health nursing staff in maintaining a therapeutic environment and high-quality client relationships and the quality of care as perceived by service users must first be given more detailed consideration because, to date, these constructs remain underexplored. This is perhaps because the ward milieu is responsive to continuous organisational and social changes, that is, the client group and staff mix change on a daily basis and so this is a complex construct to measure. This model was considered in an add-on study carried out by a nurse researcher as part of her PhD and is presented in the appendix (see *Appendix 3*).

Objective measures of inpatient care, reviewed by Sharac *et al.*,[15] have many drawbacks. They neither capture the complexity of ward dynamics nor reflect the quality of inpatient care. Measures of perceptions have an important role to play in evaluating complex situations because they allow for objective stressors and their appraisal by subjects within the social environment, as well as linking cognitions with affective and behavioural components.[55]

Recently there has been a focus on patient-reported outcome measures (PROMs) as a measure of quality and appropriateness of services and therapies. Despite service user involvement being considered an essential element in improving mental health services,[56] PROMs are rarely developed using an inclusive methodology and research suggests user dissatisfaction with many outcome measures currently in use.[57]

Service users can often have different perspectives to professionals and can provide insight into how services and treatments feel.[58] Redefining outcomes according to users' priorities can help to make greater sense of clinical research and develop a more valid evidence base.[36,59] Studies comparing the impact of traditional and user researchers in conducting research show some differences in qualitative data analysis[60] but none on quantitative research findings.[17,61] Given this, we believe that research methodologies should aim to be as inclusive as possible.

There are also some staff measures in existence that address attitudes to occupational stress (e.g. The Mental Health Professionals Stress Scale[48] and occupational stress indicator[62]) that highlight staff reactions to organisational pressures and team and client relationships. The well-known Daily Hassles Scale[63] has also been used to appraise these stressors. However, these measures neither adequately explore whether staff view daily pressures as having an impact on the therapeutic milieu (and, therefore, on the quality of the therapeutic relationship) nor have they been developed using qualitative data generated by staff working within those areas.

In nursing studies, there has been limited focus on assessing how social, emotional and psychological barriers affect the implementation of new innovations. The relationship between psychological well-being and the acute ward setting in relation to planned changes may, in nursing, be distinct from other professions because nurses remain on the ward for the duration of their shift and are often expected to play a key role in delivering changes. There are several measures developed in health care that focus on the uptake of evidence-based practice in clinical areas by mental health providers and nurses.[64–71] However, there are no measures that focus on general changes in a mental health setting.

What is needed in the literature on acute care are psychometrically robust, brief, self-report measures reflecting service users' experiences of receiving care and staff experiences of delivering/innovating that care. Measures of this type would allow a clear assessment of changes to inpatient care following specific interventions to improve the environment and therapy provided. Our study was designed to generate such measures.

The purpose of this section is to:

- describe a model of service user and nurse involvement in developing two psychometrically sound measures. The service user measure captures service user perceptions of acute inpatient wards (VOICE) and the nurse measures capture staff perceptions of the daily pressures working on an acute ward (VOTE)
- provide a secondary qualitative analysis of the data from VOICE and VOTE.

Methods

This study was designed to develop and test two self-report measures. The measures were generated through a process of stakeholder (service user and nursing staff) involvement.[17,19–21,59] This approach directly involved users of the mental health services and nursing staff in order to ensure measures were produced that captured an accurate picture of an acute care ward from both perspectives. Following the measure generation phase that used qualitative methods, prototypes of the new measures were assessed for their content validity, acceptability and feasibility prior to psychometric testing on two separate large samples of inpatient nurses and service users. Ethics approval was awarded by Bexley and Greenwich NHS Research Ethics Committee (07/H0809/49) for the study to be carried out in four boroughs within a London NHS Mental Health Trust.

Focus groups were conducted to generate data for VOICE and VOTE.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Phase 1: mapping out the dimensions of inpatient care

In this initial phase, three stakeholder/reference groups were held to map out the dimensions of inpatient care as a topic guide. In developing VOICE, the stakeholders were service users with experience of research and inpatient care who were not currently using local services. National voluntary organisations were also represented. In developing VOTE, the stakeholders consisted of senior nursing staff. Each group built a topic guide that was flexible in order to encourage as creative a list as possible. Members of these two groups also then constituted the reference groups for the user-led and staff parts of the study.

Phase 2: measure generation

For both VOICE and VOTE, purposive sampling was used to reflect local inpatient demographics with regard to gender, age and ethnicity when recruiting participants who were either staff or service users of an inner London UK mental health trust.

VOICE sample

The sample consisted of people who had used inpatient services within the previous 2 years. One of the groups was specifically for participants who had been detained under the Mental Health Act (2007)[72] as it was anticipated that they may have had different experiences.

VOTE samples

Nursing staff from all grades (health-care assistants, entry-level qualified staff, clinical charge nurses and team leaders) were asked to participate in the development of VOTE. They were purposively sampled from acute inpatient mental health wards in an inner London UK mental health trust.

Focus groups

A pilot focus group was used to test the items on the topic guide. Then four two-part focus groups were run to enable respondent validation. This meant that the data from the first focus groups were analysed and results presented back to respondents in the second wave of focus groups to see if they agreed with the analyses or wished to amend, add to or change them. The aim was to ensure that the analyses were an accurate representation of the participants' views.

The service user and nurse researchers then refined the thematic analyses and constructed the draft measure.

Expert panels

As a penultimate step, two expert panels met to discuss the design of the new measures and to inform the 'instructions for use'. For the service users, one expert panel was made up of focus group members and the second was independent. Changes were made to the items, the layout of the measure and the language on the basis of this feedback.

Final consultation

The draft measure was finally presented back to the original reference group for their comments.

Psychometric assessment

The framework for psychometric assessment was that reported for the Health Technology Assessment on adequacy of outcome measures which used 10 criteria, including feasibility and acceptability as well as the usual categories of reliability and validity.[73]

Samples

Service user participants for the feasibility study were recruited from acute wards and psychiatric intensive care units (PICUs), and test–retest participants were engaged on acute wards. For the larger psychometric testing phase, participants were recruited from acute wards. Service users were eligible if they could

provide informed consent and they had been present on the ward for at least 7 days during the 4-week data collection phase. All nursing staff from all grades (health-care assistants, entry-level qualified staff, clinical charge nurses and team leaders) were asked to participate in the psychometric testing of VOTE. All potential participants gave written informed consent following an explanation of the study.

Phase 3: feasibility and acceptability

The final measures were intended as self-report tools, so studies of feasibility and acceptability were conducted to evaluate the burden of administering and completing the measure. In the feasibility study, participants completed the measure including two additional questions assessing whether or not the measure was easy to complete and understand. Acceptability was assessed using two additional questions collected from participants. Two tests which indicate readability, the Flesch reading ease test (the recommended score is between 60% and 70%), were assessed.

Phase 4: psychometric assessment

Scoring

Scoring methods for each measure were decided at this stage.

Reliability

Test–retest reliability was assessed with participants who completed the measures on two occasions with an interval of 6–10 days. The test–retest reliability was assessed by Lin's concordance correlation coefficient,[74] kappa and proportion of maximum kappa[75] to measure the level of agreement between total scores and individual item responses at times one and two.

Generally, kappa scores of 0.21–0.4 indicate fair agreement, scores of 0.41–0.60 indicate 'moderate' agreement and scores > 0.60 indicate substantial agreement.[76] Lin's concordance correlation coefficient can be interpreted similarly to a correlation coefficient.

Internal consistency was assessed using Cronbach's alpha, which refers to the correlation between items on the scale. Low scores would suggest that the items are not contributing to the same latent construct, those with a score of ≥ 0.7 are considered acceptable.[77]

Face and content validity

Whether or not the new measures truly reflected the experiences of those receiving care (VOICE) and working (VOTE) in the acute inpatient services (face validity) and covered the full spectrum of views (content validity) was explored through participatory methodology during the instrument development phase.

Criterion validity

Criterion validity was assessed by testing hypothesised criteria against other relevant measures.

VOICE

We hypothesised that those with positive perceptions of acute inpatient services would also be satisfied with services. This was tested using the Service Satisfaction Scale – Residential Services Evaluation (SSS-RES).[78] This is adapted from the Service Satisfaction Scale-30 (SSS-30),[79] designed to evaluate residential services for people with serious mental illness. The original SSS-30 has been used in a variety of settings and demonstrates sound psychometric properties.[80]

VOTE

We hypothesised that those with negative perceptions of the daily ward pressures (VOTE) would have low levels of job satisfaction. This was tested using the Index of Work Satisfaction (IWS). This is a 44-item scale that can be totalled to produce a score that measures the level of job satisfaction for health professionals. Each item is answered on a seven-point Likert scale.[81]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Construct validity

The construct under study for each measure was explored using factor analysis, with both orthogonal rotation using varimax and oblique rotation using promax. The final solution was chosen from the scree plot and the factor structure was chosen from the analysis that produced the highest loading. This analysis followed preliminary checks on the strength of the pairwise correlations between items and Bartlett's test of sphericity.[82] The subscales produced were explored for internal consistency using Cronbach's alpha.[83]

Exploratory analyses

VOICE: Do demographic characteristics affect service user perceptions of the quality of acute ward care?

We expected differences in views between service users from different populations and clinical settings. Using one-way analysis of variance (ANOVA), we assessed whether or not service users' perceptions differed by borough, gender, ethnicity, age, diagnosis, admission and legal status. The majority of the analyses were exploratory. However, we had specific hypotheses relating to ethnicity and legal status, for which we expected poorer perceptions from participants who were compulsorily admitted and those from minority ethnic communities.

VOTE: Do demographic characteristics affect staff perceptions of the daily pressures of working on an acute ward?

A regression analysis with the VOTE total score as the dependent variable and all demographic characteristics as the key independent variables was carried out using random-effects regression modelling (clustering on ward) to take into account the multilevel nature of the data. The purpose of this model was to identify any significant demographic predictors of negative staff perceptions of the daily stressors in the working environment (VOTE). The demographic variables were length of employment (> 3 years or < 3 years), qualified nurse or not, education (degree level or not), ethnicity [white, or black or minority ethnicity (BME)], country of origin (UK or not), gender and age (median split at age 39 years).

Results

Phase 1: mapping out the dimensions of inpatient care

VOICE

The stakeholder group recommended five core topics:

1. the effect of acuity on service users
2. security issues
3. medication
4. the experience of admission and discharge
5. how ward routine is experienced.

VOTE

The stakeholder group recommended six core headings for the topic guide:

1. patient care
2. core interventions
3. teamworking
4. change
5. safety
6. ethical issues.

Phase 2: measure generation

VOICE sample

As *Table 1* shows, a total of 397 participants were recruited for the study: 37 for the measure generation phase and 360 for the feasibility study and psychometric testing. Schizophrenia and schizophrenia spectrum disorders were the most frequent diagnoses for both groups and approximately half of each set of participants was from the BME community. In the measure development phase, 43% of participants were men and the median age was 45 (range 20–66). In the psychometric phase, 60% of participants were men and the mean age was 40 (range 18–75).

TABLE 1 Demographic data of service users in WP1

Sociodemographic and clinical variable	Measure generation group, *n* (%) (*N* = 37)	Feasibility and psychometric assessment phase, *n* (%) (*N* = 360)
Ethnicity		
White	18 (48.6)	168 (47)
BME	19 (51.4)	185 (51)
Not disclosed	0 (0)	7 (2)
Legal status		
Formal	20 (54.1)	222 (62)
Informal	12 (32.4)	106 (29)
Not disclosed	5 (13.5)	32 (9)
Diagnosis		
Schizophrenia/psychosis	18 (48.7)	183 (51)
Bipolar affective disorder	7 (18.9)	51 (14)
Depression/anxiety	6 (16.2)	38 (11)
Personality disorder	2 (5.4)	19 (5)
Substance misuse	0 (0)	16 (4)
Other	4 (10.8)	46 (13)
Not disclosed	0 (0)	7 (2)
Employment		
Employed	0 (0)	62 (17.3)
Unemployed	0 (0)	248 (68.9)
Student	0 (0)	13 (3.6)
Retired	0 (0)	25 (6.9)
Other	0 (0)	5 (1.4)
Not disclosed	37 (100)	7 (1.9)
Admission		
First admission	0 (0)	65 (18.1)
Previous admissions	0 (0)	260 (72.2)
Not disclosed	37 (100)	35 (9.7)

Reproduced from Evans *et al.*[84] © 2012 Informa UK, Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

VOTE sample

A total of 376 nursing staff participated in the measure development (*Table 2*). The majority of nurses were qualified nursing staff, while approximately one-third were health-care assistants. In the psychometric testing phase, 47% of participants were women and the median age was 39 years (range 20–67 years).

Thematic analysis of the full data set resulted in an initial bank of 34 items, which were formed into brief statements and grouped into domains. A six-point Likert scale was chosen, ranging from 'strongly agree' to 'strongly disagree' and optional free-text sections were included to capture additional qualitative data. The items were unweighted and one question was reverse scored. The self-report measure was designed to provide a final total score, with a higher score indicating a more negative perception. The inter-rater reliability of the focus group data coding was assessed using an NVivo 7 (QSR International, Warrington, UK) coding comparison report and showed between 97% and 99% agreement. Item reduction based on relevance and preventing duplication produced 22 items.

The pilot study, which comprised a mixed group of health-care assistants and qualified staff, agreed that these topics were broad enough to allow flexibility in bringing in new ideas. Interim analysis of the pilot focus group suggested that occupational seniority might interfere with a full discussion so, in order to allow for the maximum emergence of key themes, one 'health-care assistant only' group was included.

Thematic analysis revealed that the staff participants found the following core themes the most important for inclusion on the measure: teamworking, patient care, core interventions, safety, bed management and continuing professional development. These themes provided the structure for the measure. The individual items were developed from sub-themes within these broader domains. A five-point Likert scale was reviewed at the expert panel stage, but those staff participants felt that a wider range of response options would provide more scope and reduce omissions. At this stage there were 26 items.

TABLE 2 Samples of individuals in the VOTE development in WP1

	Focus/reference group, *n* (%) (*N* = 48)	Feasibility group, *n* (%) (*N* = 40)	Test–retest group, *n* (%) (*N* = 43)	Psychometric group, *n* (%) (*N* = 245)
Staff				
Health-care assistants	16 (33)	23 (33)	14 (33)	72 (29)
Staff nurses	23 (48)	18 (45)	16 (37)	100 (41)
Clinical charge nurses	9 (19)	7 (18)	8 (19)	44 (18)
Team leaders	0	2 (5)	5 (12)	17 (7)
Ethnicity				
White British	13 (27)	15 (38)	16 (37)	67 (27)
BME	35 (73)	25 (63)	27 (63)	178 (73)
Gender				
Female	33 (69)	19 (48)	23 (54)	116 (47)
Male	15 (31	21 (53)	20 (48)	111 (45)
Age (years)				
Mean (SD)	37 (9.8)	38 (8.9)	38 (8.9)	39 (9.6)
Range	21–58	22–55	24–61	20–67

SD, standard deviation.
Reprinted from the International Journal of Nursing Studies, volume 49, Laker C, Rose D, Flach C, Csipke E, McCrone P, Craig T, *et al.*, *Views of the Therapeutic Environment (VOTE): stakeholder involvement in measuring staff perceptions of acute in-patient care*, pp. 1403–10, 2012,[85] with permission from Elsevier.

Themes from the focus groups

The main themes identified in the focus group data were staff/patient interaction and coercion. Often, staff and service users conceived these differently.

Staff and service users' interactions on acute wards

Service users In terms of the focus of group discussions, the overwhelming perception of service users was that the ward was 'untherapeutic'. A crucial contributor was the lack of available staff and helpful staff. For patients, it appears that staff are 'stuck in the office' and cannot help even when the situation is urgent, the situation is escalating or the service user is in crisis. This lack of availability was compounded by the use of 'bank' or 'agency' staff who were brought onto the ward in times of staff shortage but were usually unfamiliar with the ward and did not know the patients. They were described as 'sitting and staring' and not prepared to have a conversation with users or even ask them how they were feeling.

In the UK, every service user should have a care plan and documented notes on meeting that care plan. In fact, this is the main form of 'structured interaction' that is meant to take place on wards. But even care planning was low in reference density in the data. One of our participants had been given his discharge summary which included this documentation. It implied that he had one-to-one conversations with a named nurse every day throughout his stay; however, this was not his memory of events.

Many of the service users understood that work pressures such as under-staffing, paperwork and work-related stress impacted on staff's lack of interaction. Nevertheless, this resulted in an untherapeutic climate where emotional distress was suppressed rather than explored.

It is evident in the accounts that routine and therapeutic interaction was rare in ward life. Participants were more positive about talking therapies, although they gave instances of them being unhelpful, especially groups in which everyone, including the facilitator, sat mute for an hour. Although talking therapies were described as being helpful, very few people had actual experience of such therapy while on the ward. They reported in prospect that talking therapies would be a good way to deal with distress. One participant had observed a suicidal fellow patient and thought that if there had been group therapy available she might have got better. Another favourably contrasted talking therapies with the ever-present use of medication.

Nevertheless, there were instances when the service user recognised that the ward had done them some good or, at least, were ambivalent about it.

Nurses The nurse participants were aware that the care they were providing was less than perfect and that they did not spend enough time with patients. They perceived both administrative tasks and bed management as a barrier to interaction and the 'pressure' of these daily tasks was very frequently referenced. Instability as a result of both the distress of individual service users and the overall ward milieu were of concern.

Bed management, which involved early discharge, making space and moving clients around, provoked a strong emotional response among nurse participants as they knew it led to poor interactions with patients. The practice of 'sleeping out' when a patient spent the day on the ward and then was sent to another ward, or even a bed and breakfast, to sleep evoked angry responses. Using periods of leave to free up beds also provoked feelings of guilt and anxiety alongside the view that senior management did not care.

Nurses had complex emotional responses to the fact that they could not spend enough time with service users. They expressed frustration and concern that they were becoming deskilled. Anxiety, arising from the lack of interaction, was a common response. Explicitly, staff remained committed to the need for interaction despite the daily pressures of working in an acute setting. Implicitly, there was an underlying theme of avoidance or interaction anxiety which suggested that many staff had withdrawn from interaction with the service users. The main cause of this 'burnout' appeared to arise from limited internal

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

coping skills and from the need for staff to protect themselves emotionally from the complexities of individual service users in their care.

Coercion and control

As a barrier to therapeutic interaction and within the wider theme of violence, we will illustrate how service users and staff perceive their own power to be manifested on the ward. The two groups have differing perspectives: users feel coerced, whereas staff feel they are delivering a legitimate response to violence.

Service users For service users, coercion is a complex concept. Here we focus on control, restraint and forced medication (rapid tranquillisation), which had high reference density in the data and our patient participants devoted much time to this theme. Medication was the main (or sole) treatment but here we are concerned with medication that is given by force. Other countries use mechanical restraint but this is not deployed in the UK.

One of the main reasons given by patient participants for behaviour that might elicit restraint or forced medication was that users were cooped up in the ward and not allowed to go outside and get fresh air. Even those granted time off the ward would often find themselves without a suitable escort. Others put it more strongly, likening the environment to a prison or a cage for an animal. This, they said, provoked extreme frustration and anger which was responded to by nurses in a way they thought aggressive and unnecessary. Our participants conceptualised forced medication as violence and had a whole vernacular to describe it: 'jump on you with the needle' or 'pound on you with the needle'. However, they did not see their own behaviour as unnatural. For these participants, forced medication was not medical treatment but control.

Coercion can be experienced in more subtle ways and examples were common in the data. Medication does not always have to be given forcibly, there can be pressure to take medication under the threat that if the patient does not then it will be given by force.

There is a dichotomy in our data. Inpatient wards were routinely described as unsafe, fearful places where users felt unprotected by staff. Perceptions of coercion regarding control, restraint and medication resulted mainly from 'being done to'. However, within the context of an often chaotic ward, coercion was sometimes perceived as an appropriate staff response to other violent service users.

Nurses Both service users and nurses reported a sense of 'them and us' on the ward. In the management of violence, nurses used interventions that were perceived as coercive by service users. However, nurses expressed some awareness and sympathy towards service users around this issue and referred to the therapeutic relationship as a protective factor in the management of violence.

At the same time, it was clear that control was not absolutely one sided in favour of nurses. They reported that violence as well as intimidation was a tool that service users could use to express their resentment at their situation on the ward. The lack of support felt by nurses in the face of violence was strongly expressed in the data. This is despite the provider trust having a 'zero tolerance' policy with respect to abuse, which would imply that police intervention was sometimes warranted. Nonetheless, many felt that they had to cope alone.

There is no doubt that some nurse participants had experienced very serious incidents. As well as physical assault, two nurses had been taken hostage and locked in the nursing office for many hours. Attempts to have the police intervene were of no avail as the police saw it as a 'mental health issue'. We asked the nurses about logging incidents and their response was that they logged only the most serious ones otherwise they would spend the whole day filling out incident reports.

Official documents from the Royal College of Psychiatrists and the Royal College of Nursing (RCN) in the UK emphasise that the initial response to an imminent incident should be one of de-escalation where the staff member tries to 'talk down' the patient.[86] The reference density of de-escalation in our data was tiny with only two chunks of text coded. One person tried to explain why de-escalation was impossible owing to lack of staff; however, there was one instance where de-escalation was highlighted as best practice.

The perceptions of users, nurses and doctors on the issues of managing aggression and violence have been compared.[87,88] Staff attributed some violent incidents to factors 'internal' to service users, such as their illness or demographic characteristics. Service users, on the other hand, found responses to 'incidents' to be unsympathetic, containing and controlling. This is consistent with other findings.[89]

Expert panels

VOTE

The expert panels ($n = 13$) confirmed that the measure allowed a good range of staff perceptions of acute inpatient wards to be expressed. A five-point Likert scale was reviewed at the expert panel stage, but those staff participants felt that a wider range of response options would provide more scope and reduce omissions. At this stage there were 26 items.

VOICE

The expert panels considered the measure to be an appropriate length and breadth and, following some minor changes in wording, concluded that the measure was appropriate for use by service users in hospital.

Final consultation

The service user reference group requested that the wording of several items be clarified, that all statements should be made positive, that staff responses to emergencies and personal crises should be differentiated and that the measure should not begin with items about the staff. There were no changes to the nurses' measure as a result of the final consultation.

Psychometric assessment

VOICE: feasibility and acceptability

Feasibility took place in two waves ($n = 40$ and $n = 106$). In the first wave, 98% of participants found the measure both easy to understand and complete, and in the second wave, 82% of participants considered the measure to be an appropriate length. Two participants (2%) disliked completing the measure and six (6%) found some of the questions upsetting. VOICE took between 5 and 15 minutes to complete and was easy to administer. The measure was found to be suitable for completion by participants with a range of diagnoses and at levels of acute illness found on inpatient units. The Flesch reading ease score was 78.8 (reading ages 9 and 10 years) indicating the measure was easy to understand.[90] Following the feasibility study, one item was removed as it was considered to be a duplicate. This left the measure with 21 items.

VOTE: feasibility and acceptability

The feasibility study ($n = 40$) revealed that 95% of staff agreed that VOTE was easy to complete and easy to understand. Generally, staff found that, with minimal explanation, the measure could be completed by self-report in approximately 15 minutes. Items identified as having confused phrasing were changed and those showing poor consistency (Cronbach's alpha) or poor variability or poor reliability were dropped, leaving 20 items.

The acceptability study showed that 76% of staff thought that the length was about right and 91% that it was enjoyable and not upsetting. The Flesch reading ease score for the measure was 64%. The Flesch–Kincaid[91] grade level score was 7.6, which means that the measure can be read by a 12-year-old who is of average ability.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Phase 4: psychometric assessment

VOICE scoring

Total scores were calculated by totalling all items with no missing data. For participants responding to at least 80% of the items, a pro-rated score was calculated. Less than an 80% response was considered as a missing total VOICE score. Negatively phrased items were reverse scored so that high scores indicated a more negative perception of the quality of care.

VOTE scoring

Total scores were calculated by totalling all items with no missing data. For participants responding to at least 80% of the items, a pro-rated score was calculated. Less than an 80% response was considered as a missing total VOTE score. Negatively phrased items were reverse scored so that higher scores indicated a more negative perception of the ward. When comparisons to other measures were made, the same rule was applied to their scoring, that is, high scores indicate poorer satisfaction.

VOICE reliability

The test–retest reliability ($n = 40$) was high [ρ 0.88, 95% confidence interval (CI) 0.81 to 0.95] and there was no difference in scores between the two assessments. A total of 192 participants had complete data on the VOICE scale and were used in assessing the internal consistency. After removing items with poor reliability, this left a 19-item scale with high internal consistency ($\alpha = 0.92$).

VOTE reliability

Group 2 ($n = 43$) participated in the test–retest study using the final 20-item measure. As items tended to be skewed towards 'agree' or 'disagree', the kappa coefficient is expressed as a proportion of the maximum possible value. Six items showed substantial agreement (kappa maximum ranged from 0.60 to 0.73). Moderate agreement was shown in 14 items with kappa maximum ranging between 0.41 and 0.59. These results indicated moderate and substantial reliability. Concordance between the total scores ($n = 34$) was good (total score: ρ 0.77, 95% CI 0.65 to 0.89). The internal consistency of the measure, assessed using data from group 3 ($n = 200$), was good with the overall alpha at 0.82.

VOICE validity

The measure has high face validity. The wide range of items was determined by service users during the focus groups and the measure reflected the domains that they considered most important. The feasibility study participants felt that the measure was comprehensive and, therefore, had high content validity. Pearson's correlation coefficient showed a significant association between the total scores on VOICE and the SSS-RES ($r = 0.82$; $p < 0.001$), indicating high criterion validity.

The data were suitable for factor analysis for construct validity [Kaiser–Meyer–Olkin measure was high (0.9), Bartlett's Test of Sphericity was significant ($p < 0.001$)]. The exploratory factor analysis was conducted on 192 fully completed responses to the 19-item measure. A promax rotation, allowing the subscales to correlate, revealed two factors (*Table 3*) that accounted for 95% of the total variance. The item groupings indicated by the factors were used to determine subscales by summing the scores of the items. Items with a factor loading of > 0.4 were considered to load significantly. When items loaded onto both subscales, a decision was made to include them on the subscale with the closest conceptual fit. Subscale 1 was labelled security and subscale 2 was labelled care. Both factors were significantly correlated ($r = 0.73$; $p < 0.001$). One item ('Staff give me medication instead of talking to me') did not load highly onto either scale and so was not included in the subscales.

VOTE validity

A high level of staff involvement throughout the process of measure development ensured good face and content validity. This was achieved because staff participants provided feedback on the content of the themes arising from the qualitative data and on the language used in the item generation phase. Staff agreed that the results did capture what they had reported. The use of a flexible topic guide/interview

TABLE 3 Factor analysis of VOICE in WP1

Item	Item description	Subscale 1: security	Subscale 2: care
17	I feel safe on the ward	0.797	–0.11
21	I think staff respect my ethnic background	0.788	–0.026
9	I trust the staff to do a good job	0.765	0.065
11	I feel that staff treat me with respect	0.705	0.152
18	I feel staff respond well when the panic alarm goes off	0.641	–0.053
19	I feel staff respond well when I tell them I'm in crisis	0.592	0.223
20	I feel able to practice my religion whilst I'm in hospital	0.585	0.042
1	I was made to feel welcome when I arrived on this ward	0.414	0.343
15	I find it easy to keep in contact with family and friends	0.403	0.165
3	Ward rounds are useful for me	–0.095	0.855
5	I have the opportunity to discuss my medication and side effects	0.099	0.661
2	I have a say in my care and treatment	0.079	0.659
13	I find one-to-one time with staff useful	–0.078	0.634
4	I feel my medication helps me	0.072	0.623
12	I think the activities on the ward meet my needs	0.102	0.565
7	Staff take an interest in me	0.209	0.499
10	I feel that staff understand how my illness affects me	0.325	0.486
8	Staff are available to talk to when I need them	0.378	0.426
Cronbach's alpha		0.88	0.88
Mean (SD)		23.6 (9.8)	25.7 (9.7)

SD, standard deviation.
Shading refers to the item loadings considered for each of the two factors.

schedule maximised exploration of the construct under study and minimised omissions in the data set. A *t*-test examining whether or not VOTE scores were associated with an a priori criterion of job satisfaction, as measured by the IWS, showed that those with negative perceptions of the daily pressures of working on an acute ward also had poor job satisfaction [$t(157) = -10.34$; $p = 0.001$ ($n = 159$)]. The mean VOTE score in the low job satisfaction group was 76.5 [standard deviation (SD) 10.4] and the mean VOTE score in high job satisfaction group was 60.4 (SD 10.1). The VOTE measure was strongly and significantly correlated with the IWS [$r = 0.77$; $p < 0.001$ ($n = 159$)]. Overall, the mean satisfaction of the group was 163.6 (SD 31.3).

An exploratory factor analysis for construct validity was carried out ($n = 200$). The Kaiser–Meyer–Olkin measure of sampling adequacy was good (0.8) and Bartlett's test of sphericity was significant ($p \leq 0.001$). The scree plot[92] indicated three factors with an eigenvalue of > 1. The orthogonal rotation produced the most coherent solution because it had consistently higher factor loadings and this is presented in *Table 4*.

Factor 1 reflected staff perceptions of 'workload intensity' and appeared to be tapping issues relating to work pressures, administrative intensity and feeling unsupported (either through weak leadership or inadequate staffing). Factor 2 related to 'team dynamics' linking concepts such as staff confidence in their leadership and in the dynamics of their team. Factor 3 identified 'interaction anxiety', associated with powerlessness and a lack of confidence in the uncertain ward situation. Evaluations of the internal consistency of the subscales: factor 1 'workload intensity' (Cronbach's alpha 0.74), factor 2 'team dynamics' (Cronbach's alpha 0.75) and factor 3 'interaction anxiety' (Cronbach's alpha 0.60).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 4 Factor analysis of VOTE in WP1

		Factor loadings		
		Eigenvalue 4.18	Eigenvalue 1.22	Eigenvalue 0.98
VOTE items		Factor 1	Factor 2	Factor 3
1	It is easy to balance documentation/paperwork and spending time with the patients on my ward	0.56		
2	I feel pressured to complete tasks in my job	0.46		
3	Other than mandatory training, staff development opportunities are limited	0.3		
4	I think that the senior managers (above ward managers) understand the current realities of working on acute wards	0.62		
5	Finding enough staff to cover shifts is easy on my ward	0.57		
6	On my ward there is immense pressure to create bed space	0.31		
7	When it comes to bed management the clinical perspective of my team is always considered	0.54		
8	There are enough staff to maintain safety on my ward	0.62		
9	Patients are provided with enough information about their medication on my ward		0.59	
10	If I have concerns about patient care I am happy to address it with colleagues		0.54	
11	I benefit from regular supervision		0.41	
12	I benefit from strong leadership on my ward		0.51	
13	When it comes to patient care there are staff in my team who have a 'can't do, won't do' attitude		0.34	
14	I find that communication between the different Multi-Disciplinary Team professionals is consistently good		0.41	
15	There is a strong emphasis on promoting a sense of team spirit on my ward		0.58	
16	Patients can feel that there is a sense of 'them and us' on my ward			0.47
17	When I ask patients' to join in with activities, they say they are not interested in those on offer			0.4
18	I worry about violence and aggression when at work			0.68
19	Decisions that are made on one shift are changed on the next which makes consistency difficult in my team			0.39
20	I'd rather not address relationship issues between teammates because it will create a bad atmosphere			0.37

Exploratory analyses: how effective are the new measures in discriminating between groups?

VOICE: do demographic characteristics affect service user perceptions of the quality of care on an acute ward?

The ability of VOICE to discriminate between groups is indicated in *Table 5*. Bivariate analyses showed significant differences for legal status. Participants who had been compulsorily admitted had significantly worse perceptions ($t = -3.82$; $p < 0.001$). A multivariate analysis revealed that legal status remains significant even when adjusted for the other factors ($p = 0.001$).

TABLE 5 Differences in mean VOICE scores by demographic and clinical group in WP1

Variables considered	Number	Mean score	SD	95% CI	*p*-value
Gender					
Male	199	55.5	19.2	52.8 to 58.1	0.15
Female	147	52.5	17.8	49.6 to 55.4	
Ethnicity					
White	162	55.6	19.1	52.6 to 58.5	0.22
BME	180	53.1	18.1	50.4 to 55.7	
Legal status					
Informal	102	48.9	16	45.7 to 52.0	< 0.001
Formal	215	57.4	19.6	54.7 to 60.0	
Borough					
Borough 1	132	54.5	19.5	51.2 to 57.8	0.15
Borough 2	100	52.7	18.5	49.1 to 56.4	
Borough 3	75	57.8	17.7	53.8 to 61.8	
Borough 4	40	50.1	16.9	44.9 to 55.4	
Diagnosis					
Schizophrenia/psychosis	179	54.8	18.2	52.2 to 57.5	0.40
Bipolar affective disorder	51	56.1	21.5	50.2 to 62.1	
Depression/anxiety	38	50.9	13.8	46.5 to 55.3	
Personality disorder	18	59.3	16.6	51.6 to 67.0	
Substance misuse	13	53.7	18.9	43.4 to 64.0	
Other	42	50.3	21.3	43.8 to 56.8	
Age (years)					
≤ 20	15	61	23.7	49.0 to 73.1	0.29
21–30	77	51.6	15.5	48.1 to 55.1	
31–40	93	53.5	16.7	50.1 to 56.9	
41–50	87	55.6	21.9	50.0 to 60.2	
51–60	45	57.2	19.1	51.6 to 62.8	
≥ 61	25	50.6	18.3	43.4 to 57.9	
Admission					
First admission	62	50.7	17.6	46.3 to 55.2	0.11
Previous admissions	255	55	18.9	52.6 to 57.3	

We assessed any independent demographic predictors of VOTE using random-effects modelling and a forward selection procedure with the level of significance set at 0.05. The results showed that one variable was a significant predictor: country of birth ($n = 180$, groups = 17).

Staff who were born in countries outside the UK had more positive perceptions of the daily pressures of working on an acute ward than those who were UK born (*Tables 6* and *7*).

The final measures are available are in *Appendices 7* and *8*.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 6 Mean VOTE scores by demographic and clinical group in WP1

Variables considered	Mean VOTE scores (SD)
Staff	
Whole sample	68.54 (12.77)
Health-care assistants	67.25 (12.55)
Qualified nurses	70.81 (13.33)
Clinical charge nurses	66.55 (11.46)
Team leaders	64.23 (10.51)
Ethnicity	
White British	70.43 (13.53)
Other	–
BME	67.74 (12.57)
Gender	
Male	67.63 (12.18)
Female	69.58 (12.93)
Age (years)	
≤ 39	67.93 (13.94)
≥ 40	69.17 (11.69)
Country of origin	
UK	70.01 (13.00)
Non-UK	66.96 (12.95)
Length of employment	
≥ 3 years	69.26 (12.98)
< 3 years	67.25 (12.93)

TABLE 7 VOTE scores adjusted for country of birth (WP1)

Variables	Coefficient	SE	*p*-value	95% CI
Country of birth not UK (baseline = UK)	–4.03	1.82	0.03	–7.59 to 0.46
Between-ward variability (SD)	5.08			
Residual variance (SD)	11.61			
Intraclass correlation	0.16			

SE, standard error.

Discussion

Using a participatory methodology, we have developed both service user- and staff-generated, self-report measures of perceptions of acute care. VOICE encompasses the issues that service users consider most important, namely their care and treatment and feelings of security, and is suitable for use in research settings. VOTE identifies the daily hassles of staff that affect staff engagement with service users, colleagues and their professional identity, which are all areas that are important to stakeholders. Both measures have promising psychometric properties. In both measures, the internal consistency and criterion validity are high and test–retest data show stability over time. The full involvement of service users and staff throughout the development of each measure has ensured that VOICE and VOTE have good face and content validity and are accessible to the intended recipients.

Can VOICE and VOTE distinguish differences in views?

Detained participants were shown to hold more negative perceptions of inpatient services. This is similar to the satisfaction studies showing that service users who are admitted involuntarily are less satisfied with their care.[93] More recently, research has identified that lower levels of satisfaction are linked with the accumulation of coercive events and perceived coercion.[94,95] Our findings from the focus groups present a more complex picture and one that is worth further analysis.

We anticipated, but did not find, any differences on either VOICE or SSS-RES scores for ethnicity. Previous quantitative studies have shown differences for legal status but not ethnicity,[96,97] whereas qualitative research has revealed that BME users hold relatively poor perceptions of acute care.[16,98] Our study was set in areas of London with high proportions of people from BME groups.[99,100] Staff demographics tended to mirror those of inpatients and it may be that services were better tailored towards BME groups. In addition, asking users for their perceptions while they were in hospital may have inhibited openness and honesty, particularly on sensitive issues.

In this study, staff had slightly more positive perceptions of job satisfaction than staff in other studies.[101–103] Staff with negative perceptions of the daily pressures of working on an acute ward and barriers to change also had poor job satisfaction.

In previous studies, work stressors have differed between groups including occupational status and gender.[43,48] Of all the demographic characteristics included in this study, only country of birth showed a significant effect on staff perceptions of the daily pressures of acute ward working, with staff born outside the UK showing more positive perceptions than those born within the UK. This may be a direct result of the inner London location of the trust under study, as a large percentage of the client group and the nursing workforce are not UK born.

However, the mean VOTE scores were most negative overall in the direct care staff groups (staff nurses/nursing assistants) compared with more positive scores in those who occupied more managerial roles (clinical charge nurses and team leaders). Although we did not address whether or not the stressors differed significantly between these groups, it is true that both nursing assistants and staff nurses spend the most time in direct client contact. Therefore, their more negative perceptions might be one of the inevitable aspects of delivering therapeutic acute care for patients with severe mental illness at their most distressed. Furthermore, staff in leadership roles are more likely to be involved in the planning stages of new changes and, therefore, have an increased sense of control and responsibility over them, while those in direct care roles are more likely to be involved in delivering changes. These important issues, revealed through stakeholder involvement, require further exploration to discover key drivers for these perceptions that might then be subjected to management and other interventions.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Are VOICE and VOTE different from other measures?

The benefits of the participatory model used to develop VOICE and VOTE are visible in both the breadth of the construct investigated and the rich content of the items. Although the VOICE total scores were correlated, there were also distinct differences in content between VOICE and the comparison satisfaction measure. We believe that this is because of the use of a participatory methodology. The factor analysis of VOICE discovered two main factors: security and care. In particular, the safety and security issues were given more weight in VOICE and items on diversity were included that did not appear in the conventionally generated measure. Although items regarding the physical environment and office procedures featured in the SSS-RES,[78] these issues were not deemed as important by the service users in our study and, therefore, were not included in VOICE. However, other items referred to 'care received' overlapped with some items in SSS-RES.

It is often assumed that the only construct to measure is satisfaction with acute care. However, there are difficulties in encapsulating complex sets of beliefs, expectations and evaluations in satisfaction measures. Caution should be taken when making inferences from the results of such measures, as they may not accurately reflect the views of service users.[104] VOICE is unique in that it captures service users' perceptions and we anticipate that this will depict the inpatient experience more accurately.

The factor analysis conducted during the development of VOTE revealed three novel factors. Factor 1, 'workload intensity', seems consistent with previous work identifying the effects of poor resources such as inadequate staffing and too few beds on occupational stress for nurses.[9] Factor 2, 'teamwork', is similar in content to the concept of 'job satisfaction' characterised by leadership, educational attainment, pay and stress. Factor 3 represents a new subscale that explains the lower correlations to the IWS and Maslach Burnout Inventory (MBI) subscales. It crystallises two ideas. First, it reflects some issues relating to burnout.[46] Second, it draws together some of the feelings of fear and uncertainty that are associated with working with distressed clients in an unstable environment. This complex issue was revealed through stakeholder involvement and adds a new and unique aspect to the current library of measures that capture staff perceptions of stressors in acute inpatient settings. This issue clearly requires further exploration in the field, as one of the fundamental aspects of delivering therapeutic acute care is in the ability to work with clients with severe mental illness at their most distressed.

Further developments

Three further measures grew out of this WP; please refer to *Appendices 1–3*.

Strengths and limitations

This study was conducted in London boroughs with high levels of deprivation, ethnic minority populations and psychiatric morbidity[99,100] and, therefore, may not be directly generalisable to other settings. In addition, our focus groups and expert panels included a high proportion of participants from BME communities. Although this is a strength, it may be that different items would have been produced by other groups. Furthermore, as recruitment for VOTE was limited to staff from acute wards it is not clear whether or not VOTE would be suitable for other settings such as forensic or rehabilitation wards.

It is impossible to accurately assess inpatient care without involving the people directly affected by that service. Developing outcome measures that are valued by service users and staff is essential in evaluating and developing inpatient services. The main strength of this piece of research is that it fully exploits a participatory methodology: staff and service users were involved in a collaborative way throughout the whole research process. VOICE and VOTE are the only robust measures of acute inpatient services designed in such a way. This method has resulted in measures that encompass the issues that staff and service users consider most important and that are acceptable, easy to understand and complete by a diverse staff group and service users with a range of diagnoses and severity of illness.

Conclusion

The study has demonstrated that a participatory methodology can generate items that are prioritised by staff and users, but not included in traditionally developed measures. VOICE is the first service user-generated and psychometrically robust measure of perceptions of acute care. It directly reflects the experiences and perceptions of service users in acute settings and, as such, is a valuable addition to the PROMs library.

VOTE can be used to promote staff involvement by identifying staff-valued work pressures that might affect burnout and interfere with quality of care, as well as being used to highlight the impact of service changes in acute wards. In future research, the effects of the ward environment on the workforce might be evaluated through an exploration of staff perceptions. Whether or not VOTE can also be linked to environmental factors such as incidents, staffing levels, and the amount of time that staff spend engaged in therapeutic interaction with clients should be explored. VOTE might be a useful tool for staff to assess the feasibility of delivering new interventions and exploring whether or not they can be sustained in the long term.

Clinical implications

VOICE and VOTE are measures of the experience of living and working on acute wards. This is consistent with recent National Institute for Health and Care Excellence (NICE) guidelines 2014.[104,105] The factor structures demonstrated what was key in this experience for each group. Service users were concerned with the care provided and staff/patient interaction as well as security on the wards. Nurses were concerned with workload intensity, team dynamics and interaction anxiety. Although interaction anxiety could lead to an avoidance of service users and, therefore, compromised care, there was evidence from the post-intervention focus groups for WP3 that service users understood how pressured staff were and made allowances for it. These findings suggest that leaders in the NHS need to consider putting in place support structures for staff if care is to be optimised for service users. The 2013 Care Quality Commission (CQC) report[106] paints a mostly bleak picture of inpatient care which corresponds to our findings and they call for similar measures. The RCN ran a consultation on control, restraint and rapid tranquilisation practices in 2013, which greatly exercised our focus group participants.

Research implications

It would be worth investigating more thoroughly the variation in VOICE and VOTE and what such variation might predict. Furthermore, a more detailed investigation could be undertaken for the WP4 data. We need to find out whether or not, within the triage system, the admission ward had different VOICE and VOTE configurations (total scores and individual factors) from the long-stay wards as well as looking at this over time. We also have test–retest data from a forensic ward and these could be further examined in terms of the VOICE and VOTE scores to investigate if there are differences between wards. Other ongoing research is examining whether or not VOICE scores are sensitive to a peer support intervention on acute wards.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 3 Work package 2: CITRINE

This chapter reports findings from the health economic component of the PERCEIVE programme. It consists of (1) systematic review of studies that have investigated staff activity on psychiatric inpatient wards and particularly the time staff spend with patients, (2) systematic review of economic evaluations of therapeutic interventions delivered in inpatient settings, (3) demonstration of simulation modelling methods to provide such evidence, (4) development of a patient-based questionnaire to measure staff contacts and activities on inpatient wards and (5) results from this questionnaire in terms of staff and activity costs and predictors of these costs.

Systematic review of studies measuring staff–patient contact time on psychiatric inpatient wards

Greater staff–patient interaction and greater patient activity (i.e. participation in therapeutic or social activities rather than being socially disengaged or spending time alone) in inpatient wards improves clinical outcomes for patients with mental illness.[107] The changing composition of UK mental health-care service provision in recent years has put a strain on acute mental health resources, which has affected how nurses interact with and treat patients. This is mirrored in most western health-care systems. The greater emphasis and spending on community-based mental health care has also reduced the portion of the mental health-care budget allocated to inpatient psychiatric care, leading to a reduction in the number of acute care and long-stay beds, although the amount spent on inpatient care remains substantial.

The average daily number of available psychiatric beds in England dropped from 67,122 in 1997–8 to 26,929 in 2007–8.[108] Admitted patients are more likely to have psychotic disorders and in 2006/7 the median length of stay for these patients was 42 days. At a cost of £268 per bed-day of acute inpatient psychiatric care,[109] this average length of stay would cost £11,256.

Given the reduction in acute beds, priority naturally goes to those patients who are more severely mentally ill, particularly for those who have been involuntary admitted or have major social problems.[110] Nurses report that they feel pressure to discharge patients who may not have fully recovered in order to free up beds.[13] Owing to the severity of the patient mix, staff members find it difficult to provide interaction and therapeutic activities as much of their time is spent looking after very ill patients.

In a 2004 survey by the Sainsbury Centre for Mental Health,[111] it was found that although 64% of wards routinely offer social and leisure activities (e.g. going to the gym, bingo) and 73% offered practical therapeutic activities (e.g. money management, cooking skills), psychological therapies were not routinely offered on the majority of acute psychiatric wards. Art therapy was the most commonly available therapy, available in 49% of wards, compared with 35% for psychosocial interventions, and less than 20% for cognitive behavioural therapy, solution-focused behavioural therapy and family therapy. Besides the difficulty of staff having the time to offer activities because they are occupied with taking care of very ill patients, the reports suggest that the lack of activities and talking therapies may be partially explained by the use of agency staff, the culture of the psychiatric wards and staff shortages, especially of occupational therapists (OTs) and psychologists.

Assessing both the extent of interaction between patients and staff, and patients' participation in activities, may be important for both clinical and economic reasons. Therefore, it is helpful to determine the extent of this input. This might help in planning for the costs of specific units by determining the appropriate amount and grade of staff members. Nursing and other professional time comes with a cost and while most studies and payment systems currently focus on the average cost of an inpatient day, there are reasonable grounds for deriving per-patient inpatient costs. These grounds include the requirement in economic evaluations that costs reflect the actual resources used. A further, linked, reason for using actual

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

per patient costs is that any form of prospective payment (such as 'Payment by Results' in the UK) should link prices/tariffs as closely as possible to care received.

Most of the high costs of inpatient care are attributable to personnel costs. Beecham *et al.*[112] reported that staff costs account for 70% of the costs of child and adolescent psychiatric inpatient units. Similarly, McKechnie *et al.*[113] found that the percentage of costs attributable to staffing varied by the type of unit, but ranged from 74% to 91% of total care and treatment costs of psychiatric wards.

Collins *et al.*[107] have suggested that clinical outcomes are associated with patient activity and social interaction. If this is the case, then devoting more staff time to direct (and hopefully therapeutic) patient activities and care could have a positive effect on patient care and may be cost-effective. Indeed, a study by Dodds and Bowles[114] found that changing the allocation of nursing time from formal observation of inpatients to structured, individualised activities, which often took place on a one-to-one basis, improved patient-reported quality of care while the number of staff sicknesses and staffing costs decreased. The authors estimated that in a best-case scenario, an inpatient ward could save up to £44,878 on staffing costs over a year if they adopted this change in nursing time allocation. Furthermore, research indicates that, for patients, the most important aspect of quality inpatient care is that staff members take their time with, and care for, patients.[115] Of course it should be stressed that staff time is not synonymous with therapeutic input. Merely spending more time with patients might not improve outcomes. However, given the above findings it does seem logical to expect improvements to be linked with what staff and patients do on wards and as such recording staff and patient activities seems a reasonable exercise.

The aims of this review are to (1) identify studies that have measured nursing and patient activity and interaction on psychiatric inpatient wards, (2) compare methodologies used in these studies and (3) examine how much time is typically spent in staff–patient interaction. Finally, we make suggestions for future work based on the findings.

Methods

Reprinted from International Journal of Nursing Studies, Vol. 47 Issue 7 Sharac J, McCrone P, Sabes-Figuera R, Csipke E, Wood A, Wykes T, Nurse and patient activities and interaction on psychiatric inpatients wards: a literature review, pp. 909–17 Copyright (2010), with permission from Elsevier.[15]

We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines by setting up a search strategy as well as independently checking papers for inclusion and searching reference lists. Initially, the electronic databases EMBASE (1980–2008), MEDLINE (1950–2008) and PsycINFO (1806–2008) were searched with the following keywords: 'inpatient', 'mental', 'psychiatric', 'hospitaliz(s)ed', hospitaliz(s)ation', 'time' and 'activity'. A final check of the literature was made on 30 November 2009. Included studies were limited to those published after 1970 that were written in English and reported activity on psychiatric inpatient wards. Abstracts were reviewed by JS and PM to assess eligibility for inclusion. The reference lists of included studies were checked by PM for further relevant literature. The criteria for inclusion were that (1) the study was based in an inpatient setting and (2) that it measured or observed and recorded time spent by patients and/or staff on different activities. Given that studies differed substantially in their methods and contexts, we did not attempt to combine the results in a formal way. In addition, given the study heterogeneity we did not have predetermined measures of quality but we do discuss study characteristics in the text.

Results

The search strategy identified 86 papers of potential relevance. After reviewing abstracts, two studies were initially included. A further 11 studies were included following examination of reference lists. Study characteristics are summarised in *Table 8*.

TABLE 8 Methods of reviewed time and motion studies (WP2)

Author and year	Setting	Participants	Methods	Activities reported
Bee *et al.* (2006)[116]	Three acute psychiatric inpatient wards, England	• 15 registered nurses • One student nurse • 24 unqualified nursing assistants	• Staff were interviewed on an hourly basis over three day shifts at five points in time and reported all activities undertaken	• Administrative tasks • Communication • Domestic tasks • Staff breaks • Patient contact (containment, social care, social interaction, medical care, therapeutic care)
Ryrie *et al.* (1998)[110]	Two acute wards, England	• Eight nurses	• Survey used to collect data on nursing activity every 15 minutes over a 7-day period	• Structured observations • Reactive patient contact • Planned patient contact • Non-patient contact
Whittington and McLaughlin (2000)[117]	Three acute admission wards, Northern Ireland	• 20 psychiatric nurses	• Observational study in which nurses were observed by 2 researchers who recorded activities over a 620 minute period	• Non-patient contact (reports, administration, contact with other staff, unavailable to patients, clinical room duties, time with student nurse, doctor's rounds, other duties) • Patient contact (clinical duties, giving medication, physical care, supervising meals, group therapy, individual therapy, social conversation, interactive close observation, non-interactive close observation, ECT, escort duties)
Higgins *et al.* (1999)[13]	11 psychiatric sites, England	• 118 nurses and ward managers • 54 inpatients	• Data collected from staff and patients using interviews/questionnaires on patient activities and nursing care • Observations of all ward staff and five patients were recorded every 15 minutes through all shifts	• Nursing activities (direct, indirect, associated, personal) • Patient activities (direct, treatment, others, self, sleeping)
Sandford *et al.* (1990)[118]	Four inpatient wards in a hospital, Australia	• Nursing staff members	• Study 1: the number of staff on one ward was varied to see if staff-to-patient interaction rates were related to the staff-to-patient ratio. Observations of staff–patient and staff–staff contacts took place over 20 days, at 5-minute intervals during a 5-hour period • Study 2: natural variations on three wards were recorded and observations of contacts took place as in study 1	• Interaction with other staff or patients (individual verbal, individual nonverbal, group verbal, group nonverbal, physical contact, physical restraint) • Non-interactive behaviour (solitary task oriented)

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 8 Methods of reviewed time and motion studies (WP2) (*continued*)

Author and year	Setting	Participants	Methods	Activities reported
Radcliffe and Smith (2007)[119]	16 adult acute locality wards, England	• Inpatients	• 3- to 10-minute observations by researchers were made per day over 5 days for each ward	• Patients socially engaged (interaction with staff, interaction with others, negative interaction, receiving care, interaction in bedroom, organised activity) • Patients socially disengaged (standing/walking, alone in bedroom, self-care, sitting)
Poole *et al.* (1980)[12]	One ward in a psychiatric hospital and 1 in a general hospital, Australia	• 10 patients on each ward	• Researchers observed how inpatients spent time during a 30-minute period	• Solitary behaviour • Interactive behaviour (with staff, with patients, with others, mixed)
Sanson-Fisher *et al.* (1978)[120]	Inpatient psychiatric unit in a general hospital, Australia	• Professional staff, nursing staff, OTs, and patients	• Patients and staff were observed over two periods separated by 6 weeks. Three researchers observed activity during two randomly selected times within each hour	• Solitary behaviour (individual task behaviour, egocentric behaviour, TV watching) • Interactive behaviour (individual interaction, group task formal, group task informal, group interaction)
Fairbanks *et al.* (1977)[121]	Four psychiatric inpatient wards, USA	• Patients and staff	• Observations took place over an 8-week period. Seven researchers recorded activities on each ward for six sessions	• Behaviour categories (social, non-involvement, movement, entertainment, task oriented, self-maintenance, deviant)
Cormack (1976)[122]	10 psychiatric inpatient wards in four hospitals, Scotland	• 14 charge nurses	• Nurses were observed for 12 hours in 3-hour blocks of time. Activity lasting more than 15 seconds was described and its time was recorded	• Clinical • Administrative • Domestic • Teaching • Free time • Miscellaneous
Altschul (1972)[123]	Four inpatient wards, Scotland	• 113 patients and 40 nurses	• Nurses and patients were observed for 3 to 4 weeks on each ward for about 4 hours per day. Some observations took place during night shifts	• Interaction time • Non-interaction time
Willer *et al.* (1974)[124]	Three wards in one hospital, Canada	• Patients and staff	• Each ward observed for 1 week every hour for 10 to 15 minutes between 8 a.m. and 9 a.m.	• Social activities • Isolated passive behaviour • Mixed activities • Isolated active behaviour
Lepola and Vanhanen (1997)[125]	One hospital inpatient ward, Finland	• Seven patients	• Six patients observed for an 8-hour shift and one patient for 13 hours	• Being alone with no goal-oriented activity • Being alone while participating in daily routines • Acting together on patient's request • Acting together on nurse's initiative or in daily ward routines

ECT, electroconvulsive therapy.

Four of the studies were performed in England, three in Australia, two in Scotland and one each in Finland, Canada, the USA and Northern Ireland. Before the findings from the studies are described, we will summarise the methods used and make some inference about the quality of the studies.

Studies used data collected through interviews, questionnaires, audits of attendance registers, and observations of staff and patients on wards at varying time intervals, usually recorded with an observational instrument and coded according to specific categories.

Interviews or questionnaires

Using this methodology, staff members are surveyed or interviewed about the time spent at work. In the study by Ryrie *et al.*,[110] staff members recorded their activity every 15 minutes over a 7-day period according to a set of pre-designated categories. The authors suggest that the 7-day period may not have been sufficiently representative and that some changes to staffing levels took place after the study. They also felt that 15-minute intervals may have been too long to detect changes in activity. In another study,[116] staff members were interviewed on a daily basis and reported all activities, but not the time spent, undertaken in the past hour as well as an estimate of the number of minutes spent in direct patient contact. This study covered more time periods than Ryrie *et al.*,[110] and so may have picked up more fluctuation in activity levels. The authors considered the number of nurses surveyed ($n = 40$) to be a limitation but this is in fact a relatively large sample compared with other studies (Ryrie *et al.*[110] used a sample of $n = 8$). Neither of these studies included night shifts. The method of self-report of time spent has the advantage that it requires relatively small amounts of researcher time to complete. However, a disadvantage of this method is that, given the hectic pace on the inpatient ward, the staff members may have little time to complete the surveys or interviews and so may under or over-report certain activities. There is also the danger of response bias, in which the staff members may self-report more patient contact than actually occurs.

Audits of attendance registers

Audits of attendance registers have been used to assess patient participation in group activities. This can be used as a verification tool for patient self-report of activities or in conjunction with an observational design as was done by Radcliffe and Smith.[119] This method is helpful in measuring group activity without the time and effort required for observational studies and to double-check patients' reports, which may be inaccurate owing to confusion caused by symptoms or medication. The disadvantage is that it can be difficult to verify group attendance as some patients have difficulty concentrating and may drift in and out of group activities or attend for, say, 10 minutes of an hour-long activity. Radcliffe and Smith[119] also point out that it is difficult to establish the accuracy with which activity leaders record attendance.

Observations of patients and staff

This has been the most commonly used method in studies reviewed. Researchers observe patients, staff or both for specified periods of time and record the activities they take part in, often using an observational instrument in which activities are designated into categories and coded according to category.

Altschul[123] conducted one of the seminal observational studies of nursing activity. She observed 40 nurses in four wards of a Scottish hospital during 4-hour periods over 3- to 4-week time periods. Interactions lasting more than or equal to 3 minutes were included and timed and the nature of the interaction recorded. Another Scottish study (supervised by Altschul) observed 14 charge nurses for 12 hours each.[122] Activities were recorded and categorised. Although these were early studies, they were comparable in observation period and number of nurses included to later work.

A Canadian study observed patients and staff in three wards over a 1-week period.[124] Each ward was observed 60 times. It is likely that observations did reflect ward activity accurately during that period but, as with other studies, it is unclear whether 1 week is sufficient to draw conclusions. In an early study from the USA, patient and staff activity on four inpatient wards was observed.[121] Six 150-minute sessions of observation were conducted for each ward and one person was observed during that time. Inter-rater

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

reliability was tested and found to be reasonable, although variability was quite high for observation of social activities.

In Australia, Sanson-Fisher *et al.*[120] made observations of patients and staff on a single ward. Two 6-day periods of observation were used and during this time observers moved through the ward twice every hour and recorded activities taking place. Although limited to one ward, it was an advantage that two periods (separated by 6 weeks) were used and also that a reliability test was conducted for 10% of observations. The percentage agreement between observers was between 83% and 100%, depending on the activity type. This study was followed by a comparison of patient activity on two different wards by the same authors.[12] Ten patients were selected from each ward and observed six times for 5 minutes. This does represent a relatively short period of total observation compared with other studies. In another Australian study, Sandford *et al.*[118] observed nursing activity when nursing numbers on wards differed. Initially observations were carried out on one ward over a 20-day period. Nurse numbers varied by day with either five, six or seven nurses being on duty. The impact of this on interactions was measured. Subsequently, nurse activity was observed in three wards over 12 weeks, allowing for natural fluctuations in staff numbers. In each case, nurses were observed every 5 minutes for 5 hours.

An in-depth analysis of the activities of seven inpatients in Finland was conducted by Lepola and Vanhanen[125] as part of a more qualitative study. The patients were deliberately selected and the observation period was relatively short and as such the findings would not be generalisable. Elsewhere, Higgins *et al.*[13] observed all nurses and five patients in each of 11 sites. Activities were recorded every 15 minutes during all shifts over a 1-week period. As with other studies, this period of time might be too short to ensure representativeness. Whittington and McLaughlin[117] observed 20 nurses one at a time for one shift. This led the authors to call into question the representativeness of the data. Another limitation noted by the authors was that a convenience sample was used. However, this does appear to be the norm in the activity studies reviewed and may not be as serious a problem as the short observation period. Indeed, a convenience sample may be the most practical (and least disruptive) way of conducting such studies.

The aforementioned study by Radcliffe and Smith[119] observed activity in 16 wards and six hospitals. This was an advantage in terms of representing a wide spread of wards, but the observation period was relatively short (1 week) and only three 10-minute observations per day were made.

The observational method is popular and does appears to be the most accurate way of measuring activity as activities are recorded as they occur, so are not subject to recall bias of retrospective reports or response bias. However, the disadvantage of this method is that it is subject to the observer effect, in that staff members would know they were being watched and their actions recorded, so they may alter their usual schedules of patient care. It is especially common for observational studies to measure activity according to specific categories. Observation does, though, entail a relatively high level of input on the part of researchers.

Some studies that have assessed the use of nursing time have found reasonably similar amounts spent with patients: 48%,[116] 43%[117] and 50%.[110] Other studies have though demonstrated that higher grade nurses spend less of their time directly with patients: < 30%,[13] 27%[122] and 29%.[116] (The former figure is for G-grade nurses. The authors reported that F-grade nurses spent around 45% of their time with patients.[13] Conversely they spend disproportionately more time on administrative tasks.) In one study it has also been shown that time spent with patients has fallen over time for all grades of staff.[13] Elsewhere, Sandford *et al.*[118] reported that only 15–18% of staff time was spent with patients and 31–34% with other staff members (with the remainder being spent alone), while Sanson-Fisher *et al.*[120] found that 24% of staff time was spent interacting with patients.

Although up to a half of staff time might be spent directly with patients, relatively little of this time appears to be spent on providing specific therapeutic interventions: 4%,[116] 17%[117] and 13–20%.[110]

In her early work, Altschul[123] found that nurse–patient interaction accounted for only 1% of observed time on wards and only 30% of patients spent more than 1% in interactions with nurses. Willer *et al.*[124] later reported more positive results with patients spending only 39% of their time in 'isolated passive behaviour'.

Another observational study[121] found that inpatients in four US psychiatric wards spent approximately 50% of their time alone and only about one-third of their time engaged in social activity. Time spent alone by patients was positively related to deviant behaviours and inversely related to social activity. Staff members spent more time than inpatients on social activity (average of 46% across the four wards) and spent significantly more time with other staff members than with patients, for example they were observed near patients less than 20% of the time but spent about twice as much time near other staff members. Before this, Willer *et al.*[124] had also demonstrated that staff spent twice as much time with other staff compared with time spent with patients. Interestingly, Sandford *et al.*[118] found that increasing staff numbers does not necessarily help as staff in their study then spent more time with other staff.

High rates of inpatient solitary behaviour were also reported in two Australian studies about 30 years ago. Sanson-Fisher *et al.*[120] observed that patients spent 49% of their time alone. In a subsequent study by the authors,[12] it was found that most of the patients' time was spent in solitary behaviour in both the mental health hospital ward (83%) and the general hospital psychiatric ward (62%), and little time was allocated to staff–patient interaction on both wards (9% and 13%, respectively).

Higgins *et al.*[13] found that patients spent only 4% of their time with staff, while they spent 28% of their time talking with other patients, watching TV or doing nothing, compared with 17% of their time in therapy. Slightly more patient time (15%) appeared to be spent in organised activities in the small study by Lepola and Vanhanen[125] but this was broadly defined as 'activities related to the patient being together with other patients and the nurses'.

Radcliffe and Smith,[119] in their observational study of 16 UK inpatient wards, observed that patients spent the majority (84%) of their time socially disengaged, that is not interacting with staff members or other patients. Only 4% of patients' time was spent in organised activities such as talking groups or art therapy.

Discussion

This review identified studies that have examined how staff and patients on acute inpatient psychiatric wards spend their time and how much time is spent in staff–patient interaction. Studies published over a 35-year period from seven countries revealed quite consistent and clear findings: (1) on average, around 50% of staff time is spent in contact with patients, (2) the more senior the staff member, the smaller amount of time they spend with patients, (3) the amount of time spent delivering 'therapy' is probably in the region of 4–20% and (4) relatively little patient time is spent in contact with staff and much is spent in isolation. There is also evidence to suggest that staff time with patients is reducing over time[13] and that increasing staff numbers may not result in more time spent with patients.[118] The finding that staff spend relatively little time in direct contact with patients and that patients spend little time with staff do not go hand-in-hand as this depends on the staff-to-patient ratio.

The review also examined the methods used in the studies, most of which observed staff and/or patients over a period of time. A number of these studies were relatively small, including only limited numbers of staff. However, others were more representative in that they covered activity across different wards and sometimes across different hospitals. The majority of studies observed activity for only a limited period of time (sometimes just 1 week) and, as such, may not reflect typical activity patterns, but this is less of a concern in the multiward/site studies. This was recognised as a problem and it is perhaps inevitable given the limited research time that they would have been faced with. Most studies also focused only on daytime shifts. A further recognised problem with the studies is that observers might influence the very activity that they are measuring.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Other methods to measure activity are self-report and collection of data from audits/administrative systems. Self-report by nurses may result in recall bias and an incentive to 'inflate' time spent in activities considered more 'worthwhile'. Self-report and observational methods have been compared (on non-psychiatric wards) and differences found.[126] Nurses reported significantly more time spent on patient care (40% compared with 33% observed) and ward-related activities (7% compared with 3% observed). However, although significant, these differences are not particularly large. The authors also report that observational methods were more acceptable to staff than self-report, which perhaps reflects the busy workload on inpatient wards.

Limitations

There are a number of limitations to the review. First, the number of studies ($n = 13$) identified was relatively low. This is surprising given the broad search strategy and the long inclusion period. This suggests that there is little interest in how staff and patients spend their time on psychiatric inpatient wards, or that research resources have not been made available to address this or that studies have not been published in academic journals. Second, they related to a long period of time and specific settings; as such, it is challenging to draw firm messages from the findings. Third, we restricted ourselves to papers published since 1970 in English. Inpatient services have, of course, existed for a far longer period and it is likely that important work will have been published prior to 1970 and in non-English language journals. Unfortunately, research time did not allow for a search of these. Fourth, 'grey literature' (e.g. unpublished reports) was not reviewed. Finally, only two papers were initially obtained through the electronic search. This was surprising, but searches of the reference lists did reveal substantially more literature.

Implications

It has been known for many years that what takes place on wards is important for patient outcomes. A landmark study by Wing and Brown[11] compared three hospitals in the UK from 1960 to 1968 to determine if the ward atmosphere can affect patients' clinical outcomes. The findings suggest that poor social environments in wards, in which patients spend much of their time not engaged in any activity, exacerbate mental illness. Although patients were not considered to differ in illness severity between the wards, they exhibited fewer negative symptoms in the ward with the best social environment. There were also changes exhibited over the years in all of the wards, for as the social environments improved so did the patients' conditions. Conversely, when two of the wards' social environments deteriorated over time after initial improvements, the patients' clinical conditions worsened and they became more socially withdrawn.

In a more recent study,[127] researchers changed the structure of an inner-city acute care ward by introducing treatment planning meetings, group therapy and community meetings. Two months after implementation, patients on the programme ward reported significantly higher scores of involvement (energy in treatment and patient activity), support between staff and patients, and practical orientation (providing patients with practical skills in preparation for discharge).

The lack of structured activities and patient contact indicated in this literature review is thought to be attributable to a variety of issues. These issues include a more severe patient mix so that staff members spend more time 'containing' than offering activities to patients, staff shortages of psychologists and OTs, and a trend over time of nurses spending less of their time interacting with patients and more of their time on non-patient activities such as administrative duties.

Some options to improve social engagement and participation in group activities are that (1) activities could be offered that are able to be led by unqualified staff or unqualified staff could be trained to lead some groups, (2) efforts could be strengthened to recruit volunteers onto the wards to lead activities and (3) the introduction of information technology onto the wards might be used to streamline administrative duties and paperwork so that nurses are freed up to devote more time to caring for patients on a one-to-one basis.

Systematic review of economic evaluations of therapeutic activities on psychiatric inpatient wards

Chapter 4 will report on the evaluation of an intervention to increase access to therapeutic activities on psychiatric inpatient wards. Such interventions require resources that could be used to provide other forms of care to people with mental health problems (or indeed patients in other clinical areas) and, as such, these interventions need to be evaluated in terms of cost-effectiveness, and *Chapter 4* does include such an economic evaluation. As a precursor, we conducted a systematic review of economic evaluations of therapeutic activities on wards.

Methods

Initially, the electronic databases EMBASE (1947–2014 week 4), MEDLINE (1946–2014 week 4), and PsycINFO (1806–2014 week 3) were searched with combinations of 'inpatient' or 'admission' or 'admitted' or 'hospital$' in the title plus 'healthcare util$' or 'health care util$' or 'healthcare cost$' or 'health care cost$' or 'cost$' or 'econ$' or 'financ$' in the title plus 'mental' or 'psychiatric' in the title. Abstracts were reviewed by PM to assess eligibility for inclusion and checked by LK. The criteria for inclusion were that (1) the study was based in an inpatient setting and (2) that it reported an economic evaluation of a specific intervention rather than a model of care. As in the previous review, we did not attempt to combine the results in a formal way.

Results

The search strategy identified 236 papers. None of these met the search criteria.

Discussion

Inpatient care constitutes a substantial proportion of mental health-care expenditure and yet there appear to have been no economic evaluations of interventions that take place on inpatient wards. Although this is worrying, it was not an unexpected finding given that there is little clinical evidence regarding inpatient interventions and it is rare for economic findings to be produced in the absence of clinical results.

Modelling the cost-effectiveness of inpatient interventions: social cognition training case study

Decision modelling is a way of synthesising both economic and clinical evidence from a variety of sources in order to assist decision-makers in evaluating health technology. Thus, it can aid in understanding the trade-offs between costs and benefits of an intervention by explicitly and systematically laying out information about the health-care processes. Decision models can, for instance, be helpful if clinical trials are unable to capture the scope of an intervention (e.g. because its time horizon is beyond the typical length of a randomised trial) or if economic evaluation alongside a clinical trial is not feasible (e.g. because the intervention is already in place, other ethical issues or prohibitively high costs).

The literature review presented earlier revealed that there have been very few economic evaluations of inpatient-based interventions for people with mental health problems. The aim of this section is to demonstrate how modelling can be used to generate such information using a case study. It provides preliminary evidence regarding the cost-effectiveness of a psychological intervention for schizophrenia that is not routinely provided. Owing to the limitations of medication treatment, psychological interventions over the past few decades have gained increasing recognition in the treatment of schizophrenia.[128] However, many of these have not been empirically validated and their cost-effectiveness has not been established.[128] This case study uses social cognition training (SCT), which was provided as part of the DOORWAYS intervention in form of social cognition and interaction training (SCIT)[129] to illustrate the considerations relevant in constructing a decision model for health economic evaluations.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Methods

The target population of the decision model were inpatients with schizophrenia. SCT was compared with treatment as usual (TAU). The economic value of SCT was assessed from a health-care perspective and a time horizon of eight months was assumed so that no discounting of costs or benefits was required. There were only limited data on care pathways for this intervention and the outcomes experienced by patients. Therefore, a simple decision tree structure was chosen (*Figure 1*).

Evidence was collected on the risk of relapse for each trial arm over the time horizon of the model.

For the purposes of this exploratory model, the general psychopathology subscale of the Positive And Negative Symptom Scale (PANSS) was chosen as an outcome measure. PANSS is a commonly used, disorder-specific, non-preference-based instrument. It measures symptoms in patients with psychosis rather than quality of life. This means that cost-effectiveness estimates based on this measure are not easily comparable to alternative interventions in other disease areas. However, our model uses an equation estimated by Siani *et al.*,[130] which predicts EuroQol-5 Dimensions (EQ-5D) utility values based on PANSS scores. The EQ-5D utility values were then used to calculate quality-adjusted life-years (QALYs). QALYs combine the health-related quality-of-life values and the length of time that this quality is enjoyed into a single summary metric.

Relevant clinical trial evidence was identified based on two recent systematic reviews of the SCIT literature and studies citing these.[131,132] The evidence on the reduction in the PANSS subscale was synthesised using a fixed-effects model for continuous measures assuming that there was no correlation between baseline and follow-up values within each trial arm. The meta-analysis was conducted in Stata® (version 11; StataCorp LP, College Station, TX, USA).[133] To date, there are only a few studies investigating the durability of treatment effects in SCT.[131] For this reason, the treatment effect was assumed to decrease in a linear fashion up to the end of the time horizon of the model.

We identified only one small observational study by Hodel *et al.*[134] ($n = 19$) investigating the effect of SCT on relapse rates. It suggested that this intervention may reduce rehospitalisation at 8-month follow-up. Conservatively, the reduction in relapse rate was assumed to lead only to cost saving but not improvements in quality of life. Based on results from elsewhere in the programme, the average length of admission was assumed to be approximately 45 days. To determine the cost of inpatient stay, data were derived from a nationally recognised source.[134] The length of SCT in the literature was variable, ranging from 6 to 24 weeks.[135] Based on the studies used to inform the model, treatment duration of 8 weeks with two 45-minute sessions per week led by a psychotherapist was assumed. The cost of a clinical psychologist was used as a proxy for the cost of a psychotherapist.[136]

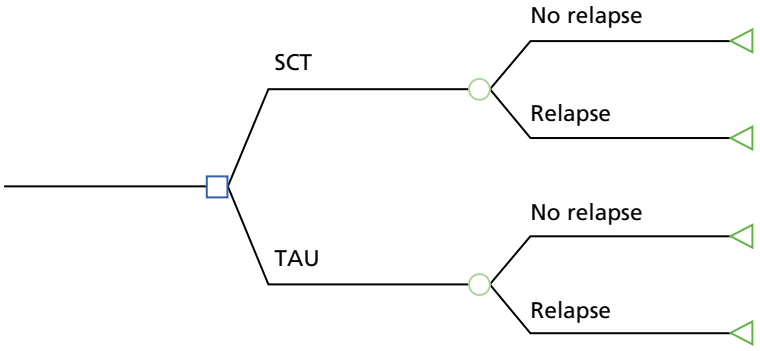


FIGURE 1 Decision tree structure.

In addition to calculating the differences in costs and benefits of the intervention, the cost-effectiveness of the interventions was compared using a net monetary benefit approach.[137] If λ is defined as the value that the decision-makers assign to increasing QALYs by one unit and E and C are the incremental costs and differences between the two treatments, the net benefit of the intervention can be defined as $NB = \lambda E - C$. Based on this framework, a probabilistic sensitivity analysis was conducted for a range of λ values from £0 to £40,000 per QALY (i.e. the probability that SCT was cost-effective was calculated based on random draws from the distribution of the input parameters). The results of this analysis were displayed on a cost-effectiveness plane as well as a cost-effectiveness acceptability curve. In addition, a threshold analysis was conducted by examining at which point the net benefit of SCT would be zero assuming that the reduction in relapse rates in the observation study was biased upwards.

Results

Three relevant studies were identified in the literature review.[138–140] The weighted mean reduction in PANSS score attributable to SCT compared with TAU was estimated to be –3.49 (95% CI –8.01 to 1.04) (*Figure 2*). The study by Hodel *et al.*[134] suggested that SCT reduced the risk of relapse by 26% (95% CI –70% to 17%) percentage points. The cost of the intervention was calculated to be £1632 and the savings for an averted hospitalisation were £15,615. In the base-case scenario, SCT was estimated to both increase QALYs by 0.009 (95% CI –0.007 to 0.018) and decrease the costs by £2532 (95% CI –£3282 to £8110); *Figures 3* and *4* show the joint distribution of costs and benefits. In other words, the model indicated that it was the dominant treatment option. In addition, the probability that SCT was the most cost-effective was found to be high (> 75%) at all levels of willingness to pay for a QALY improvement. The threshold analysis suggested that in order to be at least as equally cost-effective as TAU, SCT needed to reduce the risk of hospitalisation, on average, by 8.5% to 10.5% depending on the willingness to pay for a QALY improvement.

Discussion

One should be cautious not to overstate the results of this exploratory model because some strong assumptions were used to be able to construct a coherent framework based on relatively sparse data. For example, the mapping or 'cross-walk' approach employed in this study has been described as a 'second best' solution and introduces additional uncertainty and errors into estimates[141] Although we showed that the reduction in risk of hospitalisation of less than half of the base-case scenario may be sufficient for SCT to be cost-effective, it is highly uncertain to what extent the findings by Hodel *et al.*[134] are generalisable. As noted by Henderson,[135] research on SCT 'remains very much in the developmental stage'. Based on current evidence, there is little robust evidence regarding its effectiveness but the results of this model suggest that it might have the potential to be a cost-effective intervention. However, before making the transfer from a developmental stage to routine clinical practice, further knowledge regarding the effects of SCT on quality of life and in the long term is required.

The objective of this section has been to demonstrate the feasibility of using a modelling approach to assess the cost-effectiveness of inpatient-based therapeutic interventions. As will have been apparent, this approach is dependent on the quality of the data available. Although it can circumvent the need to collect primary data from patients, it still requires robust estimates to be made of model parameters and to explore any uncertainty around these estimates in an appropriate manner. Having said that, the modelling approach is generally feasible and, as long as it is conducted with transparency, it represents an accepted way of generating information on cost-effectiveness.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

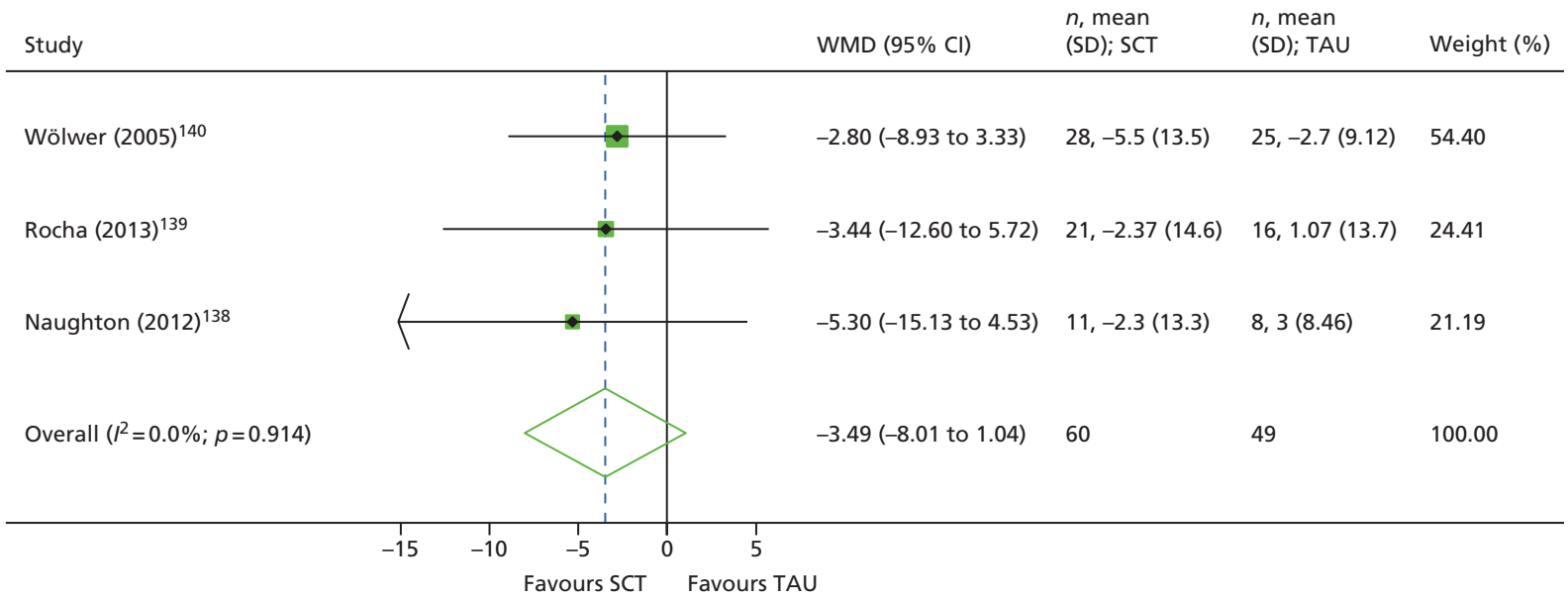


FIGURE 2 Forest plot of meta-analysis: reduction in general symptoms PANSS.

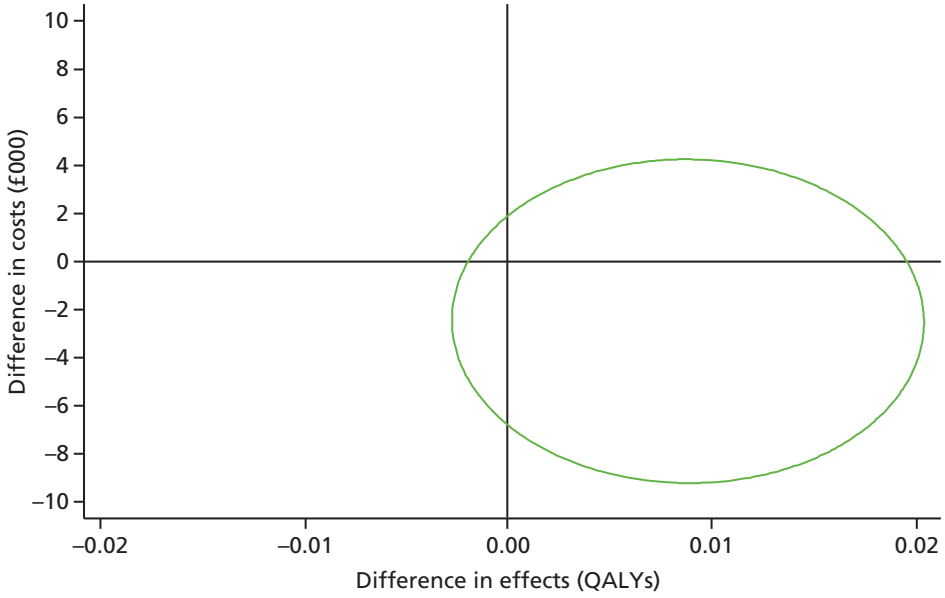


FIGURE 3 Cost-effectiveness plane: mean costs and effects of social cognition relative to TAU with 95% confidence region.

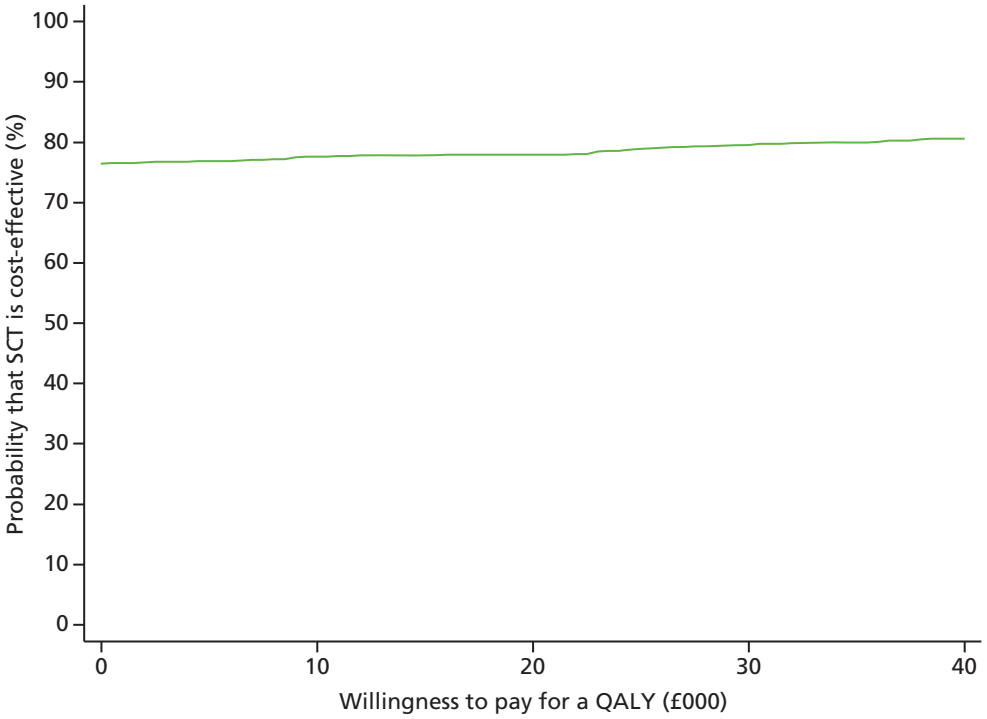


FIGURE 4 Cost-effectiveness acceptability curve.

Development of a measure of service use

The aim of this section is to describe the development and assess the reliability of a tool (the CITRINE) to achieve the objectives of collecting data on the care contacts and therapeutic activities of patients while on a psychiatric ward. This would potentially enable a more precise costing of inpatient psychiatric services.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Methods

Reproduced from Sabes-Figuera R, McCrone P, Sharac J, Csipke E, Craig TK, Rose DS, *et al.* Developing a tool for collecting and costing activity data on psychiatric inpatient wards. *Epidemiol Psychiatr Sci* 2012;**21**:393–9.[142] © Cambridge University Press 2012.

In developing the tool we conducted interviews with inpatient staff to discuss its content and structure. The professionals interviewed included two ward managers, three OTs, one nurse and one therapy and benefits co-ordinator. All were working at psychiatric inpatient units in the South London and Maudsley NHS Foundation Trust. The interviews lasted for approximately 30 minutes and the main objective was to identify the group activities that take place on the wards for which professionals were most likely to have contacts with inpatients. Staff sometimes provided detailed schedules of patient activity on their wards. The first draft of the tool was then circulated for comments and suggestions to a group of experts including psychiatrists, psychologists and nurses.

Twenty-five service users from inpatient psychiatric wards were then interviewed using the questionnaire. The objective was to determine the acceptability of the tool, service users' opinion about difficulties in answering the questions and to evaluate how practical it was to collect data in this way. Two focus groups were conducted in parallel with this feasibility component. The first included eight OTs, with seven nurses from an inpatient ward comprising the second. Participants were asked for their opinions about the tool and for any suggestions for improvements.

The process of consultation with staff resulted in the development of a tool that allows activities specific to an inpatient ward to be measured. The list of group activities was amended to include only the activities that had taken place in the previous week. Interviewers list the availability of activities (based on the ward calendar for the relevant week) and, therefore, each version of the tool is unique to that particular ward and time period. Initial questions on contacts with members of staff were shortened. The staff contacts section was also revised so that service users could indicate the name of professionals they have met even if they did not know or remember their job title. The tool is shown in *Appendix 9*.

This next study assessed the level of congruence between the information provided by service users using the final version of the tool and information that could be obtained from other sources. It took place in two stages. Initially, information on the activities attended by 41 service users over the preceding 7 days was obtained from the OT of each ward and the number of contacts with nursing staff members was collected from the patient's records for the same period. This information was also collected from the same patients and period using the tool.

Information on the number and duration of staff contacts and activities attended by 22 additional service users was collected over a 14-hour direct observation period (08.00–22.00). An adapted 1-day version of the CITRINE questionnaire was used to collect the relevant information from the same service users over this period. (This stage was added later and so the patients are different from the 41 previously included.)

Contacts with staff members were combined with unit costs measured in UK pounds for the financial year 2007/8 from an established source.[109] In the first phase ($n = 41$) we assumed an average nurse contact of 15 minutes. Activity costs were calculated specifically for the study based on session duration, preparation required, staff involved and materials for each activity provided on each ward. These data were combined with unit costs of staff time[109] and with information on the average number of service users attending each activity to obtain an estimated individual cost for each group activity. Activities were subsequently classified in categories according to their type and cost. (A list of all unit costs is available from the authors.) An average of these activity costs was used in the analyses presented here. The significance of differences in reported and observed/recorded mean resource use was assessed using a paired *t*-test and the agreement was further assessed using the concordance correlation coefficient.

Results

The CITRINE tool typically took between 5 and 10 minutes to complete and most of the service users who participated in the feasibility study found it easy or very easy to answer the questions. *Table 9* reports 7-day data from the CITRINE and from information covering the same period from OTs and case notes for 41 service users. The mean age was 37 years and 19 of the participants were women. Most participants (68%) had a primary diagnostic of psychosis or bipolar disorder and the average length of stay at assessment was 43 days. However, there was great variation in this represented by a SD of 66 days.

Service users report attending more activities than is contained in case notes (see *Table 9*). Attaching an average unit cost of activities suggests a cost difference of £10 per person. Case notes also report more one-to-one nursing contacts, resulting in a cost difference of £4 per person. Although neither difference is large, the discrepancy in the number of activities reported is statistically significant.

The comparison of data obtained from the 1-day observational study and the 1-day version of the tool shows good congruence in terms of activities attended and psychiatrist contacts. However, the congruence in terms of contact with nurses and other staff is less good. The differences in the 1-day version of the tool were significant or of borderline significance. The concordance correlation coefficient for total costs from this part of the reliability study was 0.79. The cost implications for the difference in nursing contact are important given that the observational data suggest that this accounts for one-third of the total cost.

Discussion

The process of developing the CITRINE tool has involved input from a wide range of staff involved in providing care to those receiving psychiatric inpatient services (nursing staff, OTs, ward managers and

TABLE 9 Mean number and cost (£) of activities and nursing contacts reported by service users and from OTs/case notes (WP2)

Type of activity	Mean (SD) number of contacts/duration		Mean (SD) cost (2007/8 £)		Statistics	
1-week stage (n = 41)	***Data provided by service user***	***OT/case note data***	***Data provided by service user***	***OT/case note data***	**t-value[a]**	***Significance[a]***
Activities (number)	5.2 (5.6)	3.8 (3.0)	39.5 (43.2)	29.2 (23.2)	2.159	0.037
One-to-one contacts with nursing staff (number)	2.8 (2.7)	3.2 (3.9)	30.4 (29.4)	34.3 (47.3)	–0.501	0.619
1-day stage (n = 22)	***Data provided by service user***	***OT/case note data***	***Data provided by service user***	***OT/case note data***	**t-value[a]**	***Significance[a]***
Activities (number)	1.0 (0.8)	1.1 (1.1)	7.7 (6.3)	8.4 (8.2)	–0.624	0.539
One-to-one contacts with nursing staff (duration in minutes)	7.1 (13.8)	29.8 (23.0)	5.1 (9.9)	21.3 (16.5)	–5.394	< 0.001
One-to-one contacts with psychiatrists (duration in minutes)	5.7 (16.3)	5.7 (11.1)	30.2 (86.0)	29.9 (58.2)	0.027	0.979
One-to-one contacts with other staff (duration in minutes)	3.2 (12.9)	9.1 (13.6)	2.7 (10.7)	7.6 (11.3)	–1.982	0.061
Total cost			45.6 (92.9)	67.2 (72.7)	–1.990	0.060

a Tests of significance apply to both contacts/duration and cost.

Reproduced from Sabes-Figuera R, McCrone P, Sharac J, Csipke E, Craig TK, Rose DS, *et al.* Developing a tool for collecting and costing activity data on psychiatric inpatient wards. *Epidemiol Psychiatr Sci* 2012;**21**:393–9.[142] © Cambridge University Press 2012.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

psychologists). Furthermore, key input to the process has been provided by the recipients of this care. The tool collects data directly from service users. From a pragmatic perspective, this is appropriate as it is the most practical way of obtaining the breadth of activity data that we require for economic studies. In addition, it emphasises the service user's perspective in reporting activity. Of course this could mean that activities that are entirely unmemorable will not be recalled or costed (e.g. conversations with nurses that could be very similar every day such as 'you have a telephone call', 'do you want to go for a walk?'). However, this type of very brief contact that is not remembered by service users is not the focus of the tool.

The result is a tool that has the objective of collecting information on the therapeutically relevant activities and staff contacts that take place on psychiatric wards and at the same time is acceptable to service users and relies on them as the main source of the relevant information. There have been other attempts to describe interventions and services provided in inpatient mental health settings, such as the International Classification of Mental Health Care.[143] However, this instrument is a tool to classify services according to type of care, rather than one that provides information on the specific services that patients receive.

The main function of the tool will be its use in economic studies, combined with information on relevant unit cost of the services and staff time. Therefore, its level of accuracy in reflecting what is actually taking place within psychiatric wards is paramount. The reliability study carried out offers some reflections on this issue. In particular, the questionnaire (1) is a good source of information on the number of activities attended by service users and on contacts with psychiatrists, (2) may provide acceptable information on interactions between service users and non-nursing care staff (OTs, psychologists, social workers, etc.) and (3) may be more limited in recording time spent in service user–nursing staff interaction.

In relation to nursing staff contacts, there are some aspects that should be considered. First, this type of contact is common and, consequently, creates difficulties for accurate recall. Second, service users might report only contacts that they think are significant or meaningful for them. This classification is subjective and can result in some contacts being labelled as 'non-contacts' by service users. For example, a member of staff may have spent some time asking how a patient was and would regard this as a contact, but it may not have been recognised as such by a patient if it was very brief or unwanted. This may have been the key reason for the discrepancy between the observed nurse–patient contacts and the patient-reported contacts. The mental health status of some of the service users may affect their ability to provide accurate information, although data on this are lacking. Although time spent with nursing staff needs to be measured correctly, the difficulty seems to apply only to one-to-one contacts. Contacts as part of organised activities are more readily measured and, therefore, the disagreement over total cost is limited.

The alternative sources of information on inpatient psychiatric activities and staff contacts are not free of problems. Registers and electronic databases are designed to support clinical care and not to record activities within wards. Furthermore, there might be intra- and interward variability on the level of completion of these and on the accuracy of the information recorded. Observational data are an alternative but require too many resources to be a realistic when studies involve large numbers of individuals/wards with data collected at several time points.

There are limitations with this work. First, we did not conduct the full range of psychometric tests, although we did assess validity and reliability. It should be stressed that we do not regard this as a clinical instrument but rather as a simple recording schedule to ascertain more accurately the use of care on inpatient wards. We anticipate that the tool being adapted for different circumstances and extensive tests on one particular version would have limited usefulness. Second, and related, we did not assess inter-rater reliability. The content of the tool, though, is such that ambiguity of responses should be limited and, hence, inter-rater reliability is not as important as with a tool for which interviewer interpretation is required. Third, the activities relate to wards in one particular hospital. It may be that in another setting a more complex array of activities would present more of a challenge. However, the tool does require the listing of specific activities at the start and so this problem is reduced. Fourth, the relatively poor concordance on contacts

with nurses is a cause for some concern. This has been discussed, but the wording of the question may not be clear. Greater definition of what constitutes a contact may be required. Fifth, service users may not always have been clear about the profession of the person they had contact with. However, we have dismissed this because no service users mistook a psychiatrist for nurse and this is likely to be the category of error that would most skew the costs data. Self-report data are associated with such problems. Sixth, to rely on self-report data might be a limitation of the tool although a number of studies have suggested that patient recall of service use is acceptable.[144–146] Asking for the duration of contacts may be difficult but these durations are meant to be approximations.

Analysis of baseline costs

This section has the following objectives: (1) to describe the care contacts, therapeutic activities and related costs for patients on psychiatric wards in a large inner-city hospital and (2) to identify patient predictors of these costs.

Methods

Measures

The measures were (1) sociodemographic data: gender, age, ethnicity, level of education, marital status, and living and employment situation, (2) clinical information: diagnosis, whether or not the current admission to hospital was under the Mental Health Act[72] (i.e. involuntary), at what age the service user had their first contact with mental health services and time since their first psychiatric admission, (3) functioning: Global Assessment of Functioning (GAF)[147] scores were obtained for each service user. In addition, in order to explore the impact of service users' perceptions on service use, the VOICE was included. Data on use of services were collected using the CITRINE.

Service costs

The service use data obtained with the questionnaire were combined with appropriate unit costs for 2008 obtained from national sources for care professionals.[109] Similar data were not available for the unit cost of therapeutic activities and these were therefore calculated specifically for the study. Staff members responsible for organising activities were asked to provide details on the session duration, preparation required, staff involved and materials for each activity provided on each ward. These data were combined with unit costs for these resources and with information on the average number of service users attending each activity to obtain an estimated individual cost for each group activity. Activities were subsequently classified in categories according to their type and cost.

Statistical analysis

Multiple regression models were constructed to identify factors that explained variations in costs. Dependent variables were the cost of one-to-one care contacts and costs of therapeutic activities. A simple regression model for each potential independent variable was run and those variables with statistical significance for any of the two dependent variables were included in the final model. All selected variables were entered in a single block in a fixed-effects regression to take account of data being obtained from service users on 17 different wards. Cost data are often positively skewed and this may result in regression residuals that are similarly skewed, which is a violation of the assumptions underlying the linear regression model. Consequently, non-parametric bootstrapping methods were used.[148] Bootstrapping involves resampling with replacement from the original data a sufficiently large number of times so that the population from which the sample is drawn can be approximated. Here, 5000 samples were automatically generated and bootstrapped 95% CIs were generated.

Results

Of those service users who fulfilled the inclusion criteria, 62% agreed to participate in the study. Service use data were available for 402 service users; however, some values for sociodemographic and clinical variables were missing for some participants or their length of stay was < 7 days. It was decided to include

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

in the analysis only participants with complete data and the final sample size was 334. It was found that those with incomplete data had lower GAF functioning scores (35.6) than those included (39.8) ($p < 0.01$).

Characteristics of the sample are shown in two tables (see *Tables 10* and *11*). Two-thirds had a diagnosis of schizophrenia or bipolar disorder and a similar proportion had been admitted under the Mental Health Act. The average length of stay at the point of assessment was almost 7 weeks, but with a wide range. The GAF scores indicate a high level of disability. As stated above, the VOICE was designed so higher scores reflected lower quality of care. Therefore, those service users with a score of 38 or lower felt that the quality of care received was good. Alternatively, those with scores of 95 and higher, perceived that inpatient care was of low quality. In our sample, 21% had a VOICE score of 38 or lower and 5% had a score of 95 and higher. Therefore, most of the sample felt that the quality of care was neither good nor bad.

Table 12 summarises the data collected with the CITRINE instrument, providing information on the number of therapeutic activities attended and number and type of one-to-one care contacts reported by participants during the period of 7 days. Almost 10% of the sample stated that they had not had any one-to-one contact and more than one-fifth reported that they had not attended any activities. The average number of one-to-one care contacts for the whole sample was less than one per day.

TABLE 10 Sample characteristics: binary variables ($n = 334$) (WP2)

Characteristic	*n* (%)
Female	143 (42.8)
Non-white ethnicity	176 (52.7)
≥ 12 years of education	131 (39.2)
Schizophrenia or bipolar disorder	218 (65.3)
Current admission to hospital under the Mental Health Act	216 (64.7)
Current admission is first admission	68 (20.4)

Reproduced from Sabes-Figuera R, McCrone P, Sharac J, Csipke E, Craig TK, Rose DS, *et al.* Developing a tool for collecting and costing activity data on psychiatric inpatient wards. *Epidemiol Psychiatr Sci* 2012;**21**:393–9.[142] © Cambridge University Press 2012.

TABLE 11 Sample characteristics: continuous variables ($n = 334$) (WP2)

Characteristic	Mean (SD)	Maximum	Minimum
Age (years)	40.0 (12.8)	75	18
VOICE scale score	54.8 (19.5)	114	19
GAF symptoms	42.4 (14.3)	80	1
GAF functioning	39.8 (10.9)	81	11
Length of stay at assessment (days)	48.4 (94.0)	1261	7

TABLE 12 Therapeutic activities (WP2)

Activities	*n* (%)	Mean (SD)	Median	Maximum	Minimum
Therapeutic activities attended	262 (78.4)	3.9 (4.7)	2	25	0
One-to-one contacts with care staff	301 (90.1)	5.6 (5.4)	4	43	0

More than 40% of participants reported no contact with a member of nursing staff in the past week and for those who did have contact, their frequency was slightly higher than one contact every 2 days (*Table 13*). Participants with no nursing contacts were not more likely to have contact with other staff members (and so this does not indicate a problem of distinguishing between care staff). The number of participants who reported contacts with the other types of care staff was low, except for psychiatrists, with whom almost three-quarters of the sample had a weekly average of 1.6 contacts.

The service use data were combined with the unit costs for 2008 to obtain the costs of therapeutic group activities and one-to-one care contacts during the previous week. There was significant variation within the different group activities in the cost per participant attendance. This variation is the result of three factors: first, the different qualifications and number of professionals involved; second, the duration of the activities and the time necessary for their preparation; and, finally, the number of participants who attended each activity. This number varied from 2 to 14.

Table 14 shows the average cost of group activities and one-to-one care contacts over the past week for the 334 participants. Costs were highest for time with a psychiatrist and accounted for more than half of the cost of care contacts. The cost of group activities represented a small proportion of the total cost, as did contacts with nursing staff. If the total figure, cost of group activities and one-to-one care contacts for the 1-week period is translated to a cost per day, the value is slightly more than £30 per day. There is substantial variation between individuals, as reflected by the SDs.

TABLE 13 One-to-one contacts (WP2)

Type of one-to-one contacts	*n* (%)	Mean (SD) contacts for those in receipt	Mean (SD) duration (minutes) for those in receipt
Nursing staff	198 (59.3)	4.1 (3.8)	16.8 (19.1)
Psychiatrist	247 (74.0)	1.6 (1.0)	18.6 (14.5)
Other doctor	84 (25.1)	1.7 (1.3)	16.2 (14.3)
OT	68 (20.4)	2.6 (2.9)	21.8 (24.3)
Care co-ordinator	91 (27.2)	1.6 (1.4)	23.4 (25.3)
Other care staff	110 (32.9)	1.9 (2.2)	36.8 (32.7)

TABLE 14 Cost of services and activities (2008/9 £) (WP2)

Service/activity	Mean	SD	% of total
Nursing staff	29.30	69.20	12.90
Psychiatrist	109.00	131.00	48.10
Other doctor	17.40	45.00	7.70
OT	6.50	21.00	2.90
Care co-ordinator	15.20	48.60	6.70
Other care staff	19.60	52.60	8.70
All one-to-one contacts	197.00	196.30	86.90
Therapeutic activities	29.70	38.80	13.10
Total cost	226.70	206.10	100.00

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The regression analysis of the variation in costs of one-to-one contacts (*Table 15*) showed that older participants had lower costs. A higher score on the VOICE instrument was related to lower cost, indicating that participants with a worse perception of inpatient care also use fewer inpatient resources. For instance, a decrease in the VOICE score of 10 points would imply, assuming other variables do not change, an increase in costs of £11. There were significant cost differences between inpatient wards. Regarding the activities model, the VOICE score was, again, significantly associated with cost (see *Table 15*). Service users with a higher level of education had higher activity costs. (Subsequent analyses revealed this to be caused by more engagement with activities rather than use of higher-cost activities.) The results of the two models also show the positive relationship between the two types of cost.

Discussion

It was striking that 40% did not report contact with nursing staff. The results for therapeutic activities attended do not indicate that the low level of contacts with staff was compensated by more attendances. We have estimated that the cost of care contacts and therapeutic activities is around £30 per day, which is just slightly higher than 10% of the average cost per bed-day of £288 for 2008/9.[149] This figure of £288 includes the costs of other resources not considered in the present study such as cleaning, maintenance, food, laundry, drugs and other treatments, etc. Crucially, the figures here do not include staff time that is not spent in contact with patients. In relation specifically to nursing staff costs, Bowers and Flood[150] analysed expenditure data (i.e. a 'top-down' approach including non-contact time) for 136 wards and they found a cost per bed-day equal to £90, while in the current study we found that the daily cost of nursing staff contacts was £4.20. The difference could be partially caused by an under-reporting by service users of any contacts and, specifically, nursing staff contacts. However, our review of studies that measured nursing and patient activity and interaction on psychiatric inpatient wards found proportions of staff time spent in direct contact with patients (24–48%) and/or providing specific therapeutic interventions (4–20%).

TABLE 15 Regression of one-to-one contact costs and activity costs on demographic and clinical characteristics (clustering for ward) (WP2)

	Model of contacts cost[a]			Model of activities cost[a]		
Variable	B	SE	95% CI	B	SE	95% CI
Age (years)	–1.75	0.66	–3.23 to 0.64	–0.15	0.16	–0.44 to 0.18
Gender (female vs. male)	7.28	30.96	–74.01 to 48.65	9.36	6.32	–0.90 to 23.87
Ethnicity (non-white vs. white)	–42.53	22.9	–84.85 to 4.31	0.84	2.43	–2.98 to 6.50
First admission (yes vs. no)	34.18	28.86	–10.52 to 99.91	1.92	8.03	–15.90 to 15.38
Education (≥ 12 years vs. < 12 years)	46.99	21.1	–0.85 to 81.32	6.89	5.49	0.62 to 21.88
Diagnosis (schizophrenia and bipolar vs. other)	3.24	18.13	–32.17 to 39.01	–2.95	5.99	–19.57 to 2.48
GAF symptoms	–0.55	0.76	–1.96 to 1.04	–0.27	0.18	–0.65 to 0.02
GAF functioning	0.47	1.07	–1.99 to 2.23	0.12	0.24	–0.21 to 0.70
VOICE (prorated if ≥ 16 items answered)	–1.13	0.33	–1.72 to 0.44	–0.21	0.11	–0.47 to 0.05
Length of stay (at assessment)	–0.12	0.14	–0.47 to 0.02	0.03	0.03	0.00 to 0.10
Activities cost	0.51	0.36	0.03 to 1.47	–	–	–
Contacts cost	–	–	–	0.02	0.01	0.00 to 0.04
Ward variable	$F(16 \text{ to } 306) = 2.66$; $p = 0.001$			$F(16 \text{ to } 306) = 5.55$; $p < 0.001$		

SE, standard error.
a Costs in 2008/9 Great British pounds.

High scores on the VOICE questionnaire, representing low satisfaction, were related to low costs of staff contacts and activities; however, the direction of the association is not clear. It is possible that attending fewer activities and having fewer contacts with staff is a result of a worse perception of care received, or low levels of care received may lead to poor perceptions of care. In relation to sociodemographic characteristics of the patients, age was inversely related to staff contact costs, while service users with ≥ 12 years of education had activity costs 23% higher than those with fewer years of education. Nevertheless, it is interesting that McCrone *et al.*[151] found that older people and those with lower levels of education had higher costs of community mental care in a sample of individuals with severe mental illness. The age effect is the opposite of what was found here and may be related to other factors (e.g. physical health in the elderly and/or chronicity relating to poor function). In addition, it should be recognised that these costs may be influenced by the structure of services; staff are present on inpatient wards (i.e. they are available), whereas in the community it may be more complex to arrange a contact.

These findings need to be considered while taking account of the limitations of the study. First, a relevant proportion of inpatients, almost 40%, refused to participate in the study. It was not possible to collect information on demographic or clinical characteristics for this group of patients, therefore preventing a comparison with those who agreed to participate. However, subsequent data collection as part of WP4 indicates that those who consented were representative of people admitted to the wards. There are also some limitations of the CITRINE questionnaire which were discussed earlier.

The low costs found reflect low levels of reported interaction between care professionals and inpatients. This is concerning because it has been established that in inpatient care, outcomes are associated with the level of attention patients received from staff.[152] The relationship found in this study between costs and service users' views of the quality of care, measured by the VOICE instrument, emphasises these previous findings.

Conclusion

This chapter has summarised the existing evidence from studies exploring staff–patient interactions on inpatient wards. The key findings of this review are that despite evidence of the benefits of and official guidance for therapeutic interaction, reports of low activity and social engagement for patients have remained stable for 35 years and that limited nursing time is spent in direct contact with patients. It provokes the need for better exploration of how psychiatric inpatient staff members allocate their time on the ward within the constraints of limited acute care resources and how their time could be spent in more effective ways. This is especially important given that relatively few studies have addressed this topic which is likely to be of key interest to mental health professionals and patients.

Increasing therapeutic activities is a core part of this programme and it is of interest that no existing economic evaluations of such interventions have been identified. Increasing the evidence base is necessary and one approach is to use modelling methods, and we have provided a demonstration of these techniques. The main alternative to modelling is to collect primary data and we have developed a measure of staff contacts and activities to do this. The CITRINE is a tool that, despite some limitations, provides adequate information on the activities that take place within psychiatric wards. Therefore, its use is recommended, alone or in combination with other sources, in economic analyses of inpatient care. This chapter provides analyses of costs data based on this measure. This is the first study, to our knowledge, to present a detailed analysis of the cost of care contacts and activities on adult psychiatric inpatient wards. The results show that service users report low levels of quality contacts with staff members.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 4 Work package 3: DOORWAYS

Introduction

Advances in psychopharmacology, changing attitudes towards mental illness and growing concerns about the quality of life in asylums have led to the gradual transition from the long-term hospitalisation of those with mental health problems to an increase in community treatment. However, hospitalisation for acute episodes is still an indispensable feature of mental health care and consumes the greatest proportion of mental health service costs. Despite this heavy investment, service users and front-line staff continue to complain about the quality of inpatient care.[10] More than 50 years ago, when institutionalism was being re-evaluated, Wing and Brown[11] carried out their well-known 'Three Hospitals' study, which confirmed that the hospitals that provided richer social environments and social opportunities had markedly less disturbance in verbal and social behaviours.[11] Between 1960 and 1964, the social environment improved in all three hospitals, while the clinical condition of patients improved alongside it. Those patients with the least social interaction, the least access to the outside world and with the fewest interesting activities to take part in and who spent the majority of their time doing nothing were the most unwell, but as their environments became more stimulating, they too showed some improvements. These associations between environmental stimulation and acuity remain.

In addition, as bed numbers continue to fall, wards are generally now reserved for the most acutely ill with consequent increases in the proportion of inpatients who are compulsorily detained. Both these factors are likely to increase levels of behavioural disturbance.[7–9] This can lead to a very fraught atmosphere on the wards, with nurses spending the majority of their time dealing with crises as they emerge rather than engaging in therapeutic activities or interactions with service users. The Department of Health[153] has acknowledged that in spite of advocating therapeutic environments, this fire-fighting activity, along with an abundance of administrative work and lack of support and supervision, make this very challenging indeed. Taken together, this often brings about ward environments that appear to be more custodial than therapeutic.

Apart from the stressful ward atmosphere, service users also report that there is often very little to do with their time while they are inpatients, which results in intense boredom.[10,19] Service users recognised that the lack of therapeutic activities (or the cancellation of those on offer) is often a result of staff shortages and/or too many demands being placed on staff. Walsh and Boyle[34] found that service users do, or try to, actively seek out entertainment and company as a means of providing purpose and routine for their days. Nolan *et al.*[154] found that the camaraderie and social support from fellow service users, as well as their ability to provide normalisation and positive examples of recovery, were very highly valued and, therefore, fostering such opportunities is important. Some service users have indeed found the support that they need from other service users who are more available than nurses.[155] These interactions can most easily be encouraged through the provision of organised ward activities.

Echoing the Department of Health findings, nurses report the primary reason for not spending time on therapeutic activities or direct patient contact on the need to focus on resolving crises for a small number of patients, as well as the increase in the administrative duties expected of them.[40,41] There have been a number of attempts to improve this, such as changing the allocation of nursing time from formal observation only to specific structured, individualised activities (often on a one-to-one basis), which improved patient-reported quality of care while the number of staff sicknesses and staffing costs decreased.[114] Hansen and Slevin[127] changed the structure of an acute care ward by introducing regularly held care planning meetings, group therapies and community meetings. Post implementation, service users on the ward reported significantly greater involvement in their own care, greater support between staff and service users, and greater opportunities to practise practical skills in preparation for discharge. A solution suggested by Mind[10] was to

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

attempt to improve staff retention which would hopefully lead to a decreased reliance on agency staff who may not have the skill or familiarity with the ward to run such activities.

A Sainsbury Centre for Mental Health[16] report provided a more upbeat picture, finding that 73% of ward managers across England reported that both practical therapeutic activities (e.g. cooking skills) and talking therapies occurred regularly on their wards, and 64% reported leisure activities (e.g. film nights) occurring regularly. The most commonly offered therapy was art therapy, on nearly half of all wards, compared with 35% for psychosocial interventions. Specifically, therapies with a strong evidence base, such as CBT, occurred on only 20% of wards. The recommendations include increasing the availability of training in therapeutic activities that have an evidence base. What the reports do not discuss is whether these activities were simply in a ward timetable or whether they actually occurred in reality. The lack of more evidence-based activities taking place is primarily as a result of not having sufficient numbers of nurses to run the wards and also fulfil other ward duties, the high number of temporary staff who may be unfamiliar with the activities and a shortage of OTs and psychologists.

Although not equivalent with formal therapies, service users also value informal ordinary conversations with nursing staff[34] and research indicates that, for service users, the most important aspect of quality inpatient care is that staff members take time and care with them.[115] The CQC[156] specifically said that when staff were questioned about activities, 96% reported activities on weekdays and 61% on weekends and evening, but when service users were questioned this dropped to 65% and 46%, respectively. Furthermore, Walsh and Boyle[34] and the Royal College of Psychiatrists[157] report that there is a dearth of activities in the evening and weekends, when service users would like them. This is echoed by NICE guidelines,[158] which state that meaningful activities should take place outside Monday to Friday and should not be restricted to 09.00 to 17.00. NICE guidelines further recommend psychological talking therapies for psychosis and other disorders for those who are hospitalised. A voluntary organisation called Star Wards offered tips and advice on a variety of things wards can do to improve their services, among which are both formal and informal activities.[159]

Bowers'[160] review of the changing nature of the types of service users and their needs specifically recommends that it would be within nursing staff's remit to provide evidence-based psychotherapy in inpatient settings. Cowls and Hale[161] found that inpatients also valued input from staff other than nurses and appreciated the activity-based groups as well as more psychoeducational groups. Indeed, the Royal College of Psychiatrists[157] go as far as to say that all staff should share the responsibility of engaging service users. Many training/accreditation programmes have been set up to address these very issues (e.g. Cowls and Hale[161]) in a variety of different ways. The Department of Health[153] and Clarke[162] outlined some of the ways in which many such programmes could have been improved. They suggested more service user and carer involvement in the design, implementation and subsequent evaluation of acceptability of inpatient services. They also suggested involving acute inpatient staff themselves at all stages, a broad range of health-care professionals, not just nurses, and adequate training and supervision for ward-based staff in addition to ward-based learning. The aims of our project were to implement such a programme to improve inpatient services and evaluate its impact rigorously.

The aims of the study are to:

- investigate in detail the difference that increasing therapeutic activities makes on service user and staff perceptions of the environmental milieu
- explore the sustainability of positive effects (particularly on staff morale and the level of increased activities)
- examine the impact of the programme on the perceived amount of care received and its costs.

Methods

The effects of the intervention were measured by service user and staff ratings made on three, five or seven occasions. The study was organised as a stepped-wedge design (a type of cluster randomised trial in which it is the timing of the intervention that is randomised). As all clusters eventually receive the intervention, this design is often used in situations when the researchers do not wish to deny the intervention to any particular cluster (*Figure 5*). Please see *Statistical methods* for more details and *Appendix 5* for a fuller discussion of this design. All service users entered the data set once only and were unaware of the condition (pre or post intervention) they were allocated (blind assessment) and all assessments were blind rated. Staff entered the data set if they had been on the ward during the assessment period. Non-permanent staff must have completed seven shifts in the past month. At the time of the assessment, staff working on control wards were unaware of when they would receive the training. Clearly, on those wards that had already been allocated to the intervention condition the staff knew what group they were in; however, all staff assessments were self-report. Service users were allowed to enter the data set only once, even if they had multiple or long admissions. Staff completed assessments each time they were available and had consented and so could have multiple entries in the data set.

Procedures

Ethics approval for the study was granted by Bexley and Greenwich Research Ethics Committee (reference: 07/H0809/49). Researchers approached all eligible patients following consultation with ward staff. Researchers gave each participant an approved information sheet explaining the study, answered any questions and gained signed informed consent. Participants were free to withdraw at any time. Handovers and ward meetings were

Ward	Time								
	1	2	3	4	5	6	7	8	9
1	0	1	1	1	1	–	–	–	–
2	0	1	1	1	1	–	–	–	–
3	0	0	1	1	1	–	–	–	–
4	0	0	1	1	1	–	–	–	–
5	0	0	0	1	1	–	–	–	–
6	0	0	0	1	1	–	–	–	–
7	0	0	0	0	1	–	–	–	–
8	0	0	0	0	1	–	–	–	–
9	0	0	0	0	0	1	1	–	–
10	0	0	0	0	0	1	1	–	–
11	0	0	0	0	0	0	1	–	–
12	0	0	0	0	0	0	1	–	–
13	–	–	–	–	–	–	0	1	1
14	–	–	–	–	–	–	0	1	1
15	–	–	–	–	–	–	0	0	1
16	–	–	–	–	–	–	0	0	1

FIGURE 5 Intervention schedule for DOORWAYS. 0 represents a ward for pre intervention, but for which we did collect data. 1 represents when a ward had received the intervention and data were collected. Time is measured in periods of 6 months.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

used to publicise the study with the nursing staff. Recruitment lasted for a period of 4 weeks at each of the time points. Data were stored in a locked office and entered into the database, and every entry was double checked by a different researcher to ensure accuracy.

The setting

This study was carried out in two mental health trusts and covers several distinct geographic areas.

In the first trust, we studied wards serving three borough catchment areas. The first borough serves an inner city population that has a deprivation index of 29.73 (ranking 41 out of 326 English boroughs[163]). Five wards participated in this study: three men's wards and two women's wards. Each had 18 beds. The catchment area also provided an intensive care ward comprising14 beds but this was not included in the study.

The second borough was located in a more suburban area, with a deprivation index of 22.76 (ranking 107 out of 326 boroughs[163]). Three wards participated in our study: one men's ward and two women's wards. Two wards had 22 beds. One of the women's wards was unique in that, although considered part of the acute service, it was a 'ward in the community'. This ward had only eight beds and accepted service users only if they were stable and not on a section. The ward was a detached house with a large garden located in a suburban street and the ward ethos was to provide as many therapeutic activities as possible.

The third borough had an average index of deprivation score of 31.24 (ranking 29 out of 326 boroughs[163]). Four 18-bedded wards provide acute inpatient care. One ward provided beds for women and one ward for men. One of wards was a specialist service, an early intervention unit accepting service users of both genders with psychosis from across the entire trust.

The final four wards were in a separate trust. This area was also mixed socioeconomically. Two of the wards served an urban borough with a deprivation score of 31.94 (ranking 28 out of 326 boroughs[163]). Both wards were mixed gender and had 18 beds. They also provided a PICU (not included in the study). The final two wards served two boroughs, both of which were suburban. The first had a deprivation score of 16.71 (ranking 174 out of 326 boroughs[163]) while the second had a score of 14.95 (ranking 203 out of 326 boroughs[163]). These were also mixed gender and provided 18 beds.

Participants

Service users using inpatient services on the participating ward at the time of data collection were eligible to take part in the study if they met the following inclusion criteria: they had been on the ward for a minimum of 7 days, they were able to communicate in English and they were well enough to consent. The exclusion criterion was that the participant had previously participated in the trial. There were no diagnostic exclusion criteria.

Nurses of any grade were eligible to be included in the study. Temporary agency staff were required to have completed seven shifts in the previous month in order to be eligible so that we could be sure that they were sufficiently familiar with the ward environment to complete the measures.

Main outcomes

For all measures obtained by questionnaires it was a requirement that > 80% of the questions were completed. Responses with < 80% completion were counted as missing.

Service user measure: VOICE[84]

Service user researchers constructed a measure from qualitative data, which was validated by expert panels of service users and tested for its psychometric properties. VOICE is easy to understand and complete and, therefore, is suitable for use by service users while in hospital. The 19-item measure has good validity and internal and test–retest reliability. The main outcome is the total score and we will carry out secondary analysis on the subscale scores.

Staff measure: VOTE[85]

VOTE is a 20-item measure that captures staff perceptions of the daily pressures of working in acute inpatient mental health wards. Each item is answered by choosing from a six-point Likert scale. The outcomes investigated were the total scores and, for secondary analyses, the factor scores as set out in *Chapter 2*. This score represents each staff participants' view and low scores are positive. Test–retest reliability and the internal consistency of the measure and subscales were good.

Secondary outcomes

Service user measures

Service Satisfaction Scale – Residential Services Evaluation[78]

This is a derivative measure adapted from the SSS-30,[79] designed to evaluate residential services for people with serious mental illness. The original SSS-30 has been used in a variety of settings and demonstrates sound psychometric properties.[164]

Rates of violence

For each service user who consented to participate in the study, we extracted data from routinely collected logs of violent incidents from the preceding 7 days. In this context, 'violent' can refer to either physical or verbal acts of aggression and includes both when the patient is the perpetrator of the violence and when they are the victim. Each incident was classified according to a 5-point scale; for this study we only considered incidents that were in one of the most serious three categories. It was possible for a service user to have been involved in more than one incident during the 7-day period and so the outcome measure of interest was the rate of these violent incidents. We collected these data within one trust only, so numbers of individuals will differ between the samples.

Global Assessment of Functioning[147]

The GAF is a 100-point rating scale of a patient's overall level functioning that is based on observation, interview and medical records.

Positive and Negative Syndrome Scale[165]

PANSS is a 30-item instrument that evaluates the presence, absence and severity of positive (e.g. hallucinations and delusions), negative (e.g. blunted affect and passive social withdrawal) symptoms and general psychopathology (e.g. anxiety, depression). Researchers were trained in how to conduct the interviews and were subsequently approved by a senior member of the research team (TC).

Nurses Observational Scale for Inpatient Evaluation[166]

Nurses Observational Scale for Inpatient Evaluation (NOSIE) is a 12-item scale designed for nurses, largely focused on the assessment of socially unacceptable behaviour. Nurses were asked to rate participants on items such as 'Sits, unless directed into activity'.

Client Services Receipt Inventory-Inpatient[142]

This instrument collects activity data on psychiatric inpatient wards. It is used to collect both ward activities and individual time that service users spend with health professionals. It assesses how much engagement service users report as well as having the ability to calculate the cost of such engagement using unit cost data from Curtis.[136]

Staff measures

Maslach Burnout Inventory-Human Services Survey[49]

The MBI can be totalled to produce a score which measures work related 'burnout' over 22 items, answerable by a six-point Likert scale. It has good psychometric properties and is widely used.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Index of Work Satisfaction[81]

The IWS is a 44-item scale that can be totalled to produce a score that measures health professionals' levels of job satisfaction. It is answerable using a seven-point Likert scale.

Background information

Service user measures

Demographic and clinical information: each participants' age, gender, marital status, ethnicity, education, employment status, legal status and number of previous admissions were recorded.

Staff measures

Demographic information on each participants' age, gender, ethnicity, employment band, length of employment and education, and previous employment in wards/trusts was gathered.

Ward data

Ward Atmosphere Scale:[167] The Ward Atmosphere Scale (WAS) is a well-established and validated measure of the ward environment consisting of 100 yes or no items. It was completed once a month throughout the entire study by two senior members of staff.

For each ward, we recorded the number of patients who had been on the ward in the last 30 days, gender of all patients, number of nurses on the ward, proportion of shifts undertaken by temporary staff, average length of stay, number of incidents on the ward, and ward acuity as indexed by average GAF, NOSIE and PANSS. We collected information on the number of activities provided on the wards and the number of service users participating in these from ward managers at each time point. We compared average frequencies before and after the intervention. In addition, we accounted for ward effects using a fixed-effects framework (see *Statistical methods*).

Intervention

Following consultation of NICE guidelines, trust clinical leads, ward managers and nursing staff directly involved with each ward, eight interventions were chosen, based on evidence of feasibility and outcome in wards. Ward managers suggested that three of these interventions should be compulsory. These were SCIT,[129] CBT-based communications training for nurses (cofacilitated by a service user educator) and computerised cognitive–remediation therapy (CRT) (to involve OTs). In addition, pharmacists were recruited to run a medication education group. The wards then had a choice to receive training in two of the following five interventions: Hearing Voices Group, Emotional Coping Skills Group, Problem Solving Skills, Relaxation/Sleep Hygiene and Coping with Stigma group. Each ward made their choices based on what activities were already occurring on the ward or the client mix. Please see www.perceive.iop.kcl.ac.uk/doorways.html[168] for a summary of each. A modified SCIT intervention was carried out on forensic wards as a value added study (see *Appendix 4*).

Statistical methods

Sample size

The total number of measurements originally envisaged to be available was seven wards and 15 patients per ward sampled over three time points after baseline, that is, a total of 315. The first six wards were to be randomised in pairs (30 patients each time). Assuming an intraclass correlation (ICC) of 0.05 with clusters of size 30, this sample size would have given approximately 80% power to detect a standardised effect size of 0.5 (moderate), using significance tests with alpha 0.05 (double sided). The patients were to be different at each time point so there was no additional clustering at the patient level to be taken into account. During the course of the project the number of wards was increased to eight and finally to 16. This design gave increased power (90%) with the assumptions above. Because of the stepped-wedge design the actual number of wards and service users in the intervention and control groups varied according to time point, so the above calculations were approximate.

Randomisation

The time at which wards were allocated to receive the intervention was randomised by one of the statisticians. Randomisation was carried out separately within boroughs and on three occasions as wards became available to receive the intervention. Random permutations of numbers were generated to designate the order of which wards were allocated to the intervention. The order that wards within a borough appeared in the randomly generated list determined the time point at which they were allocated to the intervention. This was performed using the *ralloc* procedure in Stata. After the initial baseline assessment period, two wards were randomised to receive the intervention with a further two wards randomised every 6 months until all wards had received the training. The design is shown in *Figure 5*.

Statistical analysis

The effect of the intervention was determined through effects on (1) service users, (2) staff members, including more general ward effects as well as individual level data, and (3) the cost of perceived care contacts and cost-effectiveness. In addition, data were collected to measure the demographic make-up of service users and staff, the atmosphere on the wards and any adverse events. We also recorded data relating to how the intervention was carried out, how many staff were trained and how many therapeutic sessions were run. All analysis was carried out using Stata versions 11 and 12.

Service users

The primary outcome measure of the effect of the intervention on service users was the VOICE score. This was investigated first using a *t*-test and second using a regression model. The model used was a Normal linear regression model. To match the design of the study, we always adjusted for the effect of ward and time. We used fixed-effects for both, using an indicator variable to account for the ward and a time variable which counted 6-month periods starting from 1 January 2009. We investigated linear and non-linear effects of time. As well as this main result, we also investigated if any covariates were confounding the intervention effect. We did this by regressing each covariate with the VOICE score and, separately, with the intervention variable. If there was evidence that the covariate was associated with both the intervention and the outcome, as judged by having a p-value of < 0.1 then it was included in the final model. This analysis was run on the total VOICE score as well as its two subscales: security and care. The same strategy was used to investigate the SSS-RES scores of service users. In addition, 95% CIs are reported for all estimates.

We also investigated interaction effects on the intervention outcome for voluntary patients and involuntarily admitted patients and two other variables (gender, ethnicity) identified a priori to be associated with VOICE.

The effect of the intervention on the rate of violent incidents on the wards was analysed using a Poisson regression model. This estimates a rate ratio, comparing the rate of violent incidents on post-intervention wards with a baseline of the rate of incidents on control wards. Thus, a ratio of < 1 indicates that the intervention reduced the rate of violence. Again, the effect of ward and time was always taken account of by using fixed-effect methods and both unadjusted and adjusted (for confounding covariates) estimates are reported.

We collected information on the number of activities provided on the wards and the number of service users participating in these from ward managers at each time point. We compared average frequencies before and after the intervention. In addition, we accounted for ward effects using a fixed-effects framework. To corroborate the staff data on the change in number and participants in the activities we also undertook an analysis of the activities data that were collected using CITRINE. Specifically, we compared the number of different activities and the total number of sessions that service users reported having participated in before and after the intervention. We used a bootstrapped fixed-effects model to obtain estimates of the intervention effect adjusted for ward characteristics and to take into account the non-normal distribution of the data. In addition, we tested for potential confounders including calendar effects as in the other analyses.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

We ran two focus groups on wards for which the intervention was being implemented to elicit service users' and staff views about the intervention.

Staff

The primary outcome measure for staff was the VOTE score. This was analysed using a regression model with a random effect to account for the effect of repeated measures being taken on staff. Ward and time were, again, accounted for as fixed effects by using an indicator variable to account for the ward and a time variable that counted 6-month periods starting from 1 January 2009. Non-linear effects of time were investigated. The same model was used on the VOTE subscales: workload intensity, team dynamics and interaction anxiety. This strategy was also used to analyse the MBI and IWS scores of staff.

The following variables derived from previous studies were considered as potential confounders: gender, age, ethnicity, employment band, first language and length of employment. Confounding factors were assessed and included in the model if they were associated with both the intervention and the outcome with a significance of $p < 0.10$. We performed exploratory analyses by including an interaction between the subgroup and the staff training arm and tested the nested models using likelihood ratio tests on all outcomes to explore potential differential effects between men and women, and between white and non-white staff.

Cost and cost-effectiveness

We estimated the cost of the intervention based on the cost of employing a clinical psychologist to lead the training sessions and the opportunity cost of nurses and OTs attending these training sessions [i.e. the cost of the time they could have spent performing other activities (see *Appendix 6* for costs)]. We discounted the cost of the intervention over the average time that the wards were followed up after the intervention in this study. We also compared the cost of perceived staff contacts before and after the intervention. Our regression analysis followed a similar format as the analysis of service user data, that is, we estimated a linear fixed-effects model adjusting for calendar and ward effects as well as other potential confounders. To allow for skewness and kurtosis in the cost data we calculated bootstrapped standard errors (SEs). Furthermore, we estimated the cost-effectiveness of the intervention by linking cost data with the two outcome measures, VOICE and VOTE, in turn. We plotted their joint distribution on a cost-effectiveness plane using confidence ellipses to delineate the area that bootstrapped replications were most likely to lie in. To further characterise the uncertainty around the point estimates, we constructed cost-effectiveness acceptability curves using a net benefit approach. These indicated the probability that care in the intervention group was more cost-effective than the control group for different monetary values placed on a unit improvement in the two primary outcome measures.

Results

Baseline data

A total of 1108 service users took part in the study. This is 70% of the population of service users who were eligible to participate at the time of the assessments. A total of 560 staff members consented to take part and we had valid data from 539 (444 pre-training intervention and 280 post training). Some contributed more than one set of data.

Service user participants

Just over half of service user participants were unmarried white men. Nearly three-quarters were not in employment and had experienced multiple previous admissions to inpatient care. *Table 16* shows a full description of the sample.

Staff participants

Table 17 describes characteristics at first point of entry to the study. For continuous variables (age and length of employment) these are mean averages of the variable at first point of entry to the study.

TABLE 16 Baseline characteristics of service user participants (WP3)

Description	Total number of service users (% in category)	Number on pre-training wards (% in category)	Number on post-training wards (% in category)
Overall trial population	**n = *1108***	**n = *670***	**n = *438***
Gender			
Men	609 (55)	352 (53)	257 (59)
Women	499 (45)	318 (47)	181 (41)
Age (years)			
≤ 20	49 (4.4)	25 (3.7)	24 (5.6)
21–30	259 (23.4)	163 (24)	96 (22)
31–40	280 (25.3)	171 (25.5)	109 (26)
41–50	272 (24.6)	163 (24)	109 (26)
51–60	156 (14)	87 (13)	31 (7)
≥ 61	92 (8)	61 (9)	20 (4.7)
Marital status			
Single/unmarried	746 (69)	448 (68)	298 (69)
Married/cohabiting	144 (13)	80 (12)	64 (15)
Other	218 (20)	142 (21)	69 (16)
First language			
English	879 (80)	525 (79)	354 (82)
Not English	219 (20)	141 (21)	78 (18)
Employment status			
Paid/self-employed	165 (15)	100 (15)	65 (16)
Volunteer	41 (3.8)	26 (4)	14 (3.4)
Unemployed	766 (72)	460 (71)	305 (73)
Other	136 (12.3)	83 (12)	32 (7.7)
Ethnicity			
White	556 (50)	325 (49)	231 (53)
Black	377 (34)	250 (37)	127 (29)
Mixed	71 (6.4)	38 (5.7)	33 (7.5)
Asian	59 (5.3)	30 (4.5)	29 (6.6)
Other	41 (3.7)	23 (3.4)	18 (4.1)
Chinese	3 (0.34)	3 (0.5)	0 (0)
Educated post 16 years			
Yes	338 (41)	246 (37)	140 (32)
No	722 (65)	424 (63)	298 (68)
Sectioned			
Yes	616 (56)	386 (58)	230 (53)
No	484 (44)	280 (42)	205 (47)
Number of times previously admitted			
Mean (SD)	3.6 (5.5)	3.6 (6.0)	3.5 (4.7)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 17 Baseline characteristics of staff participants (WP3)

Description	Total number of staff (% in category)	Number on pre-training wards (% in category)	Number on post-training wards (% in category)
Overall trial population[a]	**n = *560***[a]	**n = *444***	**n = *280***
Gender			
Men	200 (36.2)	171 (39)	98 (35.3)
Women	353 (63.8)	268 (61)	180 (64.7)
Age (years)			
≤ 20	5 (0.9)	4 (0.9)	1 (0.4)
21–30	158 (28)	122 (30.9)	66 (31.8)
31–40	153 (27.3)	132 (33.4)	21 (18.1)
41–50	125 (22.3)	101 (25.6)	68 (28.1)
51–60	36 (6.4)	28 (7.1)	23 (9.5)
≥ 61	83 (14.8)	8 (2)	26 (22.4)
Mean (SD)	37.4 (10.3)	36.9 (10.1)	38.2 (10.7)
Ethnicity			
White	216 (39.6)	167 (38.5)	114 (41.6)
Black	262 (48.1)	208 (47.9)	126 (46.0)
Mixed	15 (2.8)	14 (3.2)	5 (18)
Asian	37 (6.8)	32 (7.4)	18 (6.6)
Other	15 (2.8)	13 (3.0)	11 (4.0)
First language			
English	324 (61.1)	258 (60.4)	167 (62.5)
Not English	206 (38.9)	169 (39.6)	100 (37.5)
Education group (highest achieved)			
Vocational	136 (29.4)	109 (29.3)	97 (40.4)
University degree (undergraduate)	213 (46.1)	166 (44.6)	97 (40.4)
University degree (postgraduate)	113 (24.5)	97 (26.1)	46 (19.2)
Length of employment on ward			
< 2 months	36 (6.9)	27 (6.4)	9 (3.4)
3 months to < 6 months	119 (22.7)	84 (19.9)	36 (13.6)
6 months to < 12 months	72 (13.7)	56 (13.2)	33 (12.5)
1 year to < 3 years	101 (19.2)	87 (20.6)	67 (25.4)
3 years to < 6 years	93 (17.7)	48 (15.5)	48 (18.2)
6 years to < 10 years	61 (11.6)	87 (20.6)	37 (14)
≥ 10 years	43 (8.2)	37 (8.7)	34 (12.9)

TABLE 17 Baseline characteristics of staff participants (WP3) (*continued*)

Description	Total number of staff (% in category)	Number on pre-training wards (% in category)	Number on post-training wards (% in category)
Band (highest achieved)			
2	41 (8.0)	33 (8.0)	20 (7.7)
3	139 (27.2)	113 (27.3)	69 (26.4)
4	3 (0.6)	3 (0.8)	1 (0.4)
5	231 (45.2)	184 (44.4)	117 (44.8)
6	67 (13.1)	59 (14.3)	37 (14.2)
7	28 (5.5)	21 (5.1)	15 (5.8)
8	2 (0.4)	1 (0.2)	2 (0.8)

a Some data are missing as the staff participants wanted to keep their anonymity. We also counted data on the first occasion that a staff participant entered the data set (e.g. for age).

The design of study allowed staff to be part of the study on more than one occasion. The majority only took part once, but *Table 18* shows the breakdown of how many staff contributed more than once.

On average, 51.6 nurse training sessions took place on each of the wards. This includes some nurses who attended multiple sessions.

Wards

Sixteen wards were included in this trial. Five wards were men's wards, five were for women and six were mixed. All baseline data reported in this section were collected during the first data collection period for each ward. We used the WAS to record the views of one or two senior staff of their wards and took an average of their responses as our measure of the ward atmosphere (*Table 19*).

Outcomes and estimation

Did service users perceive a change in the ward following the introduction of activities?

High scores on VOICE are indicative of a poorer view of the ward. The mean VOICE score was 56.5 in the pre-intervention wards and 54.2 for the post-intervention wards, suggesting an improvement. However, an investigation of service user perceptions of the therapeutic environment showed no improvement with introduction of the intervention. A regression model adjusting for only ward and time estimated the standardised intervention benefit as 0.19 (mean VOICE score pre-intervention = 56.5, SD = 19.1, $n = 644$; mean VOICE score post-intervention = 54.2, SD = 17.2, $n = 414$). The only confounder identified was legal status (an a priori moderator) and the adjusted model (*Table 20*) provides weak evidence for benefit

TABLE 18 Number of data collection periods to which staff contributed (WP3)

Number of data collection periods	Number (%) of staff
1	274 (48.9)
2	133 (23.8)
3	73 (13.0)
4	51 (9.1)
5	25 (4.5)
6	4 (0.7)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 19 Baseline characteristics of wards (WP3)

Ward	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
WAS	61.4	70.3	68.3	71.0	72.7	58.6	60.9	49.8	61.8	–	58.4	–	52.0	–	66.0	–
Service user age (years)	42.4 (11.6)	36.7 (11.3)	38.6 (10.5)	42.2 (10.8)	41.6 (11.9)	39.2 (14.8)	42.8 (12.5)	39.2 (10.6)	46.4 (13.5)	42.6 (12.7)	28.1 (9.0)	42.9 (12.8)	42.8 (14.5)	39.6 (12.5)	34.7 (12.8)	38.1 (9.6)
Service user gender[a]	0	0	100	100	0	100	100	0	100	0	45	46	48	54	44	43
VOICE	53.6 (18.2)	52.4 (19.2)	35.5 (10.2)	55.3 (15.7)	57.7 (21.1)	53.3 (14.5)	57.2 (14.4)	61.4 (20.3)	63.5 (19.5)	54.2 (19.9)	51.3 (18.1)	55.7 (20.6)	48.3 (17.7)	52.9 (11.9)	51.3 (12.5)	48.6 (16.9)
SSS-RES	83.3 (27.6)	92.8 (23.5)	71.1 (12.6)	93.9 (26.6)	86.2 (28.4)	88.3 (21.8)	102.0 (27.3)	98.6 (28.6)	100.7 (24.0)	89.4 (29.0)	86.1 (30.0)	96.1 (30.3)	78.0 (29.9)	87.4 (18.9)	82.7 (19.0)	83.0 (27.9
NOSIE	19.5 (7.7)	16.6 (9.4)	14.6 (9.5)	16.8 (5.9)	21.2 (8.6)	14.4 (7.1)	15.6 (9.2)	15.7 (10.5)	16.0 (7.8)	20.8 (8.1)	12.2 (6.5)	15.7 (7.7)	11.5 (7.7)	14.0 (8.3)	16.3 (6.8)	14.3 (4.7)
PANSS	49.6 (8.2)	58.5 (20.1)	55.4 (10.8)	58.3 (15.8)	54.5 (8.9)	54.0 (12.1)	56.0 (16.3)	54.1 (11.1)	69.6 (18.7)	40.7 (4.2)	66.7 (20.3)	40.8 (6.0)	62.1 (12.8)	45.6 (12.4)	46.3 (11.2)	42.3 (8.4)
GAF	49.6 (8.3)	40.9 (9.7)	48.8 (4.9)	39.9 (9.8)	39.0 (7.9)	39.6 (9.6)	43.6 (9.6)	36.2 (10.4)	40.3 (12.1)	39.3 (13.1)	43.3 (13.0)	37.7 (17.0)	42.1 (10.8)	47.0 (7.8)	54.4 (12.5)	51.7 (14.6)
Change in HONOS	–2.0 (6.7)	–2.1 (5.4)	–4.4 (5.8)	–7.2 (10.7)	–5.4 (4.9)	–3.8 (7.8)	–3.3 (11.0)	–3.5 (5.3)	–7.0 (7.7)	–3.2 (10.3)	–5.1 (6.0)	–3.1 (6.2)	–	–	–	–
Staff age (years)	37.9 (7.9)	43.5 (9.9)	44.3 (4.8)	40.0 (9.9)	40.2 (8.5)	39.9 (12.6)	35.0 (9.3)	35.7 (8.9)	36.6 (11.0)	36.4 (11.0)	32.1 (7.2)	39.4 (9.4)	35.6 (12.4)	36.1 (10.4)	40.8 (6.1)	39.4 (12.4)
Staff gender[a]	20.0	53	100	83	50	83	80	48	64	63	44	52	52	50	50	80
VOTE	70.9 (9.06)	65.3 (12.1)	61.1 (6.9)	69.5 (7.8)	69.7 (14.1)	77.6 (15.8)	72.6 (15.8)	70.5 (16.0)	71.5 (11.8)	63.9 (9.9)	60.1 (10.0)	65.8 (11.8)	73.5 (12.0)	66.8 (9.2)	63.9 (6.8)	74.2 (10.7)
MBI	42.5 (15.4)	36.9 (16.0)	40.3 (21.5)	38.0 (18.0)	42.8 (24.0)	46.4 (13.7)	41.6 (16.3)	52.2 (17.2)	51.0 (20.5)	50.7 (14.8)	47.8 (18.8)	54.2 (12.4)	44.4 (16.5)	46.4 (17.0)	38.4 (13.4)	39.0 (16.1)
IWS	169.3 (15.7)	154.5 (27.4)	164.0 (23.5)	157.5 (21.1)	163.4 (39.9)	184.6 (29.6)	171.0 (31.0)	175.2 (37.3)	157.2 (24.9)	144.5 (26.1)	143.9 (30.3)	157.2 (24.9)	162.1 (28.7)	167.6 (28.5)	168.2 (25.6)	178.3 (27.4)
Cost (£/week)	117 (184)	254 (374)	479 (261)	143 (219)	323 (185)	112 (100)	459 (863)	329 (403)	257 (171)	289 (344)	362 (367)	78 (82)	344 (317)	181 (203)	331 (439)	279 (182)

a Percentage of female service users/staff.
Values are mean (SD) or percentage of baseline characteristic.

TABLE 20 Adjusted estimate of intervention effect on VOICE standardised scores ($N = 1051$) (WP3)

Factor	Coefficient	SE	95% CI	*p*-value
Intervention	–0.18	0.10	–0.38 to 0.01	0.06
Section: involuntary	Baseline			< 0.0001
Section: voluntary	–0.27	0.07	–0.40 to –0.15	
Constant	0.10	0.12	–0.14 to 0.34	0.66

(standardised effect –0.18, 95% CI 0.38 improvement to 0.01 deterioration; $p = 0.062$). We found two other effects (independent of treatment or ward); a deterioration in VOICE score over time by 0.06 SD per month (95% CI 0.01 to 0.12; $p = 0.021$) and, over the whole trial, voluntary patients were more positive about the ward environment than involuntary patients by 0.27 SD (95% CI –0.40 to –0.15; $p < 0.0001$).

We also investigated the effect of legal status on admission (voluntary vs. involuntary) and other potential moderators. We found a significant interaction only with legal status ($p = 0.006$), with good evidence that the intervention improves VOICE scores of people admitted involuntarily [standardised improvement of –0.35 (95% CI –0.12 to –0.57; $p = 0.002$)]. Among people in hospital voluntarily, we found no evidence of an intervention effect (standardised effect = –0.01, 95% CI –0.23 to 0.22; $p = 0.955$).

We explored the components of the perceptions of the therapeutic environment by performing analyses on the two subscales of the VOICE (security and care subscales). The mean score on the security subscale was 25.0 for pre-intervention wards and 23.6 for intervention wards. Results from the adjusted analysis showed that the average intervention effect was an improvement on the security subscale of 0.27 SD, although this was not statistically significant (95% CI –0.53 to 0.00; $p = 0.05$). A similar effect with time was found in this model as for the primary analysis of the VOICE total scale.

The mean score for the second VOICE subscale (care) was 27.4 for pre-intervention wards and 26.6 on intervention wards. The analysis of this outcome provided no evidence of an improvement in the intervention arm (average reduction in intervention arm 0.18 SDs, 95% CI –0.37 to 0.01; $p = 0.07$). This model included a linear effect with calendar time, although no statistically significant trend was found.

Post-intervention focus groups

The purpose of the focus groups was to find out how service users perceived the intervention. Most participants were not aware that there was a programme of activities, although some could identify particular therapies. Participants told us that often the planned groups did not take place. For those groups that they could identify, there were mixed views. Some found the groups useful, especially in filling up the day. Conversely, some said the groups were repetitive or even 'childish'. Service users also reported that the quality and style of the activities was dependent on specific nurses. They felt some staff were more knowledgeable or better able to engage with them than others.

What effects did activities have on staff views of the wards?

A total of 677 VOTE scores were from 428 individuals on pre-training wards (mean 68.2, SD = 12.0) and 398 scores were from 271 individuals on post-training wards (mean 70.3, SD = 12.7). Results from the regression model showed no evidence of a change in VOTE scores [standardised effect size = 0.04, 95% CI –0.09 to 0.18; $p = 0.54$ (see *Table 21*)].

Our exploratory analysis examined effects for staff subgroups and suggested a difference between men and women, but not for ethnicity. The model including an interaction between trial arm and gender estimated a negative effect for men (0.19 SD, 95% CI 0.01 to 0.36; $p = 0.02$) and a non-significant benefit for women (0.05 SD 95% CI –0.20 to 0.11; $p = 0.55$).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 21 Adjusted results for VOTE and VOTE subscales (standardised scales) (WP3)

	Training intervention effect	95% CI	SE	*p*-value	Observations (staff)	ICC
VOTE	0.04	(–0.09 to 0.18)	0.07	0.54	1075 (546)	0.61
VOTE (workload intensity)[a]	0.02	(–0.13 to 0.18)	0.08	0.78	1076 (526)	0.47
VOTE (team dynamics)[b]	0.08	(–0.08 to 0.23)	0.08	0.34	996 (488)	0.52
VOTE (interaction anxiety)[b]	0.001	(–0.15 to 0.15)	0.08	0.99	996 (488)	0.59

a Adjusted for length of employment.
b Adjusted for age.

The VOTE score consists of three subscales, measuring workload intensity, team dynamics and interaction anxiety. Using the same model as above, workload intensity mean score prior to training was 31.0 (SD = 6.4) and 32.2 (SD = 7.7) post training. Lower scores indicate poorer perception of workload intensity. A regression-adjusted time and ward and the confounding factor of length of employment provided no evidence ($p = 0.78$) that the training altered nurses' views on their workload intensity. The estimated intervention effect was 0.02 on a standardised scale (95% CI –0.13 to 0.18), representing a slight worsening of VOTE scores on post-intervention wards (see *Table 21*).

Team dynamics mean score at pre-intervention was 17.7 (SD = 5.3) and 18.2 (SD = 5.5) post training. Higher scores again indicate deterioration in the perception of team dynamics. Using a regression model adjusted for the effect of age we found no evidence ($p = 0.34$) that the training led to an improvement in nurses' perceptions of team dynamics. The estimate of effect was a deterioration of 0.08 SDs (95% CI –0.08 to 0.23) (see *Table 21*).

Interaction anxiety mean score prior to training was 18.6 (SD = 4.4), and 18.5 (SD = 4.5) post training. Higher scores indicate deterioration. The regression model adjusted for nurses' age found no evidence ($p = 0.99$) that training changed staff perceptions of interaction anxiety. The estimated effect was a worsening of 0.001 SDs (95% CI –0.15 to 0.15; see *Table 21*).

Post-intervention focus groups

Nurses reported that, overall, they enjoyed running the activities and appreciated the opportunity to feel that they were doing something therapeutic rather than paperwork and crisis management. They did report that this was hard to consistently carry out, in particular when high numbers of the staff were temporary staff or there were not enough staff on duty. They did not feel that they had the necessary organisational support to do the best job. They also said that they felt that the activities were somewhat repetitious, that is, service users sometimes got bored of the activities on offer, especially when they were in hospital for long periods of time. They said that they often felt bored running the same groups, but appreciated that being trained in more numerous activities would not have been possible. They also said that service users had repeatedly reported that they would prefer to personalise the groups (e.g. bring real-life problems to the problem-solving groups rather than hypothetical ones). Overall, they had a positive view of the activities but practicalities often got in the way.

Costs and cost-effectiveness

What is the effect of the intervention on the cost of perceived contacts?

We estimated that the total cost of the intervention across the 16 wards was approximately £156,000 which amounted to a cost of £10 per patient-week given an average number of 18 patients per ward and an average post-intervention follow-up of 55 weeks (*Table 22*). About 80% of this cost was attributable to the opportunity cost of nurses attending the training sessions. The average cost of perceived staff contacts over the course of a week was £259 before the intervention and £242 after the intervention. Our regression analysis suggested that the intervention increased perceived costs by £12 (95% CI –76 to 98; *p*-value = 0.774) but this difference was not statistically significant (*Table 23*).

TABLE 22 Cost of the intervention (WP3)

Item	Number of units	(Average/total) number of participants	Number of sessions	Session duration (hours)	Cost per hour (£)	(Sub)total cost (£)
Opportunity costs of time spent training						
Nurse	16	53	1	3.5	41	121,258
OT CRT training		19	7	1	34	4522
Clinical psychologist instruction costs						
Training Sessions	16	1	4	3.5	60	13,440
Activity cofacilitation	16	1	17	0.5	60	8160
CRT		19	7	1	60	7980
Total						155,360

TABLE 23 Results for cost of perceived service use (WP3)

Model	Coefficient	95% CI	SE	*p*-value
Average intervention effect	12	–76 to 98	44	0.77
Time effect (6 months): linear	14	–36 to 5	10	0.14
Intercept	293	238 to 360	32	< 0.001

Figure 6 shows the uncertainty surrounding the effects of the intervention on cost and VOICE estimates in combination. We estimated that the intervention most likely improved VOICE scores but increased costs of service contacts at the same time (i.e. 50% of our bootstrap replicates were in the north-west quadrant). It was unclear what the value of increasing the score by one unit was but, even at a willingness to pay of £100 per improvement in VOICE, the probability that the intervention was cost-effective was low

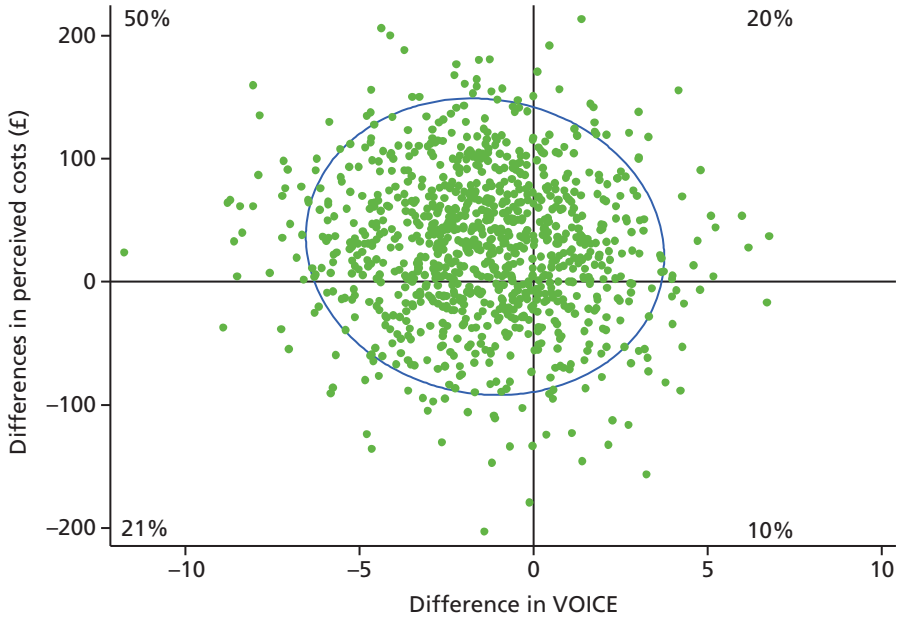


FIGURE 6 Cost-effectiveness plane using VOICE with 95% confidence ellipse (WP3).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

(*Figure 7*). In the cost-effectiveness analysis using VOTE, results differed somewhat from the previously reported separate estimations because to combine these two end points, the analysis had to be undertaken at the unit level. In the cost-effectiveness acceptability plane using VOTE, the probability that the intervention led to deterioration of staff perceptions and reductions in costs was highest with 41% followed by the probability that the intervention would lead to deterioration in VOTE and an increase in cost with 29% (*Figure 8*). Regardless of the willingness to pay for improvements in VOTE, our analysis suggested that, with a probability of being cost-effective of 40%, it was unclear whether or not the intervention was cost-effective if staff perception was the only relevant outcome measure (*Figure 9*).

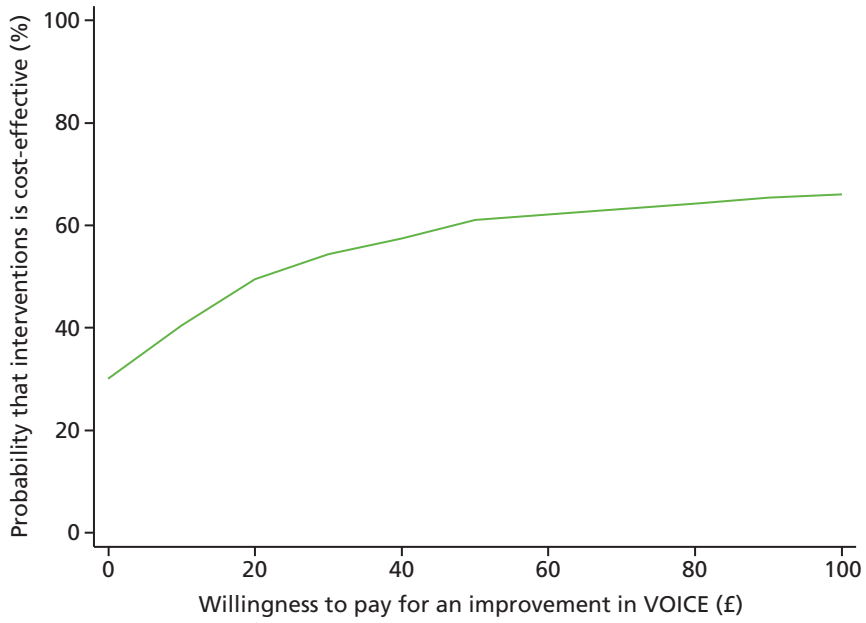


FIGURE 7 Cost-effectiveness acceptability curve using VOICE (WP3).

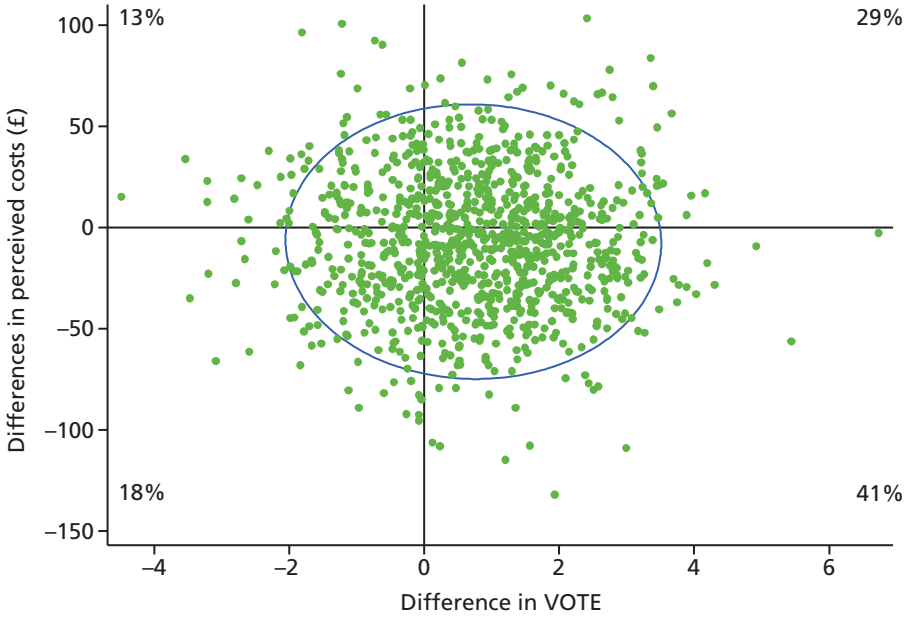


FIGURE 8 Cost-effectiveness plane using VOTE with 95% confidence ellipse (WP3).

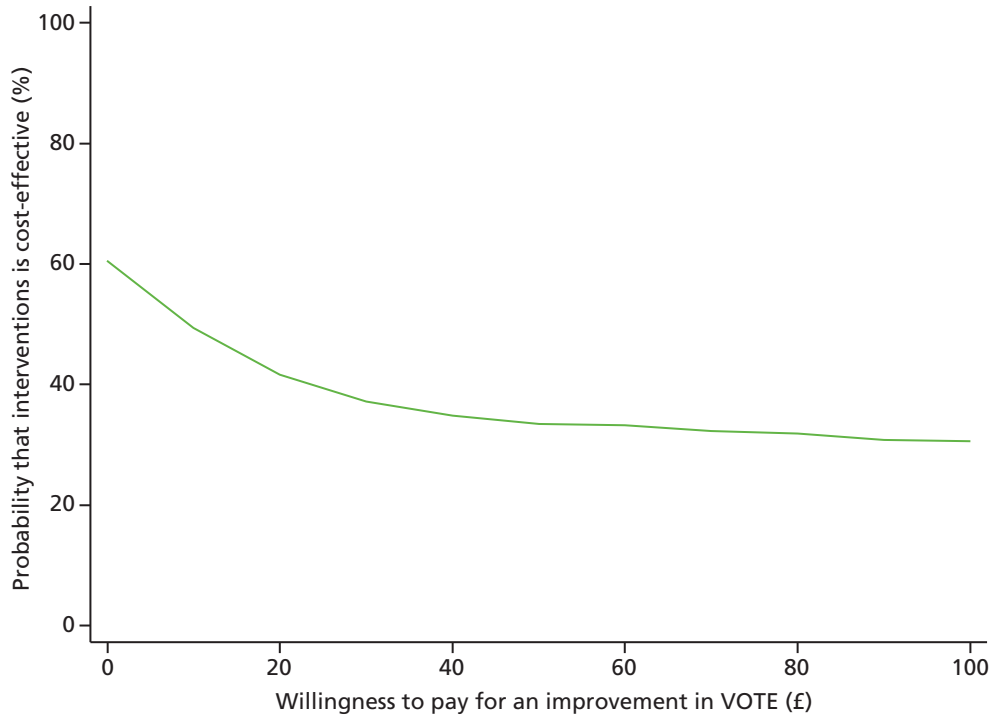


FIGURE 9 Cost-effectiveness acceptability curve using VOTE (WP3).

Ancillary analyses

Service user measures

A total of 1032 patients completed the SSS-RES [625 on pre-intervention wards (mean 91.3, SD = 27.1) and 407 on post-intervention wards (mean 86.4, SD = 24.2)]. A linear regression model suggested an intervention benefit of 4.15 points (95% CI –9.22 to 0.92; $p = 0.109$). As with the VOICE measure, there is good evidence ($p = 0.005$) for an interaction effect with legal status. For those who are compelled to accept treatment, the intervention benefit was estimated as –8.44 (95% CI –14.36 to –2.52; $p = 0.005$), but no evidence of a treatment effect in voluntary patients (0.61; 95% CI –5.39 to 6.60; $p = 0.842$).

Violence The overall level of violence recorded was low. A total of 67 violent incidents were recorded during the trial; 40 of these occurred on control wards and 27 on post-intervention wards. Accounting for the number of patients on these wards gives a naive rate ratio of 0.68.

Using a Poisson regression model adjusted for the effects of ward and time we found that the intervention reduced the rate of violent incidents by an estimated 14%. However, there is no evidence ($p = 0.73$) that this result is not due to chance. The 95% CI for this estimate is very wide, ranging from a 63% reduction to a 99% increase. We found there was only a slight time trend, with an estimated 2% increase in the rate of incidents every 6 months. Again, there was a wide 95% CI for this estimate, ranging from a 26% reduction to a 40% increase. We did not find any covariates that met our criteria for confounding (*Tables 24* and *25*).

Staff measures

Index of Work Satisfaction The mean IWS score was 160.4 for pre-intervention wards and 165.3 for intervention wards. The analysis, adjusted for time and ward, found no evidence of a change following the intervention (average standardised intervention effect = 0.02, 95% CI –0.11 to 0.16) (see *Table 26*).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 24 Number of violent incidents per service user (WP3)

Number of incidents per service user	Pre intervention	Post intervention
0	413	418
1	24	15
2	6	3
3	0	2
4	1	0
Total incidents	40	27
Total number of service users	444	438

TABLE 25 Unadjusted results for the rate of violence ($n = 1108$ from one NHS trust) (WP3)

Model	Coefficient (rate ratio)	95% CI	SE	*p*-value
Average intervention effect	0.86	0.37 to 1.99	0.37	0.73
Time effect (6 months): linear	1.02	0.74 to 1.40	0.16	0.92

TABLE 26 Adjusted results for IWS and MBI standardised scales (WP3)

	Intervention effect	95% CI	SE	*p*-value	Observations (staff)	ICC
IWS	0.02	(–0.11 to 0.16)	0.07	0.74	1028 (523)	0.67
MBI[a]	–0.09	(–0.24 to 0.06)	0.07	0.24	901 (457)	0.63
MBI (Emotional Exhaustion)[a]	–0.11	(–0.26 to 0.04)	0.08	0.16	930 (470)	0.58
MBI (Depersonalisation)[a]	–0.02	(–0.19 to 0.15)	0.09	0.83	924 (464)	0.49
MBI (Personal Accomplishment)	0.03	(–0.14 to 0.19)	0.08	0.73	1012 (521)	0.44

a Adjusted for age.

Maslach Burnout Inventory The mean MBI was 39.5 and 40.7 for pre-intervention and intervention wards respectively. The analysis adjusted by staff age showed no evidence of a significant difference in burnout following the intervention (standardised average intervention effect = –0.09, 95% CI –0.24 to 0.06) (see *Table 26*).

We explored the specific components of the MBI scale to detect differential change with time. The MBI consists of three subscales: emotional exhaustion, depersonalisation and personal accomplishment. Personal accomplishment is interpreted in the opposite direction to the first two (i.e. higher scores denote improvement).

For the factors in the MBI we compared the factors with the norms for mental health staff in the MBI Handbook.[49] We found that emotional exhaustion (EE) was higher both before and after intervention (pre intervention 21.5 and post intervention 22.3, compared with the MBI handbook, 16.89). Depersonalisation was about the same, but personal accomplishment (PA) was higher (i.e. better) (pre intervention 35.5 and post intervention 35.2, compared with the MBI Handbook, 30.87). However, no factors were affected by the training (see the analysis results in *Table 26*).

Did we increase the number of ward activities?

According to data obtained from ward records, the mean number of activities before the intervention was 6.3, whereas after the intervention it was 7.8. The regression analysis yielded a mean treatment effect estimate of 1.5 (95% CI –0.4 to 3.4; $p = 0.12$). Before the intervention, the average number of participants in activities was 28 and it was 39 after the intervention. The average treatment effect in the regression model was 6.3 (95% CI –4.1 to 16.6; $p = 0.23$). The average number of activities that service users reported to have attended before the intervention was higher than after the intervention (1.91 vs. 1.89) but, after adjusting for ward and calendar effects, there was a trend towards a positive effect on the number of activities attended (coefficient 0.58, 95% CI –0.01 to 1.21; $p = 0.06$). Women were more likely to report having taken part in activities but adjusting for the effect of gender in the regression led only to a marginal change in the treatment effect estimate (coefficient 0.59, 95% CI –0.02 to 1.22; $p = 0.59$). Although the average number of activities attended over 1 week was lower after the intervention than before the intervention (3.18 vs. 3.38), our regression model suggested that the number of sessions attended increased by 0.68 (95% CI –0.67 to 2.13; $p = 0.71$) after adjusting for ward and calendar effects. This difference was not statistically significant (*Tables 27–29*).

Adverse events

No adverse events were reported as a result of the intervention.

TABLE 27 Ward reported activities (adjusting for ward effects) (WP3)

Outcome	Average intervention effect	95% CI	SE	*p*-value
Number of activities provided	1.48	–0.40 to 3.37	0.94	0.12
Total number of participants	6.27	–4.05 to 16.59	5.10	0.23

TABLE 28 Results for self-reported (CITRINE) number of activities attended (WP3)

Model	Coefficient	95% CI	SE	*p*-value
Average intervention effect	0.59	0.02 to 1.22	0.31	0.059
Calendar effect: linear	0.22	–0.45 to –0.01	0.12	0.05
Female	0.46	0.15 to 0.69	0.13	< 0.01
Intercept	2.14	1.61 to 2.83	0.31	< 0.01

TABLE 29 Results for self-reported (CITRINE) number of sessions attended (WP3)

Model	Coefficient	95% CI	SE	*p*-value
Average intervention effect	0.68	–0.67 to 2.13	0.97	0.24
Calendar effect: linear	0.04	–0.87 to 0.04	0.46	0.08
Female	0.3	–0.04 to 0.59	0.16	0.06
Intercept	4.14	2.91 to 5.73	0.71	< 0.01

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Discussion

Summary

The DOORWAYS study sought to improve the perceptions of the ward environment among those living and working on acute inpatient mental health wards by training staff to provide activities and increase the number and quality of therapeutic activities taking place on these wards. For service users, the measured perceptions of the environment moved in the direction expected and, for those who were admitted compulsorily, this was significant. But nursing staff had more complex views on the inpatient environment. We also found that we were able to make an impact on the day-to-day life of the wards in a number of ways. However, there was only limited indication that the intervention was cost-effective at the service user level (depending on how much a unit improvement in the VOICE is valued) and even less likelihood that the intervention was cost-effective at the ward level with regard to the staff measures, especially our main outcome, VOTE.

Impact on wards

The associations between social interaction and taking part in therapeutic activities and their impact on service users' improvement is not a new one.[11] Not only do activities break up the monotony of time spent on wards,[10] they also provide a forum in which service users can interact with each other, which we know is important.[34] Therefore, we attempted to increase the activities on the wards and had moderate success in this. As the intervention consisted of training nurses to deliver activities, we had no control over the actual number of sessions that were actually run, apart from the ones the psychologist was involved in. In spite of this, we were successful in increasing the average number of activities on the 16 wards in this study. All of these wards already had a set weekly activity schedule so it is likely that some of the existing activities were dropped and replaced by the ones in this study. This may be one reason why our improvement of an average of two extra activities a week was not statistically significant. Sainsbury's Centre for Mental Health[16] specifically recommended that more evidence-based therapeutic activities are run on wards, while NICE[158] recommends psychological talking therapies be offered to those with psychosis. Although very much valued,[159] the majority of groups already being run on the ward were activities such as cookery or bingo, and so we believe that the introduction of evidence-based therapeutic activities was a positive one. The cost of the intervention was small; spending per patient did not increase and, in fact, may have decreased over time as budgets were cut. However, the number of contacts and activities did increase, which service users considered to be valuable. Furthermore, it is probably true, as others have found,[40,41] that administrative duties and crisis management often means that therapeutic activities are relegated to the bottom of the ward's to-do list. Service users in the focus groups recognised this:

> *Sometimes the nursing staff, they're quite occupied with what they're doing around the unit. They don't really have much time, they very rarely have much time to themselves to tell you the truth, they probably have a break once in the shift.*
>
> *Anonymous service user*

It is perhaps not surprising that our study was not able to singlehandedly change this. In light of that, increasing the mean number of activities from six to eight can be considered a success.

One of the results of the increased pressure on inpatient wards is the tinderbox atmosphere that is associated with violence.[8,9] During the period involved in the study, we recorded a total of 67 incidents of violence. Forty of these were on non-intervention wards while 27 were on intervention wards. We estimated this to be a 14% reduction in the rate of violence on intervention wards. Although we cannot claim statistical significance for this result, changes in relatively rare events are difficult to find and the study was not powered for this secondary outcome. Viewed in this light, and knowing about the background service pressures, this was a positive outcome and is indicative of an improved atmosphere on the wards following our intervention.

Impact on service users

The primary aim of this study was to improve the experiences of service users, which has been highlighted in many recent reports as vital in evaluating the performance of a service (e.g. Francis[18] and CQC[156]). Our study, which began in 2009, was ahead of this recent focus. Although both perceptions and satisfaction of those using services moved in the right direction, we did not reach the generally accepted threshold of statistical significance for all those on the ward and so we cannot rule out the role of chance in producing these differences. But we were able to show improvements for those who are admitted under a legal section. We chose to investigate longer-term effects as these seem to have more relevance for the NHS, whereas other studies of changes in inpatient care have tended to investigate shorter-term effects, such as a few weeks. As wards aim to discharge service users as soon as they are able to do so, service users do not typically spend 6 months on the ward and so it was not possible to use repeated measures to assess their views of how activities had improved the environment, but this would have increased the power to detect small but consistent effects. We also only analysed the intention-to-treat effects of providing an intervention (training staff) on the outcome (change in perceptions of the ward). This is shown as the link between the intervention and the effects in green in the model (*Figure 10*). However, the effects may not be direct and the model shows other potential linkages that we might investigate through our database in the future. For instance, the effects of the intervention might be mediated by the effect on the number of interventions that are provided. Moderators include the turbulence on wards, such as the changes we noticed in ward leadership.

Furthermore, this study was set against a backdrop of budget cuts and service reorganisations, which clearly had an impact on the wards, producing variability and a secular time trend. Our study also did not have sufficient power as it was not possible for us to collect data from all wards at all time points because we extended the study. Running such a complex trial on a large number of wards is a significant logistical challenge and will always present serious practical and analytical difficulties to researchers.

Our baseline analyses[169] found that both activity and one-to-one sessions were associated with service users' perceptions of the quality of the care they were receiving. These findings demonstrate that even those on an acute ward can engage in and appreciate the activities provided and the meaningful interactions they can have with staff, which is reflected in their perceptions of the ward. An often repeated justification for not running activities in the first place is that people are too ill to take part. Our study proved that this is not

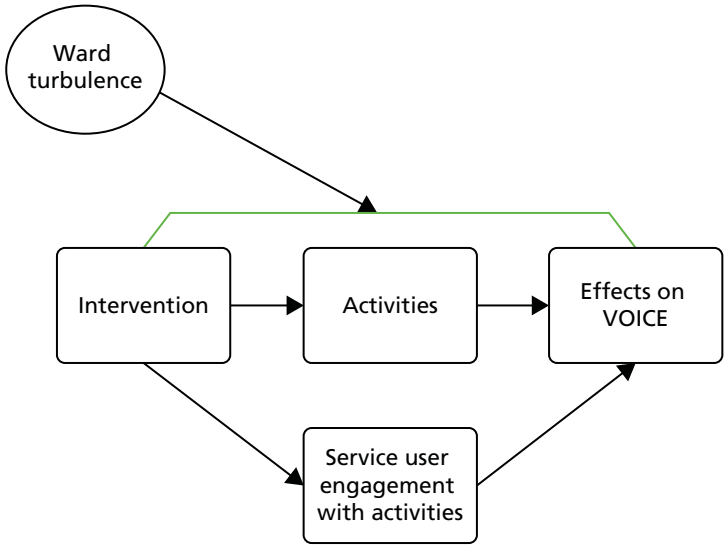


FIGURE 10 Model of potential relationships between intervention and outcome (WP3).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

the case and that symptom severity was not associated with taking part in fewer activities.[169] Furthermore, our post-intervention focus group supported this finding:

> *I found the recognition one with the faces where you start off with the blank face and the face changes quite interesting because it shows that you can't judge someone by a first look especially in a place like this because there's lots of emotions and feelings going on in the place so it's actually to get to know the person behind the, what they, I found that quite interesting.*
>
> *Anonymous service user*

Nursing staff can also take heart in this finding that their efforts are recognised, appreciated and valued.

Feedback from our focus groups showed that many service users were not aware of the activities taking place. This could be for a number of reasons: a lack of clearly displayed ward timetables, no introduction to the ward on admission, no mention of it in the community meetings, and no one inviting them to take part. Furthermore, both service users and staff reported that activities were sometimes cancelled at the last minute, but our data still suggests, on average, an increase in activities over time despite these cancellations. There was also some feedback that the activities, or the way in which they were run, could have been modified. Both groups reported that the activities became boring over time with repetition:

> *Yeah they repeat the same things again. That's the only thing that I find difficult. They say 'oh DOORWAYS meeting' and I think oh what's the point in going to a DOORWAYS meeting when it's the same thing as last week. So there needs to be a bit more of a change going on.*
>
> *Anonymous service user*

An alternating schedule of activities planned over several weeks might have overcome this problem. Service users also reported that some activities were childish or like being in school again, so perhaps the complexity/level of the activities needs to be raised.

Impact on nursing staff

The results from the nursing staff were more complex. In spite of DOORWAYS requiring nurses to engage and interact with service users more frequently, there were no effects on interaction anxiety, as measured by the VOTE subscale. Nurses on the wards in the post-intervention focus groups reported that their confidence was growing following training. In the post-DOORWAYS focus group, a nurse described how they felt after our project finished:

> *Because we were a bit unsure of ourselves . . . But I think once we got into it we didn't have a problem at all, we got more and more confident and it's not a problem anymore to, and I think it's helped us to run other groups because we're more confident in that so now we can run other groups as well without any problem.*
>
> *Anonymous nurse*

We did find that workload intensity, job satisfaction and overall VOTE scores worsened slightly but not significantly and therefore opposite to our hypothesis of an improvement. Much of this effect we consider to be a reflection of the changes in the local NHS services during the study. Three out of the four catchment areas in our study reduced the number of wards and budgets were also reduced. In one case, staff had to reapply for their own jobs and some wards did not have ward managers but were run by junior staff or overseen on a part-time basis only. In light of this turbulent background, it is not surprising that being asked to deliver more activities had an effect on their view of the ward. This consequence was also reflected in an increase in the MBI emotional exhaustion scale. However, it is interesting that there was no effect on the personal accomplishment and depersonalisation scales of the MBI, or on the interaction anxiety measure of VOTE. Therefore, the effects, although important, are not solely explained by the increases in activities. Pragmatic health services research needs to consider such issues in terms of interpreting significant (or non-significant) effects. A similar issue is reported in *Chapter 5* when we discuss that even

in the absence of any intervention, the view of the services deteriorated over time. However, these staff findings may also be a feature of a clustered trial; with relatively few numbers of clusters and a large variability in the situation of each cluster it is always possible for results to be shifted by extreme conditions on a small number of clusters. In other words, a few events or circumstances negative enough on two wards would be enough to push the results into significance. One of our future aims is to interrogate the data to see if there was a consistent perception on all wards.

Costs and cost-effectiveness

The intervention required investment in training and, when apportioned over the wards and service users, this amounted to £10 per patient-week. The cost of meaningful service contacts and activities increased (non-significantly) by £12 per patient-week following the intervention (including the intervention cost). This does not reflect an actual increase in expenditure but rather a relative increase in resources from a fixed amount. For example, if the actual cost per patient-week of an inpatient stay is £2100 (7 days at £300 per day), then £12 per week extra represents a shift of 0.5% towards per direct patient care. The cost-effectiveness acceptability curve shows that an improvement in the VOICE would have to be valued very highly in order for the intervention to be cost-effective. It was unclear whether or not the intervention could be regarded as cost-effective using VOTE.

Strengths and limitations of this study

A major strength of this study was the use of the stakeholder perspective in the main outcomes from the service users and staff members. No study to date, to our knowledge, has had such a focus. Although entirely well intentioned and very much welcomed, many of the strategies and programmes aimed at improving inpatient services are developed by senior management or bodies such as the Royal College of Psychiatrists.[161] We involved the service users and nurses using the wards on a day-to-day basis and ensured this programme delivered what mattered to them.

There were a number of reasons behind our decision to use a stepped-wedge design for this investigation. First, as the intervention was at the ward level, a cluster-based design was clearly required in this context. Second, a stepped-wedge design allows nursing staff on all wards to receive the training. This is in distinction to a regular cluster randomised trial in which only half of the clusters would be allocated to the intervention arm. Third, it was not possible for us to train all wards at the same time owing to limits in the number of researchers and psychologists employed. By staggering the timing of the intervention, we could reach more wards with the same number of researchers.

The economic evaluation was conducted from the perspective of the service user with their perception of meaningful contacts/activities being the driver of costs. Service user data relied on self-report and this may have some limitations in terms of recall accuracy. However, we have assumed that if the contact was meaningful then it would be recalled.

Research implications

Future research needs to use the tools of implementation science to uncover the factors that might predict successful implementation. Although we were able to show that activities had increased and we have anecdotal evidence of where it proved to be relatively successful, we also know that there was much variance unaccounted for. We will be investigating some of these exploratory hypotheses on our current data set to best suggest the remit for future studies. A better understanding is needed of the strategies for knowledge integration, especially in high throughput inpatient settings. We also plan to explore differences between wards with special reference to their culture and those differences produced by boroughs and trusts. We will also be carrying out further analyses investigating the impact of the intervention on clinical indicators such as PANSS scores, NOSIE and GAF. The results from the economic evaluation indicate the shift that is likely towards more direct care provision and this approach should be developed in further studies.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Clinical implications

One of the most common complaints about inpatient services is the extreme boredom and lack of therapeutic activities occurring on the wards (e.g. Walsh and Boyle[34] and CQC[170]). We successfully increased the number of activities and received positive feedback from those participating in the study, demonstrating that, with effort and dedication, it is possible to increase the number of activities. This was carried out with little increase in the costs and showed a transfer of costs to more patient valued care. Although there was variation in terms of implementation success, it should be possible to adopt our design to increase the skill and confidence of nursing staff, which has often been found to be perceived as poor.[29] This alone would be a positive outcome, not well captured in our current study but revealed in our qualitative analysis from the focus groups. Our baseline analyses also showed that symptom severity did not have an impact on the uptake of activities and participation in activities was related to more positive perceptions of the wards, demonstrating that they are a much valued component of inpatient services regardless of illness severity. This belies the belief that the acutely ill cannot take part in meaningful activities and also supports the view that more therapeutic activities could be of value and are appreciated.

Conclusions

This programme of research had some positive and some surprising results. Service users' perceptions of their care were moving in the right direction and, for those who were admitted compulsorily, there was a significant effect. Nurses appear to be working in such a pressured system that their perceptions were mixed in spite of the positive views of our project discussed in the focus group. The number of activities taking place and the number of violent incidents also moved in the right direction. A much larger study of this kind is needed in order to detect statistically significant effects, as power was low even after we increased the number of wards.

Chapter 5 Work package 4: BETTER PATHWAYS

Introduction

There has been a steady reduction in the number of acute psychiatric admission beds in many countries around the world. Since the 1960s it has become clear that with more effective pharmacological treatments, community-based alternatives to hospitalisation were possible and less costly, as well as often being preferred by service users and their families. Despite this shift towards community-based care, acute inpatient psychiatric services remain an indispensable element within an integrated mental health system. Indeed, acute wards consume the greatest proportion of the mental health budget and employ the greatest number of staff in the UK.[6] However, this high investment does not always translate into better services. Average length of stay is higher than the Organisation for Economic Cooperation and Development average (21 days for mood disorder to 38 for psychosis[171]) and many inpatient units have significant problems with the provision of treatment and unacceptably high bed occupancy levels, which are often well over 85% and sometimes over 100%.[172] Furthermore, a significant association has been shown between ward crowding and incidents of physical and verbal aggression.[173]

The traditional model of inpatient care in the UK involves service users being assigned to wards based on their geographical residence and bed availability, with no specific proposed limit on the duration of stay. The resulting case mix of acutely ill service users with others further along the pathway to recovery contributes to inefficiencies and delayed discharges. Pressures from managing newly admitted acutely ill service users also contributes to delayed discharge of those who need sustained therapeutic intervention, as the ward staff are spending time on bed management tasks as opposed to engaging with service users in potentially therapeutic communication or activities.[117] Previous studies have also shown that 27.2% of inpatients were judged by ward staff as not requiring hospitalisation, with the figure increasing to 61% for individuals with stays of > 6 months.[31] The main reason discharge was delayed in this study was the lack of suitable community accommodation or specialised rehabilitation services. Paradoxically, this system probably also contributes to the 'revolving door', as some service users (possibly those who have a home to go to and fewer social needs) experience repeated cycles of brief admissions with no resolution to their underlying problems.[174]

In this context, there have been many attempts to improve the efficiency of acute inpatient care and the care pathways that lead to it. Almost 40 years ago, Hirsch *et al.*[175] showed that service users who were randomly allocated to a specially designed short admission programme improved as much as the comparison group, despite a 33% reduction in the length of their hospital stay. Subsequently, crisis resolution and HTTs have been developed in an attempt to prevent hospitalisation and to facilitate earlier discharge for cases in which admission was unavoidable. The results have been mixed with little reduction in inpatient bed use.[176,177]

In an attempt to relieve the overoccupancy that leads to poor service user (and nurse) experience, a novel 'triage' ward system has been developed. In a triage system, all acute admissions are directed to one inpatient facility that provides a brief period of assessment and rapid stabilisation of the acute episode. The triage ward has a capped length of stay (typically 7–10 days). Service users who are assessed as needing a longer period of inpatient treatment are subsequently transferred to other inpatient settings that offer more rehabilitative care. Early experience of this model has claimed considerable benefit both clinically and in terms of greater efficiency. In a London example, 406 service users were admitted to the ward over a 6-month period with 170 (42%) being discharged home within a week. The mean occupancy rate on the triage ward during that period was 70% of capacity, which even allowed for the temporary admission of service users from neighbouring catchment areas where no inpatient beds were available.[178]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Although the primary aim of the triage system is to improve efficiency, the separation of the immediate management of the acute episode from longer-term inpatient care may also offer some advantages in terms of the experience of the service users and nurses. Rapid crisis resolution and discharge home are likely to be preferred by most and, for those with longer stays, a quieter environment that can offer a 'milieu' approach might also be conducive to better service user and staff experience. On the other hand, the triage team will be under pressure to make fast discharge decisions and to deal with high levels of acute symptoms, all of which might lead to higher levels of staff burnout. Furthermore, staff in the longer-stay locality wards may be faced with the challenges of a service user group with longer-term, more challenging behaviours and consequent delays in discharge, all of which can be demoralising. Service users may dislike rapid moves between different parts of the service and having to form relationships with yet another clinical team. User and staff views of the triage system are therefore essential to assess benefits and potential disadvantages of this service model in addition to the hard data on duration of care and costs. Despite their appeal and gaining popularity and proliferation, 'triage' services have not been subject to robust investigation. It is also unclear whether or not, when compared with a system without triage, (1) the claimed advantage of the model is sustained in the face of wider pressures on efficiency, (2) it achieves the hoped for cost benefit or (3) the savings from those service users discharged within the 7- to 10-day target are offset by even longer stays in those service users who are transferred to less acute settings.

This WP investigated whether or not the triage system contributes to reduced lengths of hospital admission overall and investigated the impact that the triage system has on the experience of service users and staff. We were also interested in whether or not one system is associated with a frequently reported benefit to service users – greater service user/staff interaction – and what this may cost.

Method

The setting

The study was carried out in one large mental health organisation responsible for all mental health services. The defined geography covered by this health-care trust comprises four catchment areas, each of which has its own inpatient hospital provision prioritised for catchment area residents but service users from any area, including other areas in the UK, may be admitted. This WP focused on two of these areas, one of which operated a triage service.

Triage service

Triage services are located in an area of London with an adult population of approximately 275,000 and an average index of deprivation score of 30.97 (ranking 31st most deprived out of the 326 boroughs in England).[163] The inpatient resource comprises 79 beds (four mixed-gender wards). There is also a PICU ward with 10 beds. The triage ward accepts all inpatient admissions and service users remain for a maximum of 7–10 days. All service users are expected to have senior medical input within 24 hours of admission and community mental health services are engaged in this process. The intensive assessment process allows for immediate discharge planning on the day of admission. Stabilisation takes place over 3 days and the HTT that is responsible for care in the community is consulted at days 3 and 4 and a decision to admit to longer-term care or home treatment is made with a resulting discharge or transfer by day 7. These assessments are carried out by health-care professionals with the ability to discharge or define the inpatient service need. There are three other longer-term treatment wards in this service system and which ward service users are admitted to depends on geography. This admission model is based on a holistic approach aiming for more efficient and continuous care, with the service user's Community Mental Health Teams (CMHTs) and HTTs significantly involved in both the admission and discharge processes. In the whole period of acute care, the HTT acts as an extension of the inpatient services providing high levels of support immediately after discharge. All four wards are mixed gender and had an average of 18 beds. All the wards are housed in a single building.

Traditional service

The comparison routine service is an area well matched for population size and deprivation (300,000 residents) with an average index of deprivation score of 31.24 (ranking 29th out of 326 boroughs[163]). Beds are provided in five wards (one of which closed during the year), with 18 beds each, which accept acute admissions. There is also a PICU with 10 beds. Service users are admitted to a ward with an available bed for their whole stay. As in the triage system, there is involvement of the service user's CMHT and a separate HTT that liaises with each of the wards to facilitate discharge. Two wards were mixed gender, one ward was female and two were male. One of the wards is a specialist early intervention unit for those having a first episode of psychosis and had a younger population than the other wards. Wards had an average of 18 beds. Initially wards were on three sites, but during the course of the study consolidated to a single site.

Two studies were carried out. One to explore the impact of a triage system on overall lengths of stay and the second partially nested within this: a survey of the views of service users and staff. These studies are reported separately.

Study 1: is a triage system more efficient than traditional care?[179]

Design

Data on all admissions and discharges were extracted from the comprehensive electronic patient record using the Clinical Records Interactive Search system. This has regulated access to anonymised information extracted from the trust electronic clinical records system which is updated daily. Data were extracted for dates of admission and discharge from all the wards in the two study areas and pathways on discharge, including any input from HTTs. Demographic and diagnostic characteristics of each admitted service user were also extracted. Two data sets were created to contain:

1. All admissions to the two systems between 1 January and 31 December 2009 for analyses of length of stay and costs. There were no admissions with lengths of stay that exceeded the date of data collection and so no censoring was necessary.
2. All re-admissions in the year following discharge from their first admission (i.e. index admission) in 2009. This allowed for a fair comparison by limiting the same period of time for re-admission to occur, regardless of the index admission date or the length of stay. There were no observation periods that exceeded the date of data collection.

Analyses

First, in order to determine whether or not there were differences between the inpatient systems we needed to identify whether or not the systems admitted different service user groups. This was investigated with *t*-test and chi-squared tests. Then we investigated whether or not the triage system discharged a greater proportion of patients back into the community within 10 days compared with routine care, that is, whether or not the triage did act as a fast throughput ward and, therefore, is different from the routine care system.

Second, we carried out an analysis of the length of stay for each admission as (1) length of inpatient care, the period from the night of admission to the day of discharge from the hospital, and (2) length of acute care, which includes any involvement of the HTT as a step down from hospital. As these analyses are by admission and some patients had multiple admissions during the period, all models accounted for non-independence of observations using random-effects linear regression including a random intercept at the patient level. Owing to the skewed distribution of length of stay, we used bootstrapping to estimate SEs and CIs. In addition, to adjust for any confounders as a result of differences in service user characteristics between the two areas, we constructed multivariable models using the same techniques and adjusted for service user characteristics that have been shown to be important in predicting length of stay in other studies.[180,181]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Third, we carried out an analysis of re-admissions to hospital. To estimate difference in odds of a re-admission (readmitted or not within 1 year) between the care systems we used a logistic regression model and added potential confounding variables using a stepwise selection technique with a threshold of $p < 0.20$. Categorical variables were tested using likelihood ratio tests.

Fourth, in an analysis of the costs of the two systems, we used the same statistical models and adjusted for the same confounders as for the length of stay analyses. We extracted data on the number of HTT visits that each service user received and multiplied the number of inpatient days and the number of HTT visits by unit cost figures from NHS reference costs[182] (£304), face-to-face contact with a crisis resolution HTT (£179)[182] (see *Appendix 6*). As it is a fairly common practice that some service users are granted periods of trial leave prior to final discharge, sensitivity analyses were conducted that repeated the multivariable analyses of length of stay and costs, but defined length of stay as the number of days spent physically present in hospital during an admission ('bed-days'), that is, excluding any leave periods before official discharge. The cost sensitivity analysis was performed under the assumption that inpatient costs while the service user was on leave were negligible and, thus, the analysis applied the same cost estimates for inpatient care (£304). Percentile-based CIs are reported for all length of stay and cost analyses.

All analyses were carried out using Stata v12.1.

Study 2: service user and staff perceptions of care

These data were collected four times at 6-month intervals over an 18-month period (largely coterminous with the comprehensive admission census described above).

This series of cross-sectional samples and their timing was deliberately chosen because of concern that a single cross sectional assessment would be open to the effects of seasonal changes and rare events, such as unexpected organisational changes that could not be controlled for in the analysis.

We carried out two sets of comparisons. In the first we compared staff and service user perceptions in the triage system (triage ward plus remaining wards) with those in routine care. In the second set of comparisons, we investigated specific differences between types of ward.

Inclusion criteria

Service users were included if they could communicate in English, if they provided informed consent and if they had been on the ward for a minimum of 7 days. For triage participants, this minimum was 3 days so as to minimise the loss of those who were rapidly discharged to the community. There were no diagnostic exclusion criteria. We interviewed service users only once in the study. This excludes 10% of individuals who agreed to take part but were in hospital at more than one time period and tended to be those who had long admissions rather than repeat admissions.

Permanent nursing staff of any grade who were working on the wards at the time of recruitment were eligible to be included in the study. Temporary staff were also eligible if they had completed seven shifts in the previous month (to ensure that they were sufficiently familiar with the ward environment).

In order to have representative views, the aim was to recruit at least 50% of staff and service users who were on the ward at the times of data collection.

Ethics approval for the study was granted by Bexley and Greenwich Research Ethics Committee (reference 07/H0809/49).

Measures

Service user measures

1. VOICE:[84] this was developed and validated as part of WP1. VOICE is easy to understand and complete and, therefore, is suitable for use by service users while in hospital. The 19-item measure has good validity and internal and test–retest reliability. A high score indicates a more negative perception.
2. SSS-RES:[78] this is a derivative measure adapted from the SSS-30[79] designed to evaluate residential services for people with serious mental illness. The original SSS-30 has been used in a variety of settings and demonstrates sound psychometric properties.[164] Higher scores indicate less satisfaction with services.
3. Demographic and clinical information: for each service user participant we recorded age, gender, marital status, ethnicity, education, socioeconomic status, accommodation and employment status, diagnoses, admission date, legal status, length of stay, time on ward, diagnosis history and medication history. For nurse participants we recorded age, gender, ethnicity, level of employment band, length of employment and education, and employment in previous wards/trusts.
4. GAF:[147] this is a 100-point rating scale of a service user's overall level functioning that is based on observation, interview and medical records. Scores can range from 0 (lowest functioning) to 100 (highest functioning).
5. NOSIE:[166] a 12-item scale designed for nurses largely focused on the assessment of socially unacceptable/unusual behaviour. Nurses were asked to rate participants on items such as 'sits, unless directed into activity'.

Staff measures

1. VOTE:[85] this is a 20-item measure that captures staff perceptions of the daily pressures of working in acute inpatient mental health wards. It was developed and validated in WP1. Each item is answered by choosing from a six-point Likert scale. All item scores are added together to form a total score. This score represents each staff participants' view and high scores are negative. Test–retest reliability and the internal consistency of the measure and subscales were good.
2. MBI:[49] this can be totalled to produce a score that measures work-related 'burnout' over 22 items, answerable by a six-point Likert scale. High scores indicate higher levels of burnout. The MBI comprises three individual subscales: emotional exhaustion (high scores mean high burnout), depersonalisation (high scores indicate high burnout) and personal accomplishment (low scores indicate high burnout). It has good psychometric properties and is widely used.
3. IWS:[81] this is a 44-item scale that can be totalled to produce a score that measures health professionals' levels of job satisfaction. It is answerable by a seven-point Likert scale.

Activities and costs

CITRINE:[142] this was developed and validated as part of WP2. This instrument records activity data on psychiatric inpatient wards including ward activities and time spent with health professionals that is perceived to be meaningful by the service user. In combination with data on the unit cost of staff time, this can be used to calculate the cost of such perceived contacts.[136] For more detailed explanation of the costing approach, see *Chapter 3*.

Analyses

There were four objectives:

1. To investigate whether or not service users' perceptions of the therapeutic environment (VOICE) and service users' satisfaction with services (SSS-RES) differed between the two systems.
2. To investigate whether or not staff perceptions of the therapeutic environment (VOTE), staff burnout (MBI) and work satisfaction (IWS) differed between the two systems.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

3. To establish whether or not the perceived amount of care received and its cost (CITRINE) differed between the routine and triage systems and to establish the cost-effectiveness of the triage system according to the perception of service users.
4. To establish whether or not our sample was representative of all the service users admitted to inpatient services in the two catchment areas.

For all of these objectives we carried out two sets of analyses. The first identified the key question: does the triage system as a whole produce any advantages or disadvantages? The second set explored three groups of wards: the triage, the remaining wards in the triage system and the routine system wards.

The service user outcomes and the cost of service use were modelled using linear regression in models that included phase of sampling as a covariate. Phase included as a trend was preferred if a likelihood ratio test between the two models provided no evidence that the categorical variable provided a better fit for the data, and phase by system interactions were tested in the same way. When interactions were significant, they are presented as the effect on phase compared with the baseline (reference) phase for each system. Sensitivity analyses were performed using mixed-effects regression models to take into account the clustering of the service users by sampling ward. In the cost regression, Eicker–Huber–White SEs were calculated to allow for non-normality of residuals. In addition, regression estimates were weighted by the ward-specific inverse probability weight (i.e. the inverse of the proportion of service users admitted to each ward that were contained in the sample). In the adjusted cost analyses, a number of potential confounders were controlled for. The choice of these variables was based on a previous study by Sabes-Figuera *et al.*[142]

For staff measures, when some staff members were interviewed more than once, a random-effects regression was fitted. Phase, and phase by system, covariates were tested using the same methods as in the service user models. Ward was included as an additional random effect in sensitivity analysis. The different exclusion criteria per system (length of stay < 7 days, length of stay < 3 days in the triage system), as defined in our protocol, resulted in the inclusion of seven participants (3%) in the triage system who would have been excluded from the routine system. We performed each model excluding these seven participants to test the influence on the findings and found that none of the results changed.

Exploratory analyses included a comparison between three categories of ward: triage ward, non-triage wards in the triage system and wards in routine care. This new three-level categorisation of system provided us with the specific contrasts of interest: (1) the triage ward versus the remaining non-triage wards in the triage system and (2) the contrast between routine system wards and the non-triage wards in the triage system. In the service user models, these contrasts were estimated with and without indicators of service user functioning on the ward to adjust for any differences between the systems.

To explore the relationship of the differences in costs with the differences in benefits between the two systems, we constructed two cost-effectiveness planes using bootstrap techniques. The first plot showed the joint distribution of mean costs and VOICE differences using the patient-level data. To obtain the second plot, we simulated the distribution of mean cost and VOTE differences using ward-level averages. In addition, we calculated cost-effective acceptability curves to show the probability that the net benefit of the triage system was higher than the routine care system at different levels of willingness to pay to improve VOICE and VOTE scores, respectively.

Results

Study 1: is the triage system more efficient than traditional care?

There were a total of 935 admissions to the triage system and 899 to routine care over our census year beginning on 1 January 2009. There were few differences and none of the differences between the characteristics of people admitted in the two systems reached statistical significance (*Table 30*). Psychosis accounted for the largest diagnostic group in both systems and around 40% were involuntary admissions.

TABLE 30 Description of admission characteristics to the routine and triage systems

Characteristics	Routine care system admissions, *n* = 899; individuals, *n* = 737	Triage system admissions, *n* = 935; individuals, *n* = 757
Demographic and clinical characteristics		
Age (years), mean (SD)	39 (12)	39.2 (12.7)
Gender: male	506 (56)	502 (54)
Ethnicity: white	385 (43)	476 (51)
Missing	13 (1)	22 (2)
Primary diagnosis		
Drug-related disorders	74 (8)	107 (11)
Psychosis	451 (50)	411 (44)
Bipolar	128 (14)	110 (12)
Depression	90 (10)	88 (9)
Anxiety	46 (5)	85 (9)
Personality disorders	22 (2)	59 (6)
Other/unspecified	49 (5)	41 (4)
Missing	39 (4)	34 (4)
Involuntary admission[a]	354 (39)	401 (43)
Admission characteristics		
Admission resulting in home treatment	236 (26)	301 (32)
Admission from outside catchment area	188 (21)	225 (24)
Rapid discharge (< 10 days)	259 (29)	403 (43)
Length of inpatient stay (days), mean (SD)	47 (84.7)	44.7 (75.2)
Length of inpatient stay (days), median (IQR)	22 (8–53)	16 (4–57)
Length of acute care (including HTT) (days), mean (SD)	51.8 (90.8)	54.5 (79.3)
Length of acute care (including HTT) (days), median (IQR)	26 (11–58)	27 (6–76)
Number of bed-days (inpatient stay excluding periods of leave from ward), mean (SD)	45.3 (85.7)	36.1 (57.2)
Number of bed-days (inpatient stay excluding periods of leave from ward), median (IQR)	20 (7–49)	15 (4–46)
Re-admission characteristics		
Any re-admission within 1 year of discharge from index admission	249 (35.6)	259 (35.4)

a Informal admissions included those individuals voluntarily admitted, those admitted under the Mental Health Act[183] Section 2/3 were formally detained.

Values are *n* (%) unless otherwise stated and figures presented are from the length of stay data set unless otherwise stated.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

For a fair comparison of the two systems, all admissions were analysed regardless of whether or not service users were resident in the catchment area. Admissions from individuals recorded to be homeless were assumed to be inside the catchment area. In both areas, a similar proportion of all admissions were for 'out of area' residents (routine care 21% and triage 24%).

Are there differences in length of stay between the two systems?

Although more admissions to the triage system had rapid discharge, both systems have service users with lengthy admissions. A univariable analysis showed that although the length of inpatient stay was, on average, lower in the triage system, the difference was not statistically significant (mean difference –3.87 days, 95% CI –12.20 to 4.47 days; $p = 0.364$). A multivariable analysis (*Table 31*) taking account of service user characteristics also failed to provide evidence of a difference between the systems (mean difference 0.53 days, 95% CI –9.77 to 7.21 days).

TABLE 31 Length of inpatient stay (WP4)

	Days	Bootstrap SE	*p*-value	95% CI (percentile based)
Service system				
Routine (in catchment)	Reference			
Triage (in catchment)	–0.53	4.37	0.90	–9.77 to 7.21
Interaction between system and catchment area				
Admission from outside catchment area (for routine care, reference)	–13.39	5.86	0.02	–24.20 to 1.25
Service system by catchment area (interaction, additional effect of triage)	10.02	7.83	0.20	–5.98 to 23.93
Diagnosis				
Schizophrenia	Reference			
Other psychotic disorders	–12.32	5.53	0.03	–24.18 to 2.50
Bipolar and related disorders	–4.76	8.96	0.60	–20.78 to 14.74
Depressive disorders	–9.34	5.82	0.11	–21.04 to 2.19
Neurotic and anxiety disorders	–19.58	10.61	0.07	–36.37 to 3.68
Personality disorders, etc.	–22.28	8.63	0.01	–39.54 to 6.14
Substance misuse	–27.02	4.9	< 0.001	–37.74 to 17.84
Other/unspecified	–2.17	9.54	0.82	–17.64 to 18.56
Demographic and clinical factors				
Age	0.06	0.16	0.70	–0.25 to 0.38
Previous number of admissions (log-scale)	9.36	3.33	0.005	3.94 to 16.59
Admission status				
Voluntary	Reference			
Compulsory assessment	1.67	2.66	0.53	–3.60 to 7.31
Involuntary treatment section	69.83	5.65	< 0.001	57.92 to 79.55
Fixed residence				
Yes	Reference			
No	69.54	9.65	< 0.001	51.35 to 89.76

Multivariable linear regression model including a random intercept to account for clustering effects by service user. Bootstrapped SEs and percentile-based CIs were estimated using 1000 repetitions.
Number of admissions, $n = 1722$; number of individuals, $n = 1367$; average number of admissions per person, $n = 1.3$. ICC = 0.49.

The use of HTT also differed between systems with a greater proportion of service users from triage followed up by the HTT (32%) at discharge than from the routine care system (26%, OR 1.50, 95% CI 1.12 to 2.01; $p = 0.006$). Repeating the multivariable analysis of length of stay but including the additional days spent with HTT showed that the mean length of acute care is actually somewhat longer in the triage system (mean difference 7.83 days, 95% CI –1.62 to 15.63 days; $p = 0.076$).

Figure 11 shows probability density distributions of length of inpatient stay with and without involvement of the HTT in both systems. It is apparent that more admissions to the Triage system resulted in a rapid discharge but both systems had a long 'tail' of admissions that were in excess of 60 days. The use of HTT differed. In the routine care system, HTT duration was largely focused around one to two weeks following discharge from hospital whereas in the Triage system service users generally had longer contact with the HTT. When the two components were considered together, as a period of acute care, there were no differences between the systems. Looking in detail, we noticed that for the acute care episodes of short duration, both systems used a similar percentage of days on home treatment to days in inpatient care. However, for the longer care episodes we noticed that the Triage system discharges individuals to HTT care for longer.

A sensitivity analysis that excludes the use of leave periods (where a service user is allowed home temporarily and his/her bed re-used) from the measure of length of stay, shows the average number of days spent on the ward was lower in the Triage system (mean difference: –8.36 days, 95% CI –16.54 to –1.51; $p = 0.030$).

About one in three of the admissions in each system resulted in a re-admission within 1 year of discharge. A multivariable model showed no significant difference in the probability of re-admission between the two systems (odds of re-admission: triage vs. routine = 1.05, 95% CI 0.77 to 1.44; $p = 0.758$) (*Table 32*). Results from the univariable analysis were (odds of re-admission: triage vs. routine = 0.90, 95% CI 0.80 to 1.22; $p = 0.925$).

Do costs differ between the two systems?

Similar to the analysis of length of stay, our exploratory cost analysis suggested that there were no statistically significant differences between the two systems. The mean cost of an acute care episode (inpatient length of stay and number of home treatment visits) was £15,233 in the Triage system compared with £15,476 in the routine system. Adjusting for potential confounders, the mean difference in the cost of acute care for service users treated within their catchment area was estimated to be higher in the triage system (cost difference £391, 95% CI –£2535 to £2748; $p = 0.77$).

Study 2: the experience of service users and staff

Four hundred and fifty-four service users were recruited, which is 61% of all those eligible to take part. They were mostly single men of non-white ethnicity, aged, on average, 40 years and who were unemployed. Most had experienced a previous admission with a median of 4 years since their first admission, and the most common diagnosis was psychosis (56%). Mean NOSIE and GAF scores were similar between systems. The mean NOSIE score of the sample was 16.0 and the mean GAF (level and symptom scores) were 43.3 and 42.2, respectively (*Table 33*).

Are service user participants representative of those using inpatient care?

We also utilised available anonymous data of the inpatient population on the same wards and over the same period. A comparison of the clinical and demographic profiles demonstrated that they were very similar (see *Table 33*). We excluded individuals with very short lengths of stay in the current sample (< 3 days) and so the length of stay sample will not mimic that reported for the total population but with that exclusion there were no major differences between the triage and routine care systems in admission duration.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

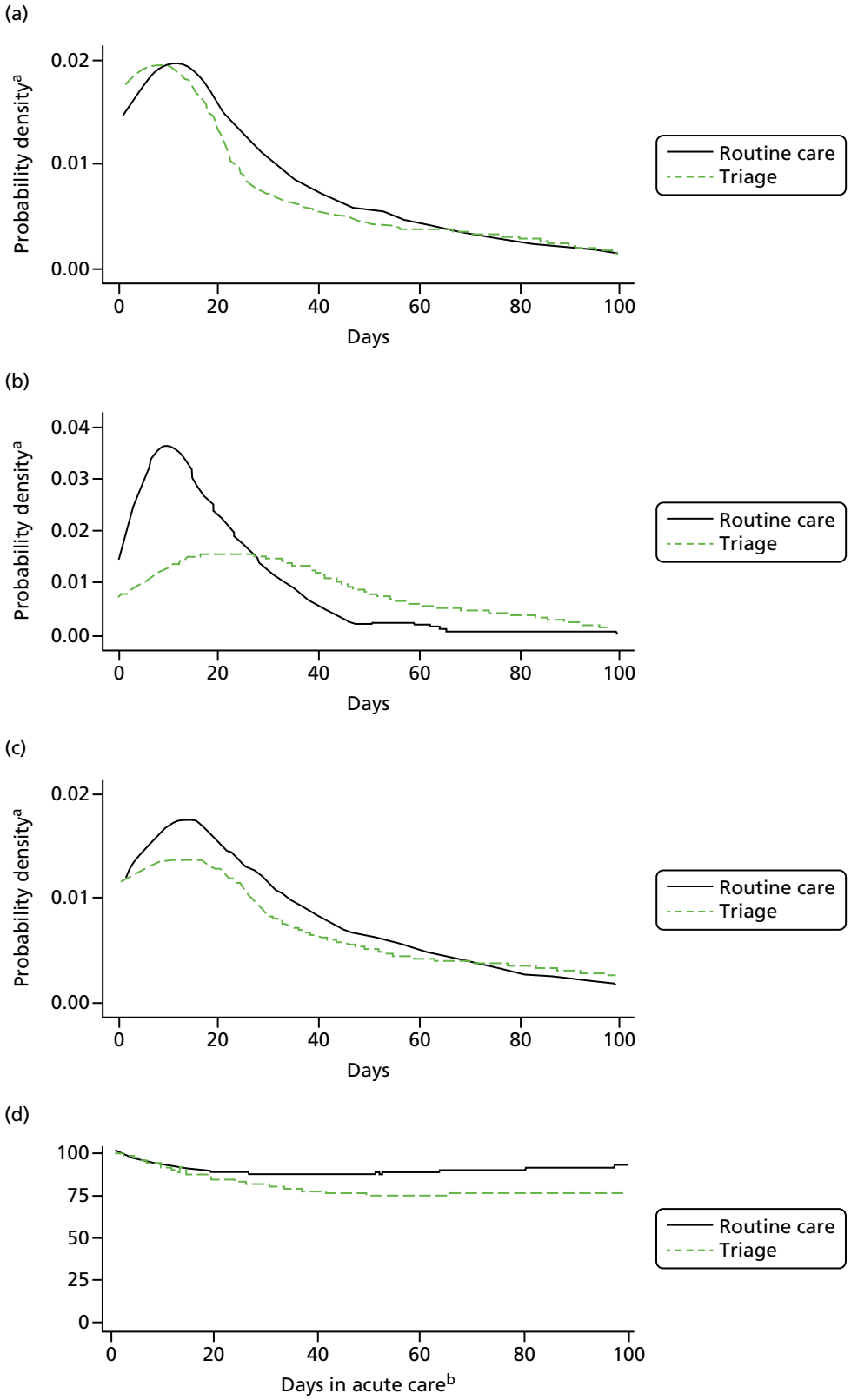


FIGURE 11 Length of care components (WP4). (a) Inpatient stay; (b) home treatment; (c) acute care; and (d) inpatient care/acute care. a, Probability density functions estimated using a kernel-smoothing function calculated on full sample, distribution censored at 100 days; and b, locally weighted scatterplot smoothing curves, censored at 100 days. Acute care = length of stay including any use of Home Treatment. Acute care admissions lasting ≤ 100 days accounted for 85% of the full sample. Reproduced from Efficacy of a triage system to reduce length of hospital stay. Williams P, Csipke E, Rose D, Koeser L, McCrone P, Tulloch AD, *et al. Brit J Psychiat* 2014;**204**:480–5.[179] Reproduced with permission of Royal College of Psychiatrists via PLSClear.

TABLE 32 Re-admission to inpatient care (WP4)

	Odds ratio	SE	*p*-value	95% CI
Service system				
Routine care system	Reference			
Triage system	1.05	0.17	0.758	0.77 to 1.44
Ethnicity				
White	Reference			
Black	1.2	0.21	0.291	0.85 to 1.70
Other	0.58	0.17	0.071	0.32 to 1.05
Employment				
Unemployed/retired/education	Reference			
Employed	0.72	0.18	0.187	0.43 to 1.18
Marital status				
Single	Reference			
Divorced/separated/widowed	0.51	0.12	0.005	0.32 to 0.82
Married/cohabiting	0.39	0.13	0.004	0.21 to 0.75
Diagnosis				
Schizophrenia	Reference			
Other psychotic disorders	1.3	0.29	0.238	0.84 to 2.03
Bipolar and related disorders	1.42	0.37	0.181	0.85 to 2.37
Depressive disorders	0.6	0.19	0.1	0.32 to 1.10
Neurotic and anxiety disorders	0.64	0.24	0.23	0.30 to 1.33
Personality disorders, etc.	2.34	1.03	0.054	0.99 to 5.55
Substance misuse	1.34	0.37	0.29	0.78 to 2.31
Other/unspecified	0.38	0.17	0.028	0.16 to 0.90
Any admissions prior to the first admission in 2009				
No	Reference			
Yes	2.7	0.43	< 0.001	1.97 to 3.69
Residential mobility				
No	Reference			
Yes	0.49	0.13	0.006	0.29 to 0.81

Multivariable logistic regression model shown was constructed using a stepwise variable selection procedure retaining all variables that had an association with re-admission ($p < 0.2$). Candidate variables that were tested and removed ($p > 0.2$) were residency outside catchment area, use of home treatment at the end of index admission, gender and age.
$n = 781$ individuals.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 33 Service user demographic characteristics by care system (WP4)

	Routine care system (*N* = 227)	Triage system (*N* = 227)	Inpatient population[a] (admissions = 2467 among 1571 patients)
Age (years), mean (SD)	39.4 (–13.6)	40.7 (–12.8)	39.5 (–12.7)
Gender: female, *n* (%)	99 (44)	99 (–44)	1037 (42)
Ethnicity: non-white, *n* (%)	134 (–59)	125 (–55)	1412 (–58)
Education, *n* (%)			
Primary education or less	11 (–5)	15 (–7)	
Secondary education	121 (–56)	104 (–48)	
Tertiary/further education	83 (–38)	93 (–43)	
Other general education	1 (–0.50)	6 (–3)	
Occupation, *n* (%)			
Paid or self-employed	37 (–17)	30 (–14)	129 (–9)
Marital status, *n* (%)			
Single/unmarried	150 (–67)	145 (–65)	1892 (–78)
Married/cohabiting	26 (–12)	28 (–13)	211 (–9)
Separated/divorced/widowed	47 (–21)	51 (–23)	321 (–13)
Socioeconomic status, *n* (%)			
Higher professionals (ONS category 1–4)	42 (–20)	44 (–20)	
Lower professionals (ONS category 5–8)	96 (–45)	111 (–51)	
Semi- and non-skilled workers (ONS category 9)	38 (–18)	43 (–20)	
Never worked and long-term unemployed (ONS category 10)	38 (–18)	19 (–9)	
Accommodation: owner occupied or renting	193 (–86)	187 (–84)	
Primary diagnosis, *n* (%)			
Drug-related disorders	14 (–6)	9 (–4)	142 (–6)
Psychosis	125 (–55)	124 (–57)	1431 (–59)
Bipolar	29 (–13)	30 (–14)	353 (–15)
Depression/anxiety	25 (–11)	21 (–10)	301 (–12)
Personality disorders	5 (–2)	10 (–5)	81 (–3)
Unspecified[b]	20 (–8)	10 (–5)	32 (–1)
Other	8 (–4)	14 (–6)	76 (–3)
Previous admission: yes, *n* (%)	165 (–73)	175 (–77)	–
Years since first admission, median (IQR)	3 (0–11)	5 (1–12)	–
NOSIE, mean (SD)	15.9 (8.3)	16 (7.5)	–
GAF level, mean (SD)	42.4 (12.1)	44.3 (13.3)	–
GAF symptoms, mean (SD)	41 (13.9)	43.4 (14.4)	–

ONS, Office for National Statistics.

a Data extracted from a case register of admissions to the two systems. Sample included contained the same wards, time periods and length of stay exclusion criteria.

b Many participants with an 'unspecified' primary diagnosis were young people with first episode psychosis.

Staff

A total of 484 observations were collected from 284 different staff members during the four phases of the study. This represents, on average, 57% of nursing staff who were eligible to take part at each time point. Most were women, of non-white ethnicity, aged 36 years on average. They had been employed for approximately 20 months on wards in the routine care system and 18 months in the triage system. Most staff were qualified nurses in both systems. It was not possible to collect data on those nurses employed on the wards who did not agree to participate.

Do experiences differ between the triage and routine care systems?

Service users

The detailed results for these analyses are shown in *Table 34*. No significant difference was found in perceptions of the ward between the two systems (VOICE scores adjusted mean difference –0.77, 95% CI –4.44 to 2.90; $p = 0.68$). However, this analysis demonstrated that perceptions of wards deteriorated over time. For every increase in phase (6-month interval), the mean VOICE score for the whole sample increased (coefficient 2.72, 95% CI 1.00 to 4.43; $p < 0.005$). This result remained significant after adjusting the model for clustering. The results remained stable in the two sensitivity analyses adjusting for the number of days spent on the ward until interview. The exploratory model that separated the triage ward from the remaining triage system wards also found no evidence of differences between the three types of ward.

TABLE 34 Staff demographic characteristics by care system (WP4)

	Routine care system ($N = 150$)	Triage system ($N = 134$)
Age (years),[a] mean (SD)	36.09 (10.22)	35.41 (9.14)
Gender: female, *n* (%)	85 (57)	87 (65)
Ethnicity: non-white, *n* (%)	85 (58)	89 (67)
Education group (highest achieved)		
Vocational	53 (40)	50 (42)
University degree (undergraduate)	60 (46)	59 (49)
University degree (postgraduate)	18 (14)	11 (9)
Length of employment on ward (months)	20.5 (1–300)	18 (1–207)
Band (highest achieved)		
2	10 (8)	6 (6)
3	36 (27)	22 (21)
4	2 (2)	0 (0)
5	53 (40)	51 (48)
6	24 (18)	23 (22)
7	8 (6)	5 (5)

a For staff with repeated measures, the mean of the time points responded was taken.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

No significant differences or associations were found in satisfaction between triage and routine care or between routine care and the remaining wards in the triage system (*Figure 12*). However, there was a trend towards lower satisfaction in the triage ward compared with the remaining locality wards (mean difference –6.91, 95% CI –14.66 to 0.84; $p = 0.08$). This remained stable in the two sensitivity analyses adjusting for the number of days spent on the ward until interview. A trend of deteriorating satisfaction with phase of the study was also observed ($p = 0.07$).

One potential explanation for the deteriorating service user perceptions is that we also found a trend of decreasing length of stay with phase. As the phase increased by one unit, the average total length of stay for participants decreased by 16.16 days (coefficient –16.16, 95% CI –26.84 to –5.49; $p = 003$). This finding remained in the sensitivity analysis which included clustering by ward (*Figure 13* and *Table 35*).

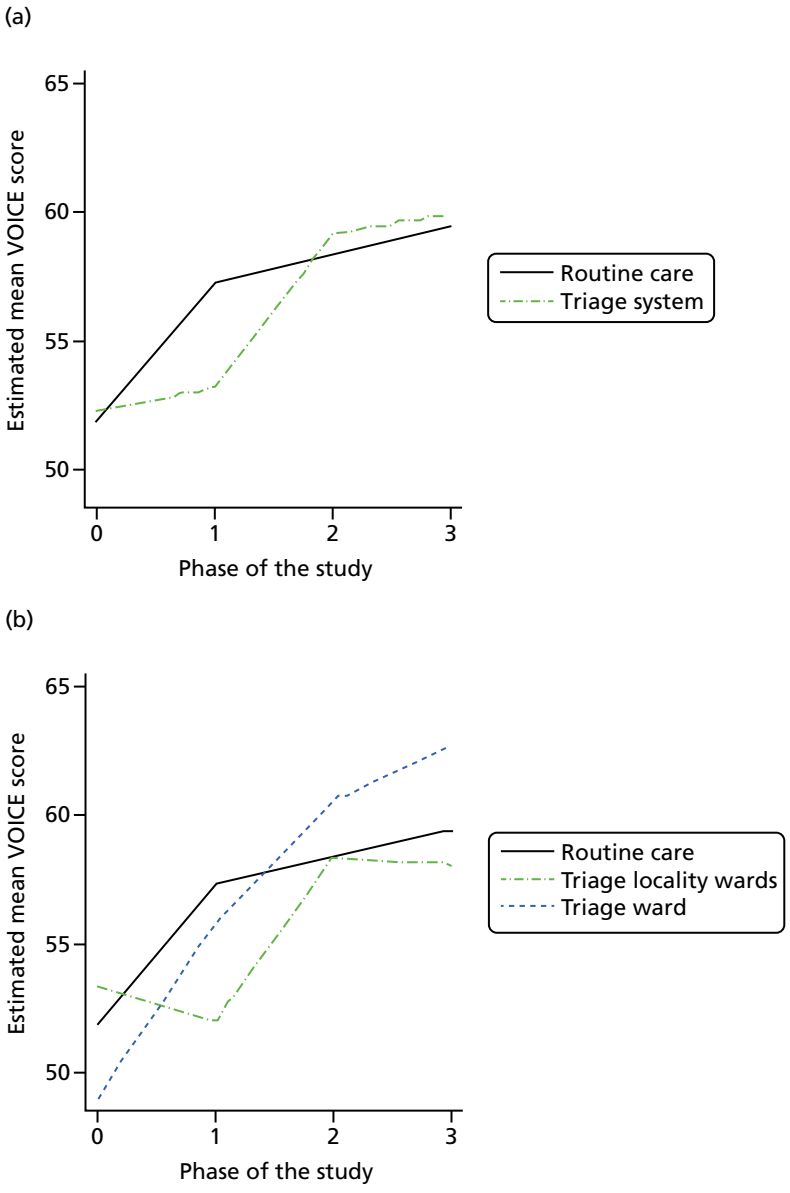


FIGURE 12 Service user VOICE for each system by phase (WP4) by (a) system; and (b) system and ward type.

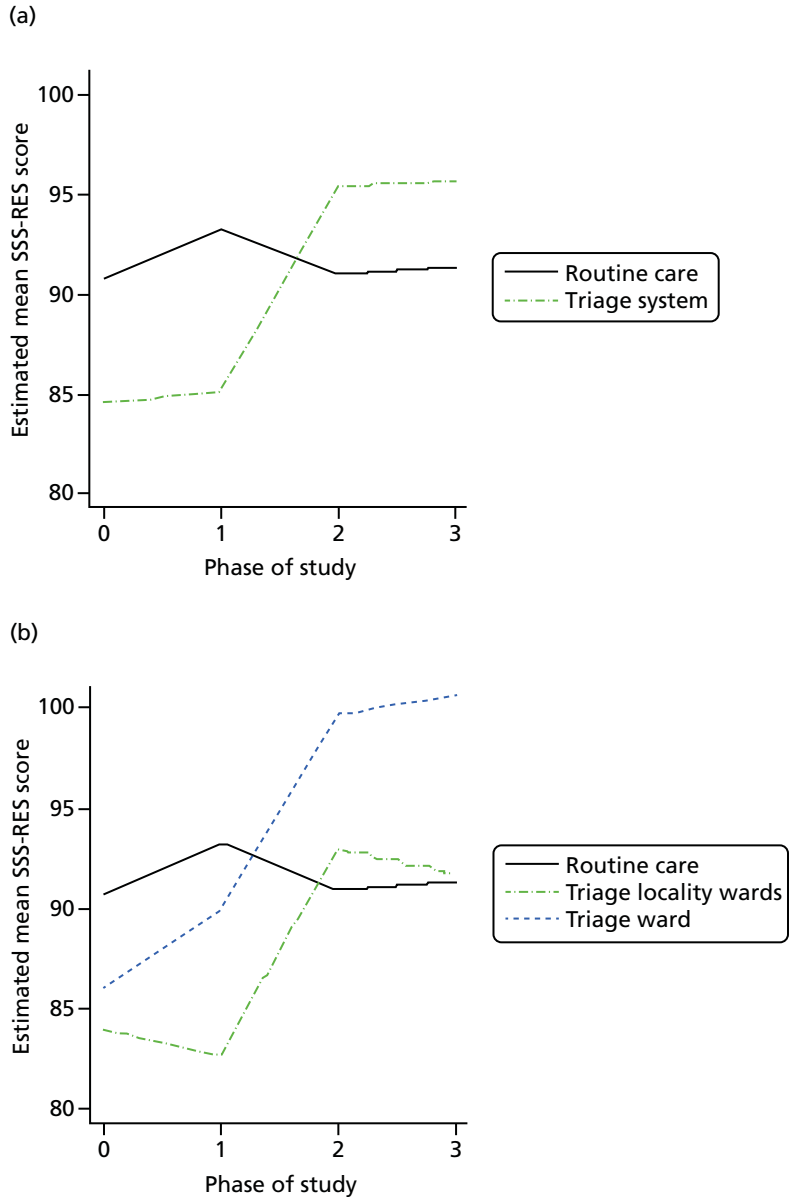


FIGURE 13 Service user SSS-RES for each system by phase (WP4) by (a) system; and (b) system and ward type.

Table 36 provides results of a number of exploratory analyses of the potential predictors presumed to be associated with service user experience (VOICE) from prior research. Worse service user experience (i.e. higher VOICE scores) was associated with higher skilled worker groups, increasing severity of symptoms and those who were admitted for compulsory treatment. A strong association with the number of contacts service users make with staff and better service user experience is associated with more staff contacts. The trend with phase and with severity of symptoms remained significant in this adjusted model.

Nursing staff

The same models were examined for staff but in these models we used time as a categorical variable. Interactions between system and phase were significant for VOTE [$\chi^2(3) = 11.03$; $p = 0.01$] and for MBI [$\chi^2(3) = 8.89$; $p = 0.03$] but not for the IWS model [$\chi^2(3) = 5.80$; $p = 0.12$].

There was no significant difference in staff experience at baseline between the two systems (adjusted mean difference: 1.68, 95% CI –2.05 to 5.41; $p = 0.38$) but staff experience in both systems changed over the period of the study and these changes differed between systems. In the routine care system, there was a deterioration in the VOTE measure at 6 months (in phase 1) (mean difference: 6.58, 95% CI 3.83 to 9.33;

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 35 Results for primary service user outcomes (WP4)

	Coefficient	SE	*p*-value	95% CI
VOICE model[a]				
Triage system	–0.77	1.87	0.681	–4.44 to 2.90
Phase	2.72	0.87	0.002[b]	1.00 to 4.43
Sensitivity analyses[a]				
Triage system (excluding seven participants with number of days until interview of < 7 days)	–0.54	1.88	0.774	–4.23 to 3.16
Triage system (adjusted for days until interview)	–0.78	1.87	0.678	–4.45 to 2.89
SSS-RES model[a]				
Triage system	–1.77	2.7	0.512	–7.07 to 3.53
Phase	2.29	1.24	0.066	–0.15 to 4.74
Sensitivity analyses[a]				
Triage system (excluding seven participants with number of days until interview of < 7 days)	–1.53	2.71	0.572	–6.85 to 3.79
Triage system (adjusted for days until interview)	–1.79	2.7	0.508	–7.09 to 3.51
Total length of stay[c]				
Triage system	–12.35	13.72	0.368	–39.25 to 14.54
Phase	–16.16	5.45	0.003[b]	–26.84 to 5.49

a Linear regression covarying for time assuming a linear relationship.
b $p < 0.05$ after performing a sensitivity analysis of the same model with robust SEs adjusted for clustering at ward level. ICC coefficient (ward) were 0.06, 0.07 and 0.06, respectively. Significance of all other results did not change.
c Linear regression covarying for phase assuming a linear relationship. SEs were adjusted using bootstrapping with 10,000 replications. Five different seeds were used and all estimates were consistent to 1 decimal place. Seven participants with < 7 days from admission until interview were removed.

Reproduced from Csipke *et al.*[184] © The Royal College of Psychiatrists 2016. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY) licence.

$p < 0.001$), but this was not seen in the triage system (mean difference 0.94; $p = 0.503$) and the difference between the two systems was significant (mean difference –5.64, 95% CI –9.53 to –1.75; $p = 0.005$). By the end of the study (phase 3), there was evidence of deterioration compared with baseline in both the routine care system (mean difference 4.80, 95% CI 1.30 to 8.30; $p = 0.007$) and the triage system (mean difference 4.78, 95% CI 1.93 to 7.63; $p = 0.001$) and no evidence that this was a differential change ($p = 0.994$). These results remained in a sensitivity analysis adjusting for clustering by ward (*Figure 14* and *Table 37*).

Similar results were seen for the other staff measures. Staff 'burnout' as measured by the MBI in the triage system appeared to be relatively stable throughout the study timeframe but an exploratory analysis separating the triage ward from the locality wards in that system showed that burnout was highest on the locality wards, even when compared with staff in the routine care system (mean difference 3.76, 95% CI –7.28 to –0.25; $p = 0.04$). Staff were also more dissatisfied in the triage system than in the routine care system (IWS mean difference 8.06, 95% CI 0.95 to 15.16; $p < 0.05$) and this increased over the study for both systems, with the most dissatisfaction in the final phase compared with baseline (mean difference 11.90, 95% CI 6.36 to 17.43; $p < 0.001$). The largest difference appeared to be between the routine care system and the triage locality wards (mean difference 9.03, 95% CI 1.19 to 16.87; $p = 0.03$) that had the lowest level of satisfaction. No evidence was found for a difference between the routine care system and the triage ward itself (mean difference 5.77, 95% CI –4.82 to 16.37; $p = 0.29$).

TABLE 36 VOICE exploratory model (WP4)

	Coefficient	SE	*p*-value	95% CI
System (triage vs. routine)	–1.59	1.8	0.379	–5.13 to 1.96
Phase	2.27	0.84	0.007[a]	0.63 to 3.92
Gender (female vs. male)	–3.79	1.8	0.036[a]	–7.33 to 0.25
Socioeconomic status				
Lower professionals vs. higher professionals	–1.52	2.37	0.523	–6.18 to 3.15
Semi- and non-skilled workers vs. higher professionals	–6.76	2.93	0.021[a]	–12.52 to 1.01
Never worked and long-term unemployed vs. higher professionals	–6.08	3.15	0.054	–12.28 to 0.12
Compulsory treatment orders				
1 month vs. informal	2.84	2.62	0.28	–2.32 to 7.99
≥ 6 months vs. informal	7.63	2.04	< 0.001[a]	3.62 to 11.64
Severity of symptoms (NOSIE)	0.28	0.12	0.021[a]	0.04 to 0.52
Number of contacts with service staff (in last 7 days from CITRINE)				
Simple effect	–2.35	0.45	< 0.001[a]	–3.23 to 1.47
Quadratic effect	0.07	0.02	0.001[a]	0.03 to 0.11

a $p < 0.05$ after performing a sensitivity analysis of the same model with robust SEs adjusted for clustering at ward level. ICC coefficient (ward) = 0.017. Significance of all other results did not change.
Linear regression of VOTE on sociodemographic and clinical variables selected using a stepwise selection procedure. Candidate variables that were tested in the model but removed owing to low ability to predict outcome were as follows: ethnicity, age, diagnosis, functioning (GAF level), number of previous admissions, number of days on ward, and number of recorded activities (in part 7 from CITRINE).

What services do people receive in the two systems, what are their costs and what is the cost-effectiveness?

Despite similar levels of nurse staffing, service users in the triage systems reported about half the number of contacts with nurses than those in the routine care system (*Table 38*). The cost of meaningful contacts with OT and activities was also found to be statistically significantly lower in the triage system. The total costs of staff contacts were estimated to be £41 lower in the triage system than the routine care system in the unadjusted analysis (95% CI 7 to 75), but this estimate was highly unstable over time and between wards. *Figure 15* shows the development of the costs of perceived service use across the study. There was a marked reduction in staff contacts in the routine care system; therefore, after adjusting for potential confounders, the triage system was estimated to have lower costs in the first two phases of the study only. On the other hand, at phase 2, costs were found to be statistically significantly higher in the triage system. Two units (an early intervention ward and one of the acute wards) in the routine care system had notably higher estimated mean costs than the other wards (£294 and £305 vs. £159). Therefore, the use of cluster robust SEs increased the width of CIs considerably, such that differences were no longer statistically significant in the unadjusted comparison. In the adjusted comparison the costs of service use in the triage system remained statistically significantly lower at phase 0 and higher at phase 2 than the routine care system.

Within the triage system, service users on the triage ward were significantly less likely to take part in activities and have contact with other care professionals. The overall cost of perceived service use was marginally lower in the triage ward than other wards in the triage system (mean difference –17, 95% CI –62 to 27), but this difference was not statistically significant and equally unstable across the different phases of the study as in the comparison of the two systems. The cost of meaningful contacts was statistically lower in the triage system at phase 3, but the mean estimate was not consistently lower across

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

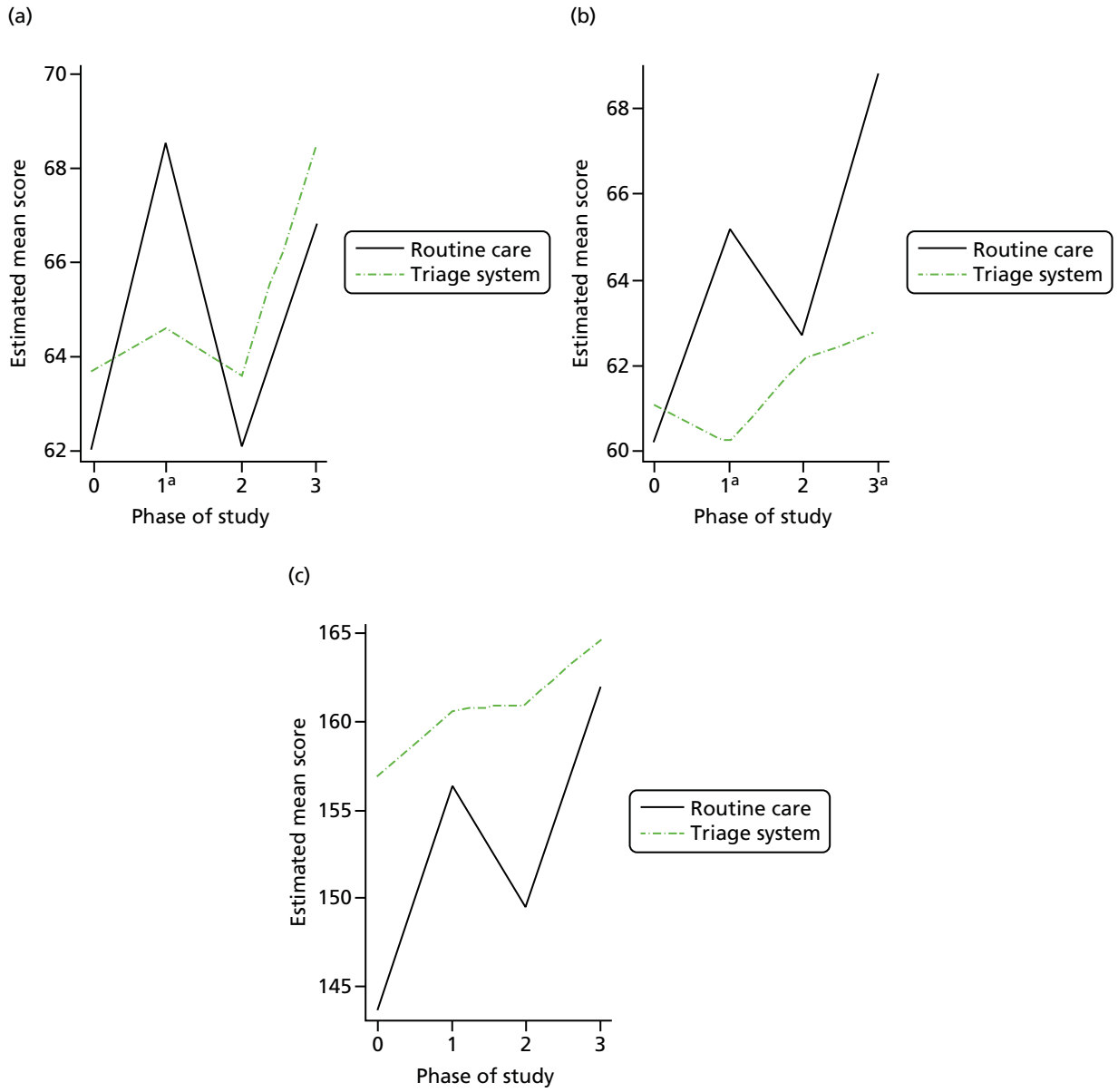


FIGURE 14 Staff outcomes for each system by phase (WP4). (a) VOTE; (b) MBI; and (c) IWS. a, Phases in which difference between systems was statistically significant.

the study. In summary, although there was some indication that service use patterns in the two systems might be different, the results of the service use analysis were not robust and these differences may be attributable to other factors. Similarly, our data suggested that there were no substantial differences between the overall cost of perceived care received in the triage ward and long-stay wards in the triage system.

The cost-effectiveness analysis also did not provide clear support for either of the two systems. Using VOICE, there was a high probability that the triage system was more cost-effective (*Figure 16*) at any willingness to pay and that it would lead to both a decrease in costs and improvement in service user satisfaction (*Figure 17*). However, the aforementioned lack of robustness in mean cost estimates over the time horizon of the study should be kept in mind. In addition, in the case of VOTE, the analysis suggested that the triage system would probably lead to a deterioration in staff perception (67%) (*Figure 18*). Therefore, only if little value was placed on this outcome measure and it was the only outcome measure would the triage system be cost-effective (*Figure 19*).

TABLE 37 Full results for the regression models of the staff primary outcomes (WP4)

	Coefficient	SE	*p*-value	95% CI
VOTE model[b]				
System				
Routine care system	Reference			
Triage system	1.68	1.90	0.38	–2.05 to 5.41
Interaction: system (routine) X phase				
Baseline (routine)	Reference			
Phase 1	6.58	1.40	< 0.001[a]	3.83 to 9.33
Phase 2	0.25	1.60	0.99	–3.11 to 3.16
Phase 3	4.80	1.79	0.007[a]	1.30 to 8.30
Interaction: system (triage) X phase				
Baseline (triage)	Reference			
Phase 1	0.94	1.41	0.50	–1.81 to 3.70
Phase 2	–0.13	1.40	0.92	–2.90 to 2.63
Phase 3	4.78	1.46	0.001	1.93 to 7.63
MBI model[b]				
System				
Routine care system	Reference			
Triage system	0.86	2.35	0.72	–3.75 to 5.47
Interaction: system (routine) X phase				
Baseline (routine)	Reference			
Phase 1	5.02	1.82	0.006[a]	1.44 to 8.60
Phase 2	2.47	2.01	0.22	–1.47 to 6.40
Phase 3	8.61	2.24	< 0.001[a]	4.21 to 13.01
Interaction: system (triage) X phase				
Baseline (triage)				
Phase 1	–0.87	1.90	0.65	–4.60 to 2.86
Phase 2	1.06	1.87	0.57	–2.60 to 4.72
Phase 3	1.76	1.89	0.35	–1.95 to 5.46
IWS model[b]				
System				
Routine care system	Reference			
Triage system	8.06	3.63	0.03[a]	0.95 to 15.16
Phase				
Baseline	Reference			
Month 6	8.13	2.50	0.001[a]	3.23 to 13.05
Month 12	5.06	2.66	0.06	–0.16 to 10.27
Month 18	11.90	2.82	< 0.001[a]	6.36 to 17.43

a $p < 0.05$ after performing a sensitivity analysis of the same model with an additional level of clustering at the ward level. Significance of all other results did not change.

b Mixed-model linear regression including a random effect for individuals and covarying for phase (categorical, coefficients use the baseline as a comparison). A system by phase interaction was tested in each model and results are presented where this interaction was shown to be significant from a likelihood ratio test.

ICC coefficients for the models were 0.71, 0.64 and 0.72, respectively.

Reproduced from Csipke *et al.*[184] © The Royal College of Psychiatrists 2016. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY) licence.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 38 Service use and costs (WP4)

	Routine care system ($N = 223$)			Triage system ($N = 226$)					
				Non-triage wards ($n = 147$)			Triage ward ($n = 79$)		
Type of contact	% accessing service	Mean (SD) number of contacts[a]	Mean (SD) costs	% accessing service	Mean (SD) number of contacts[a]	Mean (SD) costs	% accessing service	Mean (SD) number of contacts[a]	Mean (SD) costs
Nurse	64	6 (16)	31 (66)[b]	60	3 (3)	16 (27)	65	3 (2)	18 (23)
Any doctor	78	2 (1)	123 (135)	86	2 (1)	112 (101)	84	3 (2)	140 (157)
Care co-ordinator	19	2 (2)	9 (29)	34	1 (1)	15 (29)	24	1 (1)	12 (32)
OT	26	2 (2)	7 (22)[b]	2	1 (1)	1 (5)	1	1 (n/a)	0 (0)
Other care professional	25	2 (2)	17 (45)	39	2 (1)	22 (54)[c]	9	1 (0)	4 (19)
Activities	71	4 (4)	25 (33)[b]	71	2 (2)	13 (21)[c]	37	2 (1)	3 (8)
Total			212 (198)			179 (148)			160 (171)

a Among users of the service.
b Statistically significantly higher cost in routine care system than triage system at the 5% level.
c Statistically significantly higher cost in non-triage wards than triage wards in triage system at the 5% level.

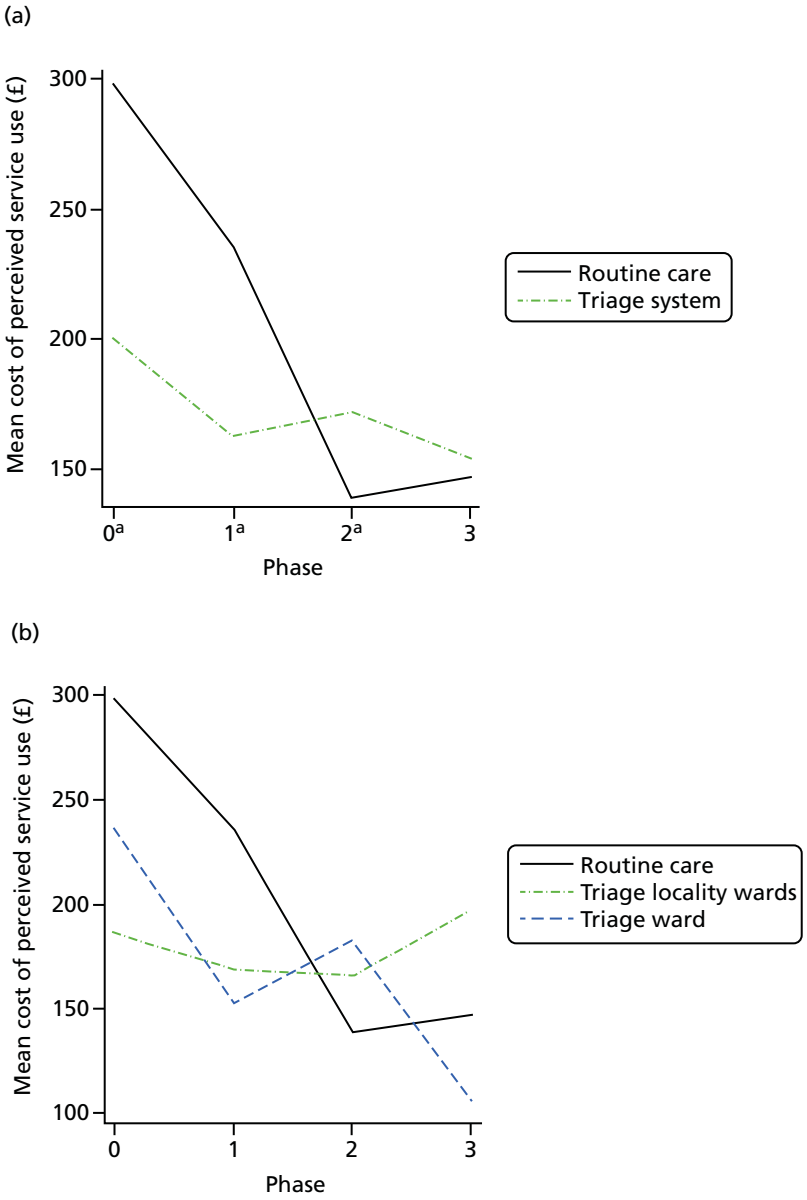


FIGURE 15 Cost of perceived service use (WP4). a, Phases in which difference between systems is statistically significant after adjusting for confounders.

Discussion

Strengths and limitations

Triage systems are one attempt to improve the efficiency and efficacy of acute inpatient psychiatric services and the current study is, to our knowledge, the first attempt at a rigorous evaluation of the approach. Our data included a comprehensive account of all admissions and discharges across two similar geographical catchment areas and linked this with a contemporaneous assessment of service user and staff perceptions of the quality of inpatient care in both systems. A limitation of our study is that it describes just one triage system in one inner city location provided by one NHS trust at one period in history. With these limitations in mind, the results of our investigation suggest that although a triage ward does achieve a rapid turnover of acute cases, this has little effect on the overall length of stay across the system and, in our data at least, a triage system cannot be said to provide a clear advantage over traditional care. There are three possible explanations for our observations. First, and most plausible, is the continuing pressure to improve efficiency in routine care that has made traditional catchment area services more like 'triage' than they might have

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

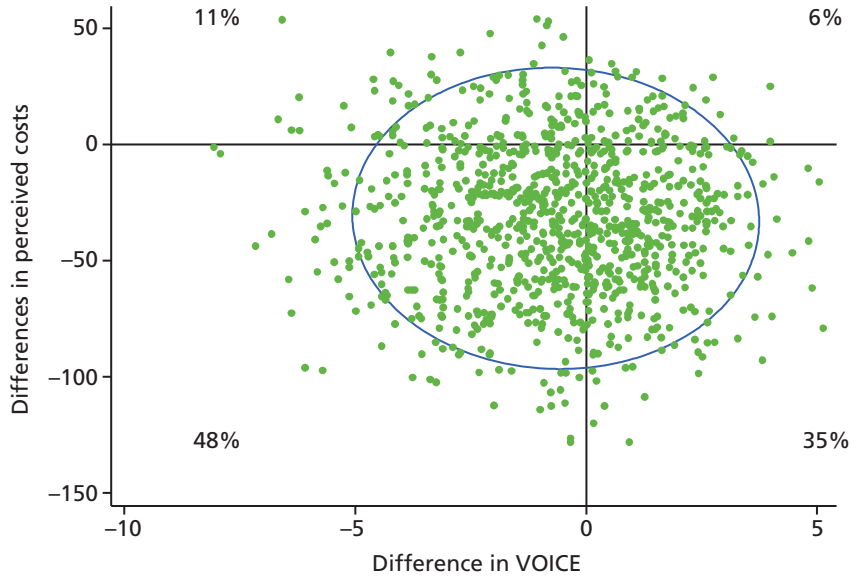


FIGURE 16 Cost-effectiveness plane using VOICE with 95% confidence ellipse (WP4).

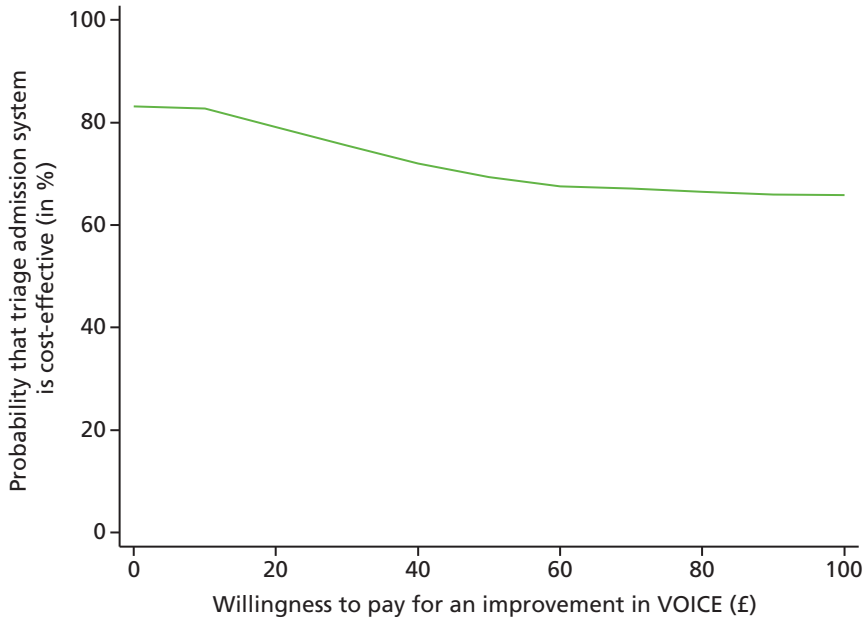


FIGURE 17 Cost-effectiveness acceptability curve using VOICE (WP4).

been had we conducted the study 10 years ago when the triage system was first promoted. Second, it may be that the major contribution to length of stay has less to do with the vigour or success of crisis management and more to do with delayed discharge at the 'other end' of the system, with substantial numbers of service users remaining in hospital because of failures in the delivery of community and social care including access to adequately supported accommodation. Finally, there is the possibility that triage wards encourage admissions of some service users who would not be hospitalised in a routine care service and that our service user data were insufficiently sensitive to pick this up.

Although it could be expected that service users might express a preference for one system over the other for a variety of reasons, we did not find this to be the case. Even when accounting for other variables

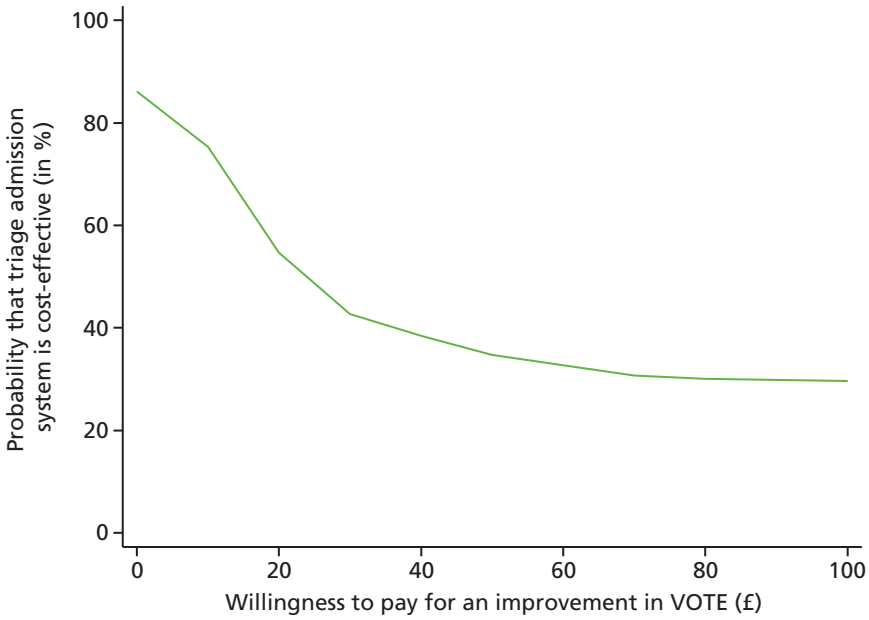


FIGURE 18 Cost-effectiveness acceptability curve using VOTE (WP4).

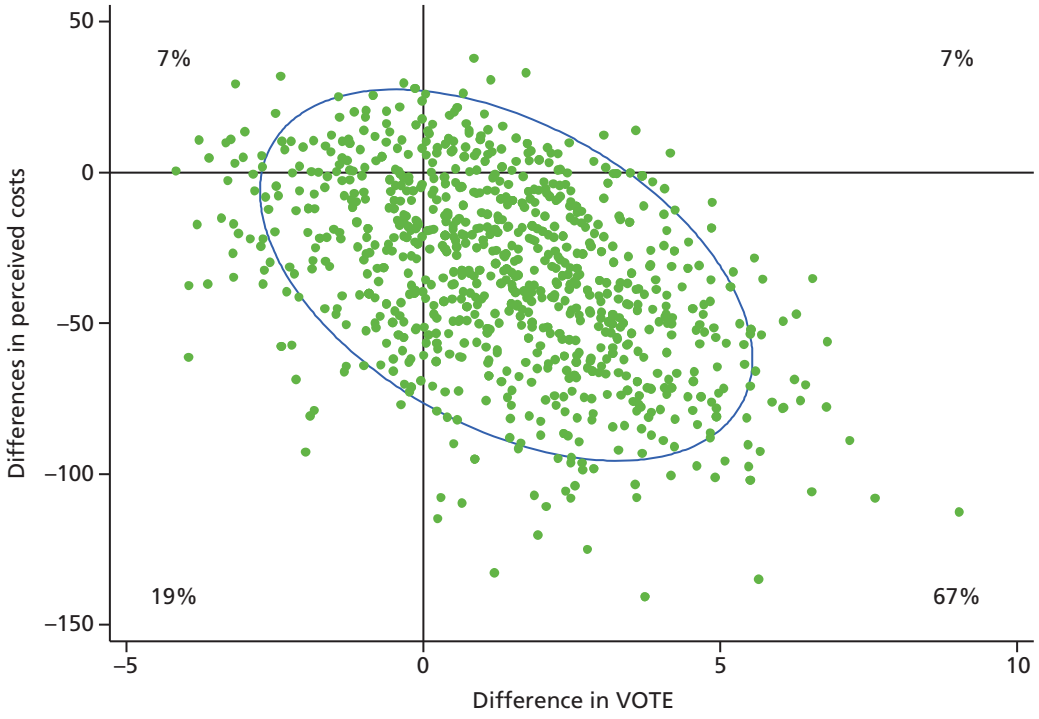


FIGURE 19 Cost-effectiveness plane using VOTE with 95% confidence ellipse (WP4).

(e.g. socioeconomic status) that are associated with differences in perceptions, there is no significant difference in perceptions of care by service users being treated in the two systems. Even when looking at the triage ward on its own, service users did not view it differently from the other wards. The only change in service users' perceptions and satisfaction was a deterioration in these measures over time that occurred in both systems and coincided with a move to shorter lengths of stay and a considerable reduction in the budgets of both systems, including a ward closure in the routine care system.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The perceptions and burnout of nursing staff did not differ between the two systems, although both got worse over time, which mirrored service user perceptions. Although we suggested that working on a high-pressure triage ward would increase burnout and dissatisfaction, we did not support this hypothesis. In fact, when looking at the wards themselves, a difference did emerge. Burnout scores were highest on the triage system's locality wards compared with the triage ward itself and the routine care wards, and the triage system had less satisfied nurses than routine care. When teased apart into types of ward, nurses on the locality wards in the triage system had the lowest satisfaction. This finding can be explained as the consequence of the triage ward system, as the service users with the more persistent and challenging behaviours tend to concentrate over time on the locality wards, particularly if community residential alternatives are slow coming forward or not available. As noted in our introduction, staff in acute wards often report that as many as one-third of the service users have delayed discharges because of failures in 'social' care, and the protracted stay contributes to both staff and service user dissatisfaction. This may be compounded in a triage system as the triage ward is perceived not just as the innovation attracting interest and attention but also as a setting in which there is a steady turnover of service users with none of the problems of long-stay service users.

In terms of the costs, the analyses suggest that the cost of meaningful staff contacts might be lower in the triage system, however, this was not significant. Similarly, there was no clear evidence for differences between the triage ward and long-stay wards in the triage system in terms of activity costs.

A striking and unexpected finding in the study was that both service users and staff reported a worsening in satisfaction with inpatient services across the period of our research regardless of whether or not they were managed in triage or routine care. Sadly, these data chime with other reports from elsewhere in the NHS that suggest a more pervasive influence on staff morale and patient experience working across the wider organisation in recent years.[18,170] They are also surely not the only data to demonstrate this decline in qualitative experience coinciding with a period of budget tightening and system pressure to increase efficiency that was reflected in our data by a universal reduction in length of hospitalisation and continuing very high levels of occupancy on the wards.

This issue was in the news as the then Deputy Prime Minister, Nick Clegg, made a speech on 20 January 2014,[185] alongside a publication by the Department of Health,[3] promising action to direct attention at improving the quality of care in our hospitals that should certainly improve the service user and carer experience. At the same time, the *Guardian*[4] also ran a feature describing the harmful effects of bed closures on inpatient services from the point of view of a former psychiatric nurse revisiting his previous place of employment and finding it much deteriorated from a decade ago.

Clinical implications

Although the results of the study do not suggest any benefit of triage system over routine care, it would be wrong to conclude that the approach does not have some merit or may not be worthy of further study. Our data show that, when considered in isolation, the triage ward is able to discharge a greater proportion of hospitalised service users in the 7- to 10-day period than is achieved by standard care and there is no apparent increase in re-admission of these patients as was thought likely. Furthermore, contrary to expectations, staff in the triage ward were no more likely to report burnout than staff in routine care. What is clear is that there may be consequences for staff morale and satisfaction on the remaining wards in such a care pathway and this may require special attention.

Implications for future research

As noted in the section *Strengths and limitations*, this is a study of one triage system in one area of the inner city in London. This was the only service available at the time. Since then, triage systems have been adopted in several areas and are also contained in our own data in WP3. Future research would benefit from the inclusion of more of these new sites and perhaps a more detailed appraisal of the clinical characteristics of service users as there are anecdotal concerns that triage might encourage hospitalisation of service users who would otherwise be managed as effectively by crisis HTTs.

In conclusion, our study suggests that introducing triage wards is not enough to address the wider challenge of shortening the length of hospitalisation, as this is also determined by the availability of social care in the community, including access to appropriate housing and community support, much of the provision of which lies outside the control of the NHS.

Conclusion

The study demonstrated that, for an inner city acute service, triage wards confer no specific benefit to length of stay and they are not preferred by either staff or service users. Contrary to our hypothesis, triage ward staff did not exhibit more detrimental effects than staff on other wards. There is also some limited evidence that nurses on wards that are part of the triage pathway experience lower morale and less satisfaction with their role.[184]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 6 Discussion and conclusions from the PERCEIVE programme

The PERCEIVE programme makes an important contribution to the evaluation of inpatient mental health services. Mental health services in the UK are in crisis with budgetary cuts, pressure on beds, staff demoralisation and service user dissatisfaction (e.g. Mind[10] and Clegg[185]). In light of the recent Francis report,[18] improving the experience of service users is of upmost importance across the entirety of the NHS. The overarching aim of the project was to evaluate efforts to improve the service user experience using measures designed by service user researchers and nurses.

Therefore, the PERCEIVE programme had several components. First in WP1 LIAISE, we developed measures of stakeholder perceptions of services for use across the programme through an iterative process involving service users and nurse researchers in their design and including feedback from service users and nurses who live and work on the wards throughout the development. We also produced for the first time, in WP2 CITRINE, a measure of patient-level costs that also involved all stakeholders. In this way, we were able to ensure that what is important to service users and nurses are the questions being asked, not what an expert researcher or clinician judges as the important questions to ask. These measures are now available for future evaluations of health services.

Second, in response to the numerous reports about how service users are dissatisfied with the lack therapeutic activities and interactions on inpatient wards, the intense boredom and the lack of anything to do on inpatient wards, our WP3 DOORWAYS study aimed to introduce more therapeutic activities and interactions onto wards. With trust management and ward staff involvement, we developed a comprehensive training programme that was unique to each ward's specific needs in order to give the nurses the skills and confidence to run evidence-based group activities. We then evaluated the feasibility of this training programme and its real-life impact on wards against a changing background of trust upheaval and reorganisation and further budgetary cuts.

Finally, in WP4 BETTER PATHWAYS, we evaluated a novel approach to deal with the pressure on wards to reduce length of stay and have enough beds for those who need them. In a pioneering move, one catchment area in our trust operated a novel system of admission and discharge, modelled on a system used in accident and emergencies around the world. The inpatient service has senior clinicians on duty 7 days a week and service users are thoroughly evaluated on arrival and discharge planning begins immediately so that the individual is treated in the way most appropriate to his or her needs. Prior to our study, the only assessment had been a pre–post evaluation, which is fraught with problems of interpretation. PERCEIVE designed the evaluation of this triage model with a rigorous comparative design. PERCEIVE evaluated stakeholders' perceptions as well as the economic benefits or drawbacks and also whether or not it was working, as it was designed when it started originally a decade earlier.

Summary of results

Excerpts from Wykes *et al.*[186] This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

This programme of work involved four distinct but interrelated studies.

LIAISE, the first WP, involved the development of two measures of the perception of living on, and working in, inpatient wards from the perspective of both service users (VOICE) and nurses (VOTE). These measures were developed by a service user researcher and nurse researcher as we believe (and there is

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

evidence that) individuals in focus groups and reference groups are more comfortable when the researcher is from the same community and that they are more candid in their views. These measures were then psychometrically validated as part of the final two WPs.

The second WP, CITRINE, developed a self-report measure of staff contacts and the uptake of therapeutic activities on inpatient wards, which was subsequently used to calculate the costs of therapeutic activities and service user and staff interactions in WPs 3 and 4. Clinical staff and service users were consulted in the process of the questionnaire's development. It was piloted to determine validity, reliability and acceptance from service users. The final questionnaire records, through self-report, service users' activities and meaningful interactions with ward staff. Self-report is important as this project sought to carry out assessments from the service user perspective and particularly what they viewed as meaningful. The concept of what constitutes a meaningful contact was much debated in the early stages of the study as it was found that service users reported fewer staff interactions as meaningful compared with the actual number (not assessed as being meaningful or not) that was observed.

DOORWAYS, the third WP, sought to increase the number of therapeutic activities service users could potentially take part in. Ward staff often report that crisis resolution, the paperwork and documentation they must complete and a lack of confidence, or a loss of skills over time, often prevent them from engaging in a therapeutic way with service users. This WP attempted to change this. First, we did this by attempting to provide the majority of ward staff with training in how to communicate effectively. This training was cofacilitated by a service user in order to present communication skills from a service user perspective. Second, clinical psychologists ran training sessions and provided in situ training in a number of evidence-based therapeutic activities. That our intervention proved less effective than we anticipated (although there was a significant improvement in people compulsorily admitted) may have much to do with wider changes in the services that were outside the control of the research. Much of the success of our approach was predicated on stable ward management and ethos specific to each ward. We found that the support of the ward managers was absolutely vital to the implementation of DOORWAYS. In several instances, wards did not have managers, were in the process of shutting completely or had job reorganisations that had an impact on whether or not our project was able to make a meaningful difference, despite the best efforts of the nursing staff.

In the final WP, BETTER, we sought to evaluate the cost-effectiveness and stakeholder perceptions of a novel method of admission and discharge. A triage service aims to facilitate the quick transition to the appropriate type of care service that users will benefit from; therefore, improving stakeholder perceptions and reducing hospital stays and costs. Our study did not provide evidence to support the premise that this was a more efficient system. However, there were a number of contextual reasons why the service did not fulfil its earlier promise. Across the entire trust, all catchment area mental health services were facing budgetary cuts and service reorganisation. During this turbulent time, the contrasting routine system may have focused on reducing admission length and did reduce the number of available beds, hence, the contrasting system was also processing admissions and discharges in a more 'triage like' way.

Patient and public involvement in the PERCEIVE study

The PERCEIVE study aimed to put service user involvement at the heart of increasing the evidence base on improving care and the experience of people using adult acute mental health services in the NHS.[187] The importance of the experience of NHS services is now a policy priority in the wake of the Francis report[18] into the failings of Mid Staffordshire Hospitals NHS Trust. At the same time, the PERCEIVE programme sought to assess whether increasing activities and psychological interventions on wards or changes in the service delivery improved both clinical outcomes and experience. Research suggests that these are priorities for service users.[20]

There is an overall emphasis on service user involvement as part of the NIHR application process with clear descriptions of the intended involvement throughout the programme. However, there is little information available on whether or not the research actually managed to achieve this involvement. As service user

and other stakeholder involvement was particularly emphasised in this programme grant, not just in the governance of the programme but also in the research objectives and the conduct of the research itself, the following sections describe how we achieved this across the programme grant as a whole and as far as possible assesses the impact of this involvement.

Service user involvement in the governance structure

Service users were members of the Steering Committee and the Data Monitoring and Ethics Committee. However, their attendance was poor and although each meeting did have service user involvement, this could have been improved. We will, in the future, encourage attendance through a more proactive approach and organise meetings at times more suitable to service users (e.g. evenings or weekends).

Design of the PERCEIVE programme

The programme had a coapplicant who is a service user researcher (DR) and has experience of psychiatric inpatient care including of detention and compulsory treatment. However, it should not be thought that what is being brought to the table is some idiosyncratic 'personal' experience. Having used inpatient services for four decades and also having a research career, DR has both conceptualised her own experience in terms of the research literature and conceptualised the research literature in terms of her experience and, crucially, others who share it. Her understanding is scientifically and socially shaped.

DR led WP1 (LIAISE), which consisted of developing outcome measures from the perspectives of service users and nurses. To do this, LIAISE utilised a model which DR has used in developing patient-generated PROMs since 2007[17] and she also contributed to the design of the other three WPs in terms of the research priorities of service users and the kinds of interventions that they would find acceptable and helpful. She helped in the advancement of the conceptualisation of DOORWAYS (WP3) as not just an increase in activity and psychological therapy but a whole-ward enhancing intervention that might be expected to lead to a better patient experience. She also supplied a dose of common sense in terms of the difficulties that could be expected and resistance to organisational change in the NHS, particularly in mental health services.

Service user involvement in measure development

Work package 1 developed a measure from the service user perspective. It was a user-led piece of work and the measure is called VOICE. It also set out to develop a measure from the nurse perspective (VOTE). These measures were not measures of 'satisfaction' as they are fraught with difficulties,[104] but aimed to tap directly into experiences and perceptions of acute care. For the service user measure, we employed a service user researcher. The first phase followed a participatory approach in that both participants and researchers had experience of service use. It began with four focus groups made up of people who had been inpatients in the local trust in the previous 2 years, with one group made up exclusively of people who had been detained. However, in the event, all groups included people who had been detained. The four focus groups each met twice and only one person across all four groups did not return for the second meeting. This indicates that they were very engaged with the process and it might be argued that this was because it was user led. The details of the rest of the measure generation can be found in *Chapter 2*. One further point worth noting is that the measure was subject to full psychometric testing and was found to be robust. This is interesting as the population on whom it was tested was an inpatient population and standard textbooks suggest that the 'subjects' for psychometric testing should be 'cognitively intact'. Many of our participants had diagnoses of psychosis which might be thought to compromise them in this task and so, again, it may be the fact of being user generated that makes the measure amenable even to people who have poor mental health because it is about things that really matter to them at that moment.

In order to further assess the impact of service user involvement on research, our programme was included as part of a study carried out by the NIHR Mental Health Research Network. The service user researcher from the LIAISE WP was interviewed by an independent evaluator for studies adopted by the Mental Health Research Network (MHRN). She had this to say:

> *There's a huge amount of benefit to employing service user researchers. It's very valuable to be able to say to service users this is who I am. When you're running a group or an interview, it breaks down*

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

those traditional barriers. Imagine if you had a psychiatrist running a focus group asking people about issues on the ward – how open are they going to be about that? There is something about a shared experience that makes people feel more comfortable and more open and honest in what they say. I remember when I was talking to one participant she said 'I feel like I can tell you everything, you understand me and I can really talk to you'. So I think it makes a massive difference.

Once VOICE was complete, it was used in WP3 (DOORWAYS), the trial of activities and psychological interventions, and WP4 (BETTER PATHWAYS), the triage ward study. The service user researcher collected data and there were other service users collecting data as well. The recruitment rates were no different between service user researchers and the rest of the recruitment team – all were uniformly high.

How does VOICE perform?

In the battery of measures for service users to complete, there was also a conventionally derived satisfaction one. In WP4 (BETTER PATHWAYS), VOICE was sensitive to differences in activities that participants were engaged in but the conventionally derived one was not. In WP3 (DOORWAYS), VOICE did not find that perceptions improved overall post interventions, which was to be expected as the number of activities taking place did not increase to a significant level. However, it was able to detect differences in those who were compulsorily detained.

Extending service user involvement to other stakeholders

INVOLVE, the programme set up by NIHR to promote patient and public involvement in research, states that mental health professionals do not count as the 'public' for involvement purposes. However, we did extend our model of measure development to nursing staff. We proceeded in exactly the same way as for VOICE to produce a measure named VOTE. In this case, we recruited nurse researchers to work with focus groups and expert panels of nursing staff and then to collect data on the wards. The participants in the groups and panels were current staff. We also employed a nurse researcher as part of the team of researchers collecting data for WPs 3 and 4.

CITRINE measure development

For the development of the measure of patient-level costs of inpatient care, CITRINE, we also included both service users and nurses. Service user researchers provided advice on the whole project and also collected data for reliability and validation exercises. In the measure development, staff members were interviewed about a variety of issues related to activities of nursing staff in relation to inpatients. Focus groups in the form of ward-based community meetings were held with service users in the initial measure development and to discuss the presentation of the self-report measure. Following advice from service users, the final measure was adapted in presentation to ensure that it would be suitable for as many people as possible by providing some support for completion (see *Chapter 3* for details).

Service user involvement in data collection

In addition to the close involvement of stakeholders in all measure developments, service user researchers were also part of the research teams responsible for data collection in all other areas of the programme grant. These staff members were recruited on the same basis as all members of the team (for their professionalism and expertise as researchers) but also required the added expertise of experience as a service user.

It is difficult to assess their impact on the team, but a number of issues did arise for those service users who had received inpatient care. In particular, some of the trauma associated with their own admission reasserted itself for them. This was dealt with through close supervision that also included a senior service user researcher who also had experience of admission as well as a senior clinical psychologist.

Although service user involvement was seen by the research team as a benefit, there was resistance by some ward staff who were concerned that service users were able to access case notes (with participant consent). However, the professionalism of our staff and close supervision of the team was able to reduce the distress to our staff and overcome the ward staff resistance.

Service user impact on analyses

Service users were involved in the development of all the analysis plans and were especially involved in the analyses in WP1, in which they specifically discussed with statisticians the best approach to the factor analyses and the assessment of the measure psychometrics. Service users were particularly important in helping to understand the statistical output and often were able to fall back on the qualitative data we had collected as well as their own experience to ensure that the team understood the meaning of tables and also to lead the secondary or exploratory analyses by providing potential hypotheses to test. A final set of focus groups were carried out, facilitated by a service user researcher, with the aim of understanding if the inputs to WP3 had been helpful and if there were any other issues that needed consideration in the analysis or follow-up of the programme grant.

Service user impact on dissemination

In keeping with our ethos, all papers had at least one service user author. Service user researchers also presented data from the programme at local, national and international research meetings. During the programme, we kept stakeholders involved in our programme by updating them on the progress through newsletters for which service users contributed and acted as editors. A further round of dissemination is planned when the current programme is completed.

Interim feedback has also been delivered and is planned to be delivered to service managers in order to ensure that the programme provides useful local data for service planning. The first of these meetings was with the Medical Director, Clinical Service Directors and Clinician Lead for the inpatient services to feedback the data from WP4. This feedback also allowed these stakeholders to suggest new analyses for future data exploration which may be of use locally. We have also planned similar exercises with service managers and ward managers as well as local service users.

Service user impact on protocol developments (add-on studies)

During the process of carrying out the programme grant protocol and as we analysed its results, other issues emerged. These became the focus of add-on studies.

Design in mind

We noted during the research that the design of wards often had an effect not only on the overall ambience and atmosphere of the ward but also the activities that it could support. Following an extensive review, it became clear that there was no hard evidence of these effects that could be used as part of our programme measures. Therefore, we undertook a further development in a user co-led study to develop a measure of mental health ward design using a participatory method and involving service users and staff. Following this development, a novel method was developed and tested in a study involving service user photographs of ward design. We asked service users to take photographs of their best and least liked features of the ward design and then we uploaded these photographs and invited those who had taken them to speak about them. This method allows more service user involvement as it does not limit the scope of the designs or their potential impact and also allows service users to be involved who may not be able to articulate clearly their preferences and dislikes in a more abstract way. The impact of service users was, therefore, not only to provide intellectual input via the review process but also to develop methodology.

Perinatal VOICE

Similar impacts occurred in an extension of the VOICE measure to perinatal services that arose out of discussions with colleagues undertaking their own programme grant assessing perinatal mental health services. After discussions with their own service user researchers and advisors, it was decided that our project would assess if women using perinatal mental health services had different concerns and needs than those service users in our programme grant. As a result, a measure specifically to be used with women in the perinatal period was developed as a modification of VOICE. This is now being psychometrically tested online with the help of a perinatal non-governmental agency and the Service User Advisory Group of the perinatal programme grant.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Strengths and limitations

The primary strength of the PERCEIVE programme grant was that its focus was on the experience of those using and working in inpatient services. The vast majority of service evaluations and improvement schemes are based on what expert academic researchers or clinical managers put forward as worthy of investigation. There is a distinct lack of stakeholder involvement in deciding what factors and questions are important in service evaluation or how to measure them. An aim of our project was to deliver service user- and nurse-generated, psychometrically tested, tools to assess services. Our use of other established measures alongside them, which picked up on the same patterns as our measures, provided further validity.

Our project was run in services that were changing as we progressed, which meant that we were attempting to change and measure systems that were already in flux. As mentioned in *Chapter 5*, WP4 was proposed as a comparison of two distinct triage services. However, one of the boroughs had such a lag in converting their service into a triage service this change did not occur. Furthermore, our project took place during a time when, nationwide, services were being squeezed. Funds were cut and, in three of five services in the study, wards were being closed, which contributed to the pressure on beds as well staff morale (especially in cases where staff had to reapply for their posts). In addition, wards lost ward managers and were managed by part-time managers or less experienced staff acting as temporary managers. Adverse events, for example the unexpected death of valued member of staff, had an impact on both ward colleagues and the service users. All of these background variables contributed to the turbulence on wards and hence the result that perceptions of the services fluctuated for both groups of stakeholders.

Research is evaluated on the generalisability of findings. We were particularly concerned with our DOORWAYS study, in which we tried to increase the number of psychologically based activities on wards. Our initial trust base is part of a longstanding and nationally recognised academic NHS partnership and has recently become part of an academic science centre. The users of these services are research active. We also tested the effects in a different trust that has fewer academic ties. The culture and ethos of the second trust, as well as the demographics in the area it serves, are very different from the original trust. The fact that we did not find differing results between these trusts leads us to conclude that our results are generalisable to comparable trusts in large cities nationwide, regardless of their research activity.

The stepped-wedge design used in WP3 led to some challenges when analysing the data. The original proposal was designed for the initial eight wards, but our Data Monitoring and Ethics Committee suggested that we increase the numbers of wards (clusters) in order to improve generalisability and allow for more accurate modelling of the between-ward variation. However, although we did increase power (in terms of the numbers of patients), once the project reached the analysis phase it was clear that our modification of the design was not as effective as we had hoped. This was because of two features of the data. The first was that the assumption of random variation between wards, as required for some of the statistical models, did not seem to be very realistic (there was at least one apparent outlier, for example). Second, external changes over time were substantial and affected the wards in different ways, increasing their variability so that even the new total of 16 wards was likely to be too low for accurate modelling of the random effects. These problems affected our ability to detect only effects in service users compulsorily admitted above background ‘noise’. However, the design did have benefits over a simple cross-sectional comparison. Neither the wards nor the researchers knew which wards would be randomised next (with the exception, of course, of the last wards). Pre-intervention measures of the variables that were to be our key outcomes were measured by self-report prior to each randomisation so neither staff nor service users in the control wards knew whether or not they would receive the active treatment (staff training). However, as a result of the practicalities of the DOORWAYS trial, researchers could not be blind to which wards had already received the intervention, as they spent time on all the wards and would see the activities occurring.

Some practicalities of research could not be avoided. Three different psychologists were employed to carry out the training of ward staff, which may have led to a variation in the quality of training. It must be acknowledged that as the research progressed, implementation methods and engagement improved. Researchers also changed over time but any interview rating was ensured to be reliable against a key rater each time staff changed. The core management team remained the same throughout and aimed to provide consistent training to all staff. Despite these changes, there were no consistent fluctuations over time in any of the outcome measures that we could account for by our changes in staff or learning.

Research implications

Our creation and validation of two stakeholder measures provides our own team as well as other research teams with useful tools in evaluating inpatient services. Colleagues have already approached us about using these assessments in the north of England and the USA, and VOICE has even been translated into Portuguese. We fully encourage the use of these measures.

The CITRINE WP has identified a number of studies that have investigated how inpatient staff spend their time and, specifically, the extent to which service users are in receipt of that time. The overall finding is that staff–service user contact time is very limited. There were no economic evaluations of inpatient-based therapeutic activities identified. This is a clear research gap, especially given that so many resources are tied up in inpatient care. As well as the need for evaluations of new activities, there is also scope (using modelling methods for example) to evaluate existing interventions. We have developed a new tool for measuring staff contacts with service users and activities that the service users undertake. Use of this tool will provide useful economic information needed for evaluating inpatient-based interventions.

DOORWAYS found that implementing such a research project in an unstable and unpredictable environment is not without its challenges. The tools of implementation science need to be used to uncover the factors that might help future work in this area succeed. Our future work includes further exploring the differences between wards, boroughs and trusts. We will also be looking in greater detail at how the absence of management and significant events that destabilise wards interfere with the implementation of the intervention. We have not explored some of the data in great depth so will be carrying out further analyses investigating the impact of the intervention on clinical indicators such as PANSS scores, NOSIE and GAF. We will also be investigating if VOTE scores for those nurses who were trained as part of our programme are different from those nurses who, although on the wards at the time, were not trained.

The BETTER PATHWAYS WP investigates a triage system in one single site in comparison with a routine care system. Since we have completed this project, further sites in our trusts and across the nation have moved to a triage system. No two sites will ever run their triage services in exactly the same way and a comparison between many of these sites would be able to tease out those practices that are working well. What is holding up discharge from locality wards, such as social reasons like housing, would also be worthy of further research.

Clinical and organisational implications for mental health inpatient units

This programme demonstrates the valuable contribution that stakeholder research makes. Our consistently high recruitment rates demonstrate that people are eager to help create evidence for good practice. In light of the Francis report,[188] trusts need to increasingly focus on stakeholder experience as a measure of service delivery.

Our study found that the higher the numbers of meaningful contacts that service users had with nurses, the better were their perceptions of services. So, despite the negative media presentation of services, including inpatient mental health services, those receiving such services appreciate quality communication with nurses. We also discovered that regardless of how acutely ill a person may be they are still able and, in fact desire, to take part in therapeutic activities. This is in opposition to the general concerns that staff had

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

about whether or not service users in such acute stages of mental ill health were able to engage in therapeutic activities. These findings suggest that NHS leaders and service directors need to consider support structures for staff to create an ethos on wards where therapeutic activities and interactions are of equal importance to the completion of paperwork and the management of crises.

Increasing activities did not lead to a worsening of staff perceptions of services. This is clearly important as the extra burdens may have affected the quality of interactions with inpatients. This was also a time during which services across both trusts were being cut and our results are perhaps not surprising. As noted above, this finding demonstrates the importance of service director and manager support to ensure that wards are staffed, organised and run adequately not only for the benefit of the services users but also for the nurses themselves. This could also contribute to greater staff retention, which is an issue across the NHS. Nurses often fed back to our team that they would like to be able to run more activities and engage more with service users as those are the things that motivate one to choose a career in psychiatric nursing. The economic analyses revealed that there was a modest cost increase following the DOORWAYS intervention. This is not an expenditure increase but rather a redistribution of already allocated resources towards more direct care. There was limited suggestion of cost-effectiveness of DOORWAYs, but only if a high value is placed on a unit improvement in the VOICE. In terms of staff perceptions, the intervention did not appear cost-effective.

Our study did not find any clear advantage of a triage system over that of routine care. We found that the introduction of a triage system did not result in a reduction in the length of stay, nor were perceptions of the wards very different between the systems. However, it did not lead to an increase in re-admissions, as feared, and costs were not significantly different from non-triage services. On its own, the triage ward functions as it should (discharging people rapidly) and, as expected, the nurses did not report more burnout. What was unexpected was that staff on the locality wards are finding it very challenging. In the locality wards, nurses reported the most burnout and this needs further exploration. The triage system does not work in isolation and the effects on locality wards especially on what can improve staff moral would be a possible next step in its evaluation.

Conclusions

In the current climate of the NHS, it is becoming increasingly clear that the experience, particularly that of service users, is a vital component of the evaluation of the quality of services. Our study will contribute to quality evaluation by providing a methodology and a set of outcome measures that will strengthen these research efforts. However, all future studies need to understand that variation in inpatient care between wards, within wards and across time are likely to undermine the surfacing of clear results. Despite the best efforts of service directors and nurses working on wards, there is clearly room for further improvement to inpatient wards. Our research makes a valuable contribution in understanding how inpatient wards cope with a changing background and offers service planners useful information in how they can continue to try and improve service users' experiences while also being cost-effective. We hope others will take forward our ideas and findings and that our large databases will provide fruitful ground for future investigations to aid our understanding of the factors that mediate and moderate ward interventions, their implementation and their sustainability.

Acknowledgements

Important information

We gratefully acknowledge the NIHR Programme Grants for Applied Research programme funding that enabled us to undertake this programme of research. We thank all the ward teams, ward management and clinical service directors who contributed to the development and implementation of the programme of research. Their advice and support throughout the project was invaluable. We also acknowledge the very valuable comments and guidance of our Project Steering Group and Data Monitoring and Ethics Committee and, in particular, all of our service user advisors. We would also like to thank Morven Leese and Clare Flach who were our initial statistical advisors and who have kept an interest in the project throughout. Finally, this programme of research could not have been successfully completed without the participation of both staff and service user inpatients who willingly gave their time to participate in all the studies.

Contributions of authors

This report was coauthored by eight team members. Each chapter had multiple contributors and all chapters were reviewed in draft form by all authors.

Til Wykes was the principal investigator and oversaw the whole programme implementation and led WP3.

Emese Csipke fulfilled the programme co-ordinator role throughout the projects and has substantially contributed to drafting the report.

Diana Rose led studies in WP1.

Thomas Craig led on WP4.

Paul McCrone led on WP2.

Paul Williams, **Leonardo Koeser** and **Stephen Nash** provided methodological advice.

Publication

Csipke E, Wykes T, Nash S, Williams P, Koeser L, McCrone P, *et al.* Changing nurses' views of the therapeutic environment on acute inpatient wards: an RCT. *BJPsych Open* 2018; in press.

Data-sharing statement

All data requests should be submitted to the corresponding author for consideration. Access to available anonymised data may be granted following review.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

References

1. Department of Health. *No Health Without Mental Health: A Cross-government Mental Health Outcomes Strategy for People of All Ages*. London: Department of Health; 2011.
2. Mental Health Taskforce. *The Five Year Forward View for Mental Health*. London: NHS England; 2016.
3. Social Care Local Government and Care Partnership Directorate. *Closing the Gap: Priorities for Essential Change in Mental Health*. London: Department of Health; 2014.
4. Filer N. *Mental Health Care: Where Did It All Go So Wrong?* 2014. URL: www.theguardian.com/society/2014/jan/25/nathan-filer-mental-health-care-where-did-it-go-wrong (accessed 27 January 2014).
5. PERCEIVE. *PERCEIVE – Patient Involvement in Improving the Evidence Base on Inpatient Care*. URL: www.perceive.iop.kcl.ac.uk/about.html (accessed 10 January 2018).
6. Mental Health Strategies. *2010/2011 National Survey of Investment in Adult Mental Health Services*. Manchester: Mental Health Strategies; 2011.
7. Dean C, Phillips J, Gadd EM, Joseph M, England S. Comparison of community based service with hospital based service for people with acute, severe psychiatric illness. *BMJ* 1993;**307**:473–6. https://doi.org/10.1136/bmj.307.6902.473
8. Priebe S, Badesconyi A, Fioritti A, Hansson L, Kilian R, Torres-Gonzales F, *et al.* Reinstitutionalisation in mental health care: comparison of data on service provision from six European countries. *BMJ* 2005;**330**:123–6. https://doi.org/10.1136/bmj.38296.611215.AE
9. Smyth MG, Hoult J. The home treatment enigma. *BMJ* 2000;**320**:305–8. https://doi.org/10.1136/bmj.320.7230.305
10. Mind. *Ward Watch: Mind's Campaign to Improve Hospital Conditions for Mental Health Patients*. London: Mind; 2004.
11. Wing JK, Brown GW. *Institutionalism and Schizophrenia*. Cambridge: Cambridge University Press; 1970. https://doi.org/10.1017/CBO9780511753374
12. Poole AE, Sanson-Fisher RW, Thompson V. Observations on the behaviour of patients in a state mental hospital and a general hospital psychiatric unit: a comparative study. *Behav Res Ther* 1980;**19**:125–34. https://doi.org/10.1016/0005-7967(81)90036-X
13. Higgins R, Hurst K, Wistow G. Nursing acute psychiatric patients: a quantitative and qualitative study. *J Adv Nurs* 1999;**29**:52–63. https://doi.org/10.1046/j.1365-2648.1999.00863.x
14. Kavanagh K, Duncan-McConnell D, Greenwood K, Trivedi P, Wykes T. Educating inpatients about their medication: Is it worth it? *J Ment Health* 2003;**12**:71–80. https://doi.org/10.1080/0963823002100005831 9
15. Sharac J, McCrone P, Sabes-Figuera R, Csipke E, Wood A, Wykes T. Nurse and patient activities and interaction on psychiatric inpatients wards: a literature review. *Int J Nurs Stud* 2010;**47**:909–17. https://doi.org/10.1016/j.ijnurstu.2010.03.012
16. Sainsbury Centre for Mental Health. *Acute Care 2004 – A National Survey of Acute Psychiatric Wards in England*. London: Sainsbury's Centre for Mental Health; 2004.
17. Rose D, Evans J, Sweeney A, Wykes T. A model for developing outcome measures from the perspectives of mental health service users. *Int Rev Psychiatry* 2011;**23**:41–6. https://doi.org/10.3109/09540261.2010.545990

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

18. Francis R. *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. London: The Stationery Office; 2013.
19. Rose D. *Users' Voices: The Perspectives of Mental Health Service Users on Community and Hospital Care*. London: Sainsbury Centre for Mental Health; 2001.
20. Rose D, Wykes T, Farrier D, Doran A, Sporle T, Bogner D. What do clients think of cognitive remediation therapy? A consumer-led investigation of satisfaction and side effects. *Am J Psychiatr Rehab* 2008;**11**:181–204. https://doi.org/10.1080/15487760801963694
21. Rose D, Sweeney A, Leese M, Clement S, Jones IR, Burns T, *et al.* Developing a user-generated measure of continuity of care: brief report. *Acta Psychiatr Scand* 2009;**119**:320–4. https://doi.org/10.1111/j.1600-0447.2008.01296.x
22. Knapp M, McDaid D, Medeiors H, Becker T, Johnson S, Kilian R, *et al. Economics, Mental Health and Policy: An Overview*. London: Mental Health Economics European Network; 2008.
23. Kinton M. *The Mental Health Act Commission: In Place of Fear?* London: Mental Health Act Commission; 2005.
24. Saxena S, Barrett T. *Atlas: Nurses in Mental Health 2007*. Geneva: World Health Organization; 2007.
25. Brennan G, Flood C, Bowers L. Constraints and blocks to change and improvement on acute psychiatric wards – lessons from the City Nurses project. *J Psychiatr Ment Health Nurs* 2006;**13**:475–82. https://doi.org/10.1111/j.1365-2850.2006.00956.x
26. Cleary M. The realities of mental health nursing in acute inpatient environments. *Int J Ment Health Nurs* 2004;**13**:53–60. https://doi.org/10.1111/j.1447-0349.2004.00308.x
27. Wykes T, Stevens W, Everitt B. Stress in community care teams: will it affect the sustainability of community care? *Soc Psychiatry Psychiatr Epidemiol* 1997;**32**:398–407.
28. Corry P. *Behind Closed Doors – Acute Mental Health Care in the UK*. London: Rethink; 2004.
29. Department of Health. *Mental Health Policy Implementation Guide. Adult Acute Inpatient Care Provision*. London: Department of Health Publications London; 2002.
30. Sainsbury Centre for Mental Health. *Acute Problems: A Survey of the Quality of Care in Acute Psychiatric Wards*. London: The Sainsbury Centre for Mental Health; 1998.
31. Ford R, Durcan G, Warner L, Hardy P, Muijen M. One day survey by the Mental Health Act Commission of acute adult psychiatric inpatient wards in England and Wales. *BMJ* 1998;**317**:1279–83. https://doi.org/10.1136/bmj.317.7168.1279
32. Shattell MM, Andes M, Thomas SP. How patients and nurses experience the acute care psychiatric environment. *Nurs Inq* 2008;**15**:242–50. https://doi.org/10.1111/j.1440-1800.2008.00397.x
33. Edwards K. Service users and mental health nursing. *J Psychiatr Ment Health Nurs* 2000;**7**:555–65. https://doi.org/10.1046/j.1365-2850.2000.00353.x
34. Walsh J, Boyle J. Improving acute psychiatric hospital services according to inpatient experiences. A user-led piece of research as a means to empowerment. *Issues Ment Health Nurs* 2009;**30**:31–8. https://doi.org/10.1080/01612840802500733
35. Faulkner A. Institutional conflict: the state of play in adult acute psychiatric wards. *J Adult Protect* 2005;**7**:6–12. https://doi.org/10.1108/14668203200500020
36. Faulkner A, Thomas P. User-led research and evidence-based medicine. *Br J Psychiatry* 2002;**180**:1–3. https://doi.org/10.1192/bjp.180.1.1

37. Gilburt H, Rose D, Slade M. The importance of relationships in mental health care: A qualitative study of service users' experiences of psychiatric hospital admission in the UK. *BMC Health Services Research* 2008;**8**:92. https://doi.org/10.1186/1472-6963-8-92

38. Gilburt H, Slade M, Rose D, Lloyd-Evans B, Johnson S, Osborn DP. Service users' experiences of residential alternatives to standard acute wards: qualitative study of similarities and differences. *Br J Psychiatr* 2010;**197**(Suppl. 53):26–31. https://doi.org/10.1192/bjp.bp.110.081075

39. Prosser D, Johnson S, Kuipers E, Szmukler G, Bebbington P, Thornicroft G. Mental health, 'burnout' and job satisfaction among hospital and community-based mental health staff. *Br J Psychiatry* 1996;**169**:334–7. https://doi.org/10.1192/bjp.169.3.334

40. Seed MS, Torkelson DJ, Alnatour R. The role of the inpatient psychiatric nurse and its effect on job satisfaction. *Issues Ment Health Nurs* 2010;**31**:160–70. https://doi.org/10.3109/01612840903168729

41. Ward M, Cowman S. Job satisfaction in psychiatric nursing. *J Psychiatr Ment Health Nurs* 2007;**14**:454–61. https://doi.org/10.1111/j.1365-2850.2007.01104.x

42. Hanrahan NP, Aiken LH, McClaine L, Hanlon AL. Relationship between psychiatric nurse work environments and nurse burnout in acute care general hospitals. *Issues Ment Health Nurs* 2010;**31**:198–207. https://doi.org/10.3109/01612840903200068

43. Jenkins R, Elliott P. Stressors, burnout and social support: nurses in acute mental health settings. *J Adv Nurs* 2004;**48**:622–31. https://doi.org/10.1111/j.1365-2648.2004.03240.x

44. White E, Roche M. A selective review of mental health nursing in New South Wales, Australia, in relation to clinical supervision. *Int J Ment Health Nurs* 2006;**15**:209–19. https://doi.org/10.1111/j.1447-0349.2006.00424.x

45. Kindy D, Petersen S, Parkhurst D. Perilous work: nurses' experiences in psychiatric units with high risks of assault. *Arch Psychiatr Nurs* 2005;**19**:169–75. https://doi.org/10.1016/j.apnu.2005.05.002

46. Bowers L, Allan T, Simpson A, Jones J, Whittington R. Morale is high in acute inpatient psychiatry. *Soc Psychiatry Psychiatr Epidemiol* 2009;**44**:39–46. https://doi.org/10.1007/s00127-008-0396-z

47. Leiter MP, Harvie PL. Burnout among mental health workers: a review and a research agenda. *Int J Soc Psychiatry* 1996;**42**:90–101. https://doi.org/10.1177/002076409604200203

48. Cushway D, Tyler PA, Nolan P. Development of a stress scale for mental health professionals. *Br J Clin Psychol* 1996;**35**:279–95. https://doi.org/10.1111/j.2044-8260.1996.tb01182.x

49. Maslach C, Jackson S, and Leiter M. *Maslach Burnout Inventory Manual*. Menlo Park, CA: Mind Garden Inc. 1996.

50. Coomber B, Barriball KL. Impact of job satisfaction components on intent to leave and turnover for hospital-based nurses: a review of the research literature. *Int J Nurs Stud* 2007;**44**:297–314. https://doi.org/10.1016/j.ijnurstu.2006.02.004

51. Coetzee C, Stanz KJ. 'Barriers-to-change' in a governmental service delivery type organisation. *SA J Indust Psychol* 2007;**33**:76–82. https://doi.org/10.4102/sajip.v33i2.380

52. Greenberg J. A taxonomy of organizational justice theories. *Acad Manage Rev* 1987;**12**:9–22.

53. Benn J, Burnett S, Parand A, Pinto A, Iskander S, Vincent C. Perceptions of the impact of a large-scale collaborative improvement programme: experience in the UK Safer Patients Initiative. *J Eval Clin Pract* 2009;**15**:524–40. https://doi.org/10.1111/j.1365-2753.2009.01145.x

54. Haar J, Spell CS, O'Driscoll MP. Organisational justice and work family policies. *South Pacific Journal of Psychology* 2005;**16**:30–9. https://doi.org/10.1017/S0257543400000055

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

55. Eiser JR. *Social Psychology: Attitudes, Cognition and Social Behaviour*. Cambridge: Cambridge University Press; 1986. https://doi.org/10.1017/CBO9780511558290

56. Department of Health. *National Service Framework: Mental Health*. London: Department of Health; 1999.

57. Crawford MJ, Robotham D, Thana L, Patterson S, Weaver T, Barber R, *et al.* Selecting outcome measures in mental health: the views of service users. *J Ment Health* 2011;**20**:336–46. https://doi.org/10.3109/09638237.2011.577114

58. Rose D. Collaborative research between users and professionals: peaks and pitfalls. *Psychiatrist* 2003;**27**:404–6. https://doi.org/10.1192/pb.27.11.404

59. Trivedi P, Wykes T. From passive subjects to equal partners: qualitative review of user involvement in research. *Br J Psychiatry* 2002;**181**:468–72. https://doi.org/10.1192/bjp.181.6.468

60. Gillard S, Borschmann R, Turner K, Goodrich-Purnell N, Lovell K, Chambers M. 'What difference does it make?' Finding evidence of the impact of mental health service user researchers on research into the experiences of detained psychiatric patients. *Health Expect* 2010;**13**:185–94. https://doi.org/10.1111/j.1369-7625.2010.00596.x

61. Hamilton S, Pinfold V, Rose D, Henderson C, Lewis-Holmes E, Flach C, Thornicroft G. The effect of disclosure of mental illness by interviewers on reports of discrimination experienced by service users: a randomized study. *Int Rev Psychiatry* 2011;**23**:47–54. https://doi.org/10.3109/09540261.2010.545367

62. Cooper CL, Sloan SJ, Williams S. *Occupational Stress Indicator*. Berkshire: Nfer-Nelson Windsor; 1988.

63. Lazarus RS, Folkman S. *Manual for the Hassles and Uplifts Scales*. Palo Alto, CA: Consulting Psychologists Press; 1989.

64. Aarons GA. Mental health provider attitudes toward adoption of evidence-based practice: the Evidence-Based Practice Attitude Scale (EBPAS). *Ment Health Serv Res* 2004;**6**:61–74. https://doi.org/10.1023/B:MHSR.0000024351.12294.65

65. Funk SG, Champagne MT, Wiese RA, Tornquist EM. BARRIERS: the barriers to research utilization scale. *Appl Nurs Res* 1991;**4**:39–45. https://doi.org/10.1016/S0897-1897(05)80052-7

66. Hutchinson AM, Johnston L. Bridging the divide: a survey of nurses' opinions regarding barriers to, and facilitators of, research utilization in the practice setting. *J Clin Nurs* 2004;**13**:304–15. https://doi.org/10.1046/j.1365-2702.2003.00865.x

67. Melnyk BM, Fineout-Overholt E, Mays MZ. The evidence-based practice beliefs and implementation scales: psychometric properties of two new instruments. *Worldviews Evid Based Nurs* 2008;**5**:208–16. https://doi.org/10.1111/j.1741-6787.2008.00126.x

68. Ofi B, Sowunmi L, Edet D, Anarado N. Professional nurses' opinion on research and research utilization for promoting quality nursing care in selected teaching hospitals in Nigeria. *Int J Nurs Pract* 2008;**14**:243–55. https://doi.org/10.1111/j.1440-172X.2008.00684.x

69. Olade RA. Attitudes and factors affecting research utilization. *Nurs Forum* 2003;**38**:5–15. https://doi.org/10.1111/j.0029-6473.2003.00005.x

70. Upton D, Upton P. Development of an evidence-based practice questionnaire for nurses. *J Adv Nurs* 2006;**53**:454–8. https://doi.org/10.1111/j.1365-2648.2006.03739.x

71. Van Mullem C, Burke LJ, Dohmeyer K, Farrell M, Harvey S, John L, *et al.* Strategic planning for research use in nursing practice. *J Nurs Adm* 1999;**29**:38–45. https://doi.org/10.1097/00005110-199912000-00008

72. Great Britain. *Mental Health Act 2007*. NHS. URL: www.legislation.gov.uk/ukpga/2007/12/contents (accessed March 2017).

73. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;**2**(14).

74. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989;**45**:255–68. https://doi.org/10.2307/2532051

75. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005;**85**:257–68.

76. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;**33**:159–74. https://doi.org/10.2307/2529310

77. Nunnally JC. *Psychometric Theory*. New York, NY: McGraw-Hill; 1978.

78. Greenfield TK, Stoneking BC, Humphreys K, Sundby E, Bond J. A randomized trial of a mental health consumer-managed alternative to civil commitment for acute psychiatric crisis. *Am J Commun Psychol* 2008;**42**:135–44. https://doi.org/10.1007/s10464-008-9180-1

79. Greenfield TK, Attkisson CC. Steps toward a multifactorial satisfaction scale for primary care and mental health services. *Evaluation Program Planning* 1989;**12**:271–8. https://doi.org/10.1016/0149-7189(89)90038-4

80. Greenfield TK, Attkisson CC. The UCSF Client Satisfaction Scales: II. The Service Satisfaction Scale-30. In Maruish ME, editor. *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment*. Mahwah, NJ: Lawrence Erlbaum Associates; 2004. pp. 813–37.

81. Stamps PL, Piedmonte EB. *Nurses and Work Satisfaction: An index for Measurement*. Ann Arbor, MI: Health Administration Press Perspectives; 1986.

82. Bartlett MS. *Stochastic Population Models in Ecology and Epidemiology*. London: Methuen & Co.; 1960.

83. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**:297–334. https://doi.org/10.1007/BF02310555

84. Evans J, Rose D, Flach C, Csipke E, Glossop H, McCrone P, *et al.* VOICE: developing a new measure of service users' perceptions of inpatient care, using a participatory methodology. *J Ment Health* 2012;**21**:57–71. https://doi.org/10.3109/09638237.2011.629240

85. Laker C, Rose D, Flach C, Csipke E, McCrone P, Craig T, *et al.* Views of the Therapeutic Environment (VOTE): stakeholder involvement in measuring staff perceptions of acute in-patient care. *Int J Nurs Stud* 2012;**49**:1403–10. https://doi.org/10.1016/j.ijnurstu.2012.06.001

86. Wing J, Marriott S, Palmer C, Thomas V. *Management of Imminent Violence: Clinical Practice Guidelines to Support Mental Health Services*. London: Royal College of Psychiatrists; 1998.

87. Duxbury J, Whittington R. Causes and management of patient aggression and violence: staff and patient perspectives. *J Adv Nurs* 2005;**50**:469–78. https://doi.org/10.1111/j.1365-2648.2005.03426.x

88. Olofsson B, Norberg A. Experiences of coercion in psychiatric care as narrated by patients, nurses and physicians. *J Adv Nurs* 2001;**33**:89–97. https://doi.org/10.1046/j.1365-2648.2001.01641.x

89. Rose D, Evans J, Laker C, Wykes T. Life in acute mental health settings: experiences and perceptions of service users and nurses. *Epidemiol Psychiatr Sci* 2015;**24**:90–6. https://doi.org/10.1017/S2045796013000693

90. FLESCH R. A new readability yardstick. *J Appl Psychol* 1948;**32**:221–33. https://doi.org/10.1037/h0057532

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

91. Kincaid JP, Fishburne RP, Rogers RL, Chissom BS. *Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy Enlisted Personnel*. Research Branch Report 8–75. Naval Air Station Memphis: Chief of Naval Technical Training; 1975.

92. Cattell RB. The scree test for the number of factors. *Multivariate Behav Res* 1966;**1**:245–76. https://doi.org/10.1207/s15327906mbr0102_10

93. Svensson B, Hansson L. Patient satisfaction with inpatient psychiatric care. The influence of personality traits, diagnosis and perceived coercion. *Acta Psychiatr Scand* 1994;**90**:379–84. https://doi.org/10.1111/j.1600-0447.1994.tb01610.x

94. Iversen KI, Høyer G, Sexton HC. Coercion and patient satisfaction on psychiatric acute wards. *Int J Law Psychiatry* 2007;**30**:504–11. https://doi.org/10.1016/j.ijlp.2007.09.001

95. Katsakou C, Bowers L, Amos T, Morriss R, Rose D, Wykes T, Priebe S. Coercion and treatment satisfaction among involuntary patients. *Psychiatr Serv* 2010;**61**:286–92. https://doi.org/10.1176/ps.2010.61.3.286

96. Bhugra D, Grenade JL, Dazzan P. Psychiatric inpatients' satisfaction with services: a pilot study. *Int J Psychiatr Clin Pract* 2000;**4**:327–32. https://doi.org/10.1080/13651500050517902

97. Greenwood N, Key A, Burns T, Bristow M, Sedgwick P. Satisfaction with in-patient psychiatric services. Relationship to patient and treatment factors. *Br J Psychiatry* 1999;**174**:159–63. https://doi.org/10.1192/bjp.174.2.159

98. Secker J, Harding C. African and African Caribbean users' perceptions of inpatient services. *J Psychiatr Ment Health Nurs* 2002;**9**:161–7. https://doi.org/10.1046/j.1365-2850.2002.00455.x

99. Kirkbride JB, Morgan C, Fearon P, Dazzan P, Murray RM, Jones PB. Neighbourhood-level effects on psychoses: re-examining the role of context. *Psychol Med* 2007;**37**:1413–25. https://doi.org/10.1017/S0033291707000499

100. Morgan C, Dazzan P, Morgan K, Jones P, Harrison G, Leff J, *et al.* First episode psychosis and ethnicity: initial findings from the AESOP study. *World Psychiatry* 2006;**5**:40–6.

101. Takase M, Kershaw E, Burt L. Nurse-environment misfit and nursing practice. *J Adv Nurs* 2001;**35**:819–26. https://doi.org/10.1046/j.1365-2648.2001.01933.x

102. Tumulty G, Jernigan IE, Kohut GF. The impact of perceived work environment on job satisfaction of hospital staff nurses. *Appl Nurs Res* 1994;**7**:84–90. https://doi.org/10.1016/0897-1897(94)90037-X

103. Jernigan III I, Beggs JM, Kohut GF. Dimensions of work satisfaction as predictors of commitment type. *J Manag Psychol* 2002;**17**:564–79. https://doi.org/10.1108/02683940210444030

104. Williams B. Patient satisfaction: a valid concept? *Soc Sci Med* 1994;**38**:509–16. https://doi.org/10.1016/0277-9536(94)90247-X

105. National Institute for Health and Care Excellence. *Safe Staffing for Nursing in Adult Inpatient Wards in Acute Hospitals. Safe Staffing Guideline [SG1]*. London: NICE; 2014. URL: www.nice.org.uk/Guidance/SG1 (accessed 13 July 2017).

106. Care Quality Commission. *Shaping the Future CQC's Strategy for 2016 to 2021*. London: CQC 2013.

107. Collins JF, Casey NA, Hickey RH, *et al.* Treatment Characteristics of Psychiatric Programs That Correlate with Patient Community Adjustment. *J Clin Psychol* 1985;**41**:299–308. https://doi.org/10.1002/1097-4679(198505)41:3<299::AID-JCLP2270410302>3.0.CO;2-A

108. Department of Health. *Hospital Activity Statistics*. 2008. URL: www.performance.doh.gov.uk/hospitalactivity/data_requests/index.htm (accessed 10 June 2009).

109. Curtis L. *Unit Costs of Health and Social Care*. Canterbury: Personal Social Services Research Unit; 2008.

110. Ryrie I, Agunbiade D, Brannock L, Maris-Shaw A. A survey of psychiatric nursing practice in two inner city acute admission wards. *J Adv Nurs* 1998;**27**:848–54. https://doi.org/10.1046/j.1365-2648.1998.00575.x

111. Garcia I, Kennett C, Quraishi M, Durcan G. *Acute Care 2004: A National Survey of Adult Psychiatric Wards in England*. London: Sainsbury Centre for Mental Health; 2005.

112. Beecham J, Chisholm D, O'Herlihy A, Astin J. Variations in the costs of child and adolescent psychiatric in-patient units. *Br J Psychiatry* 2003;**183**:220–5. https://doi.org/10.1192/bjp.183.3.220

113. McKechnie AA, Rae D, May J. A comparison of in-patient costs of treatment and care in a Scottish psychiatric hospital. *Br J Psychiatry* 1982;**141**:602–7. https://doi.org/10.1192/bjp.141.6.602

114. Dodds P, Bowles N. Dismantling formal observation and refocusing nursing activity in acute inpatient psychiatry: a case study. *J Psychiatr Ment Health Nurs* 2001;**8**:183–8. https://doi.org/10.1046/j.1365-2850.2001.0365d.x

115. Hansson L, Björkman T, Berglund I. What is important in psychiatric inpatient care? Quality of care from the patient's perspective. *Qual Assur Health Care* 1993;**5**:41–7. https://doi.org/10.1093/intqhc/5.1.41

116. Bee P, Richards DA, Loftus SJ, Baker JA, Bailey L, Lovell K, *et al.* Mapping nursing activity in acute inpatient mental healthcare settings. *J Ment Health* 2006;**15**:217–26. https://doi.org/10.1080/09638230600608941

117. Whittington D, McLaughlin C. Finding time for patients: an exploration of nurses' time allocation in an acute psychiatric setting. *J Psychiatr Ment Health Nurs* 2000;**7**:259–68. https://doi.org/10.1046/j.1365-2850.2000.00291.x

118. Sandford DA, Elzinga RH, Iversen R. A quantitative study of nursing staff interactions in psychiatric wards. *Acta Psychiatr Scand* 1990;**81**:46–51. https://doi.org/10.1111/j.1600-0447.1990.tb06447.x

119. Radcliffe J, Smith R. Acute in-patient psychiatry: how patients spend their time on acute psychiatric wards. *Psychiatr Bull* 2007;**31**:167–70. https://doi.org/10.1192/pb.bp.105.009290

120. Sanson-Fisher RW, Desmond-Poole A, Thompson V. Behaviour patterns within a general hospital psychiatric unit: an observational study. *Behav Res Ther* 1978;**17**:317–32. https://doi.org/10.1016/0005-7967(79)90004-4

121. Fairbanks LA, McGuire MT, Cole SR, Sbordone R, Silvers FM, Richards M, Akers J. The ethological study of four psychiatric wards: patient, staff and system behaviors. *J Psychiatr Res* 1977;**13**:193–209. https://doi.org/10.1016/S0022-3956(77)90016-4

122. Cormack D. *Psychiatric Nursing Observed: A Descriptive Study of the Work of the Charge Nurse in Acute Admission Wards Of Psychiatric Hospitals*. London: RCN; 1976.

123. Altschul AT. *Patient–Nurse Interaction: A Study of Interaction Patterns in Acute Psychiatric Wards*. Edinburgh: Churchill Livingstone; 1972.

124. Willer B, Stasiak E, Pinfold P, Rogers M. Activity patterns and the use of space by patients and staff on the psychiatric ward. *Can Psychiatr Assoc J* 1974;**19**:457–62. https://doi.org/10.1177/070674377401900503

125. Lepola I, Vanhanen L. The patient's daily activities in acute psychiatric care. *J Psychiatr Ment Health Nurs* 1997;**4**:29–36. https://doi.org/10.1111/j.1365-2850.1997.tb00174.x

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

126. Ampt A, Westbrook J, Creswick N, Mallock N. A comparison of self-reported and observational work sampling techniques for measuring time in nursing tasks. *J Health Serv Res Policy* 2007;**12**:18–24. https://doi.org/10.1258/135581907779497576

127. Hansen JT, Slevin C. The implementation of therapeutic community principles in acute care psychiatric hospital settings: An empirical analysis and recommendations to clinicians. *J Clin Psychol* 1996;**52**:673–8. https://doi.org/10.1002/(SICI)1097-4679(199611)52:6<673::AID-JCLP9>3.0.CO;2-L

128. Mueser KT, Deavers F, Penn DL, Cassisi JE. Psychosocial treatments for schizophrenia. *Annu Rev Clin Psychol* 2013;**9**:465–97. https://doi.org/10.1146/annurev-clinpsy-050212-185620

129. Penn DL, Roberts DL, Combs D, Sterne A. The development of the social cognition and interaction training program for schizophrenia spectrum disorders. *Psychiatr Serv* 2007;**58**:449–51. https://doi.org/10.1176/ps.2007.58.4.449

130. Siani C, de Peretti C, Millier A, Boyer L, Toumi M. Predictive models to estimate utility from clinical questionnaires in schizophrenia: findings from EuroSC. *Qual Life Res* 2016;**25**:925–34. https://doi.org/10.1007/s11136-015-1120-6

131. Fiszdon JM, Reddy LF. Review of social cognitive treatments for psychosis. *Clin Psychol Rev* 2012;**32**:724–40. https://doi.org/10.1016/j.cpr.2012.09.003

132. Kurtz MM, Richardson CL. Social cognitive training for schizophrenia: a meta-analytic investigation of controlled research. *Schizophr Bull* 2012;**38**:1092–104. https://doi.org/10.1093/schbul/sbr036

133. Harris R, Bradburn M, Deeks J, Altman D, Steichen TJ, Stern J. *METAN: Stata module for fixed and random effects meta-analysis*. EconPapers: Statistical Software Components; 2010.

134. Hodel B, Brenner HD, Merlo MC, Teuber JF. Emotional management therapy in early psychosis. *Br J Psychiatry Suppl* 1998;**172**:128–33.

135. Henderson AR. The impact of social cognition training on recovery from psychosis. *Curr Opin Psychiatry* 2013;**26**:429–32. https://doi.org/10.1097/YCO.0b013e3283642cf1

136. Curtis L. *Unit Costs of Health and Social Care 2012*. Canterbury: Personal Social Services Research Unit; 2012.

137. Löthgren M, Zethraeus N. Definition, interpretation and calculation of cost-effectiveness acceptability curves. *Health Econ* 2000;**9**:623–30. https://doi.org/10.1002/1099-1050(200010)9:7<623::AID-HEC539>3.0.CO;2-V

138. Naughton M, Nulty A, Abidin Z, Davoren M, O'Dwyer S, Kenney HG. Effects of group metacognitive training (MCT) on mental capacity and functioning in patients with psychosis in a secure forensic psychiatric hospital: a prospective-cohort waiting list controlled study. *BMC Res Notes* 2012;**5**:302. https://doi.org/10.1186/1756-0500-5-302

139. Rocha NB, Queirós C. Metacognitive and social cognition training (MSCT) in schizophrenia: a preliminary efficacy study. *Schizophr Res* 2013;**150**:64–8. https://doi.org/10.1016/j.schres.2013.07.057

140. Wölwer W, Frommann N, Halfmann S, Piaszek A, Streit M, Gaebel W. Remediation of impairments in facial affect recognition in schizophrenia: efficacy and specificity of a new training program. *Schizophr Res* 2005;**80**:295–303. https://doi.org/10.1016/j.schres.2005.07.018

141. Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. *Value Health* 2013;**16**:202–10. https://doi.org/10.1016/j.jval.2012.10.010

142. Sabes-Figuera R, McCrone P, Sharac J, Csipke E, Craig TK, Rose DS, *et al.* Developing a tool for collecting and costing activity data on psychiatric inpatient wards. *Epidemiol Psychiatr Sci* 2012;**21**:393–9. https://doi.org/10.1017/S2045796012000236

143. de Jong A. Development of the International Classification of Mental Health Care (ICMHC). *Acta Psychiatr Scand Suppl* 2000;**405**:8–13. https://doi.org/10.1111/j.0902-4441.2000.acp28–04.x

144. Calsyn RJ, Allen G, Morse GA, Smith R, Tempelhoff B. Can you trust self-report data provided by homeless mentally-ill individuals. *Evaluation Rev* 1993;**17**:353–66. https://doi.org/10.1177/0193841X9301700306

145. Goldberg RW, Seybolt DC, Lehman A. Reliable self-report of health service use by individuals with serious mental illness. *Psychiatr Serv* 2002;**53**:879–81. https://doi.org/10.1176/appi.ps.53.7.879

146. Patel A, Rendu A, Moran P, Leese M, Mann A, Knapp M. A comparison of two methods of collecting economic data in primary care. *Fam Pract* 2005;**22**:323–7. https://doi.org/10.1093/fampra/cmi027

147. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th edn. Washington, DC: American Psychiatric Association; 2000.

148. Mooney C, Duvall R. *Bootstrapping: A Non-parametric Approach to Statistical Inference*. Newbury Park, CA: Sage Publications; 1993. https://doi.org/10.4135/9781412983532

149. Department of Health. *NHS Reference Costs 2008 to 2009*. 2010. URL: http://data.gov.uk/dataset/nhs-reference-costs2008-09 (accessed 30 June 2017).

150. Bowers L, Flood C. Nurse staffing, bed numbers and the cost of acute psychiatric inpatient care in England. *J Psychiatr Ment Health Nurs* 2008;**15**:630–7. https://doi.org/10.1111/j.1365-2850.2008.01280.x

151. McCrone P, Johnson S, Thornicroft G. Predicting the costs of community care for individuals with severe mental illness in South London. *Schizophr Bull* 2001;**27**:653–60. https://doi.org/10.1093/oxfordjournals.schbul.a006904

152. Coleman JC, Paul GL. Relationship between staffing ratios and effectiveness of inpatient psychiatric units. *Psychiatr Serv* 2001;**52**:1374–9. https://doi.org/10.1176/appi.ps.52.10.1374

153. Department of Health. *Acute Adult In-patient Care: Policy Implementation Guide*. London: Department of Health; 2004.

154. Nolan P, Bradley E, Brimblecombe N. Disengaging from acute inpatient psychiatric care: a description of service users' experiences and views. *J Psychiatr Ment Health Nurs* 2011;**18**:359–67. https://doi.org/10.1111/j.1365-2850.2010.01675.x

155. Stenhouse RC. 'They all said you could come and speak to us': patients' expectations and experiences of help on an acute psychiatric inpatient ward. *J Psychiatr Ment Health Nurs* 2011;**18**:74–80. https://doi.org/10.1111/j.1365-2850.2010.01645.x

156. Care Quality Commission. *Patient Survey Report 2009: Mental Health Acute Inpatient Services Users Survey*. Newcastle upon Tyne: CQC; 2009.

157. Royal College of Psychiatrists. *Do the Right Things: How to Judge a Good Ward*. 2011. URL: www.rcpsych.ac.uk/pdf/OP79_forweb.pdf (accessed June 2012).

158. National Institute for Health and Care Excellence. *Service User Experience in Adult Mental Health: Improving the Experience of Care for People using Adult NHS Mental Health Services*. London: National Institute for Health and Care Excellence; 2011.

159. Star Wards. *Star Wards*. 2014. URL: www.starwards.org.uk/ (accessed January 2014).

160. Bowers L. Reasons for admission and their implications for the nature of acute inpatient psychiatric nursing. *J Psychiatr Ment Health Nurs* 2005;**12**:231–6. https://doi.org/10.1111/j.1365-2850.2004.00825.x

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

161. Cowls J, Hale S. It's the activity that counts: what clients value in psycho-educational groups. *Can J Occup Ther* 2005;**72**:176–82. https://doi.org/10.1177/000841740507200305

162. Clarke S. *Acute Inpatient Mental Health Care: Education, Training and Continuing Professional Development for All*. London: National Institute of Mental Health/Sainsbury Centre for Mental Health; 2004.

163. Greater London Authority. *English Indices of Deprivation 2010: A London Perspective*. London: Greater London Authority; 2010.

164. Greenfield T, Attkisson C. The UCSF Client Satisfaction Scales: II. The Service Satisfaction Scale-30. In Maruish ME, editor. *Psychological Testing: Treatment Planning and Outcome Assessment*. 3rd edn. Vol 3 Instruments for Adults. Mahwah, NJ: Lawrence Erlbaum Associates; 2004. pp. 813–37.

165. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophrenia Bulletin* 1987;**13**: 261–76. https://doi.org/10.1093/schbul/13.2.261

166. Honigfeld G, Gillis RD, Klett CJ. NOSIE-30: a treatment-sensitive ward behavior scale. *Psychol Rep* 1966;**19**:180–2. https://doi.org/10.2466/pr0.1966.19.1.180

167. Moos R. *Ward Atmosphere Scale*. Menlo Park, CA: Mind Garden Inc.; 1974.

168. PERCEIVE. *Doorways: Delivering Opportunities for Recovery*. URL: www.perceive.iop.kcl.ac.uk/doorways.html (accessed 10 January 2018).

169. Csipke E, Flach C, McCrone P, Rose D, Tilley J, Wykes T, Craig T. Inpatient care 50 years after the process of deinstitutionalisation. *Soc Psychiatry Psychiatr Epidemiol* 2014;**49**:665–71. https://doi.org/10.1007/s00127-013-0788-6

170. Care Quality Commission. *Monitoring the Mental Health Act in 2012/13*. Newcastle upon Tyne: CQC; 2014.

171. Armesto SG, Medeiros H, Wei L. *Information Availability for Measuring and Comparing Quality of Mental Health Care Across OECD Countries*. 2008. URL: www.oecd.org/els/health-systems/41243838.pdf (accessed June 2010).

172. Greengross P, Hollander D, Stanton R. Pressure on adult acute psychiatric beds. Results of a national questionnaire survey. *Psychiatric Bulletin* 2000;**24**:54–6.

173. Owen C, Tarantello C, Jones M, Tennant C. Violence and aggression in psychiatric units. *Psychiatr Serv* 1998;**49**:1452–7.

174. Powell RB, Hollander D, Tobiansky RI. Crisis in admission beds. A four-year survey of the bed state of Greater London's acute psychiatric units. *Br J Psychiatry* 1995;**167**:765–9. https://doi.org/10.1192/bjp.167.6.765

175. Hirsch SR, Platt S, Knights A, Weyman A. Shortening hospital stay for psychiatric care: effect on patients and their families. *Br Med J* 1979;**1**:442–6. https://doi.org/10.1136/bmj.1.6161.442

176. Glover G, Arts G, Babu KS. Crisis resolution/home treatment teams and psychiatric admission rates in England. *Br J Psychiatry* 2006;**189**:441–5. https://doi.org/10.1192/bjp.bp.105.020362

177. Jacobs R, Barrenho E. Impact of crisis resolution and home treatment teams on psychiatric admissions in England. *Br J Psychiatry* 2011;**199**:71–6. https://doi.org/10.1192/bjp.bp.110.079830

178. Inglis G, Baggaley M. Triage in mental health – a new model for acute in-patient psychiatry. *Psychiatr Bull* 2005;**29**:255–8. https://doi.org/10.1192/pb.29.7.255

179. Williams P, Csipke E, Rose D, Koeser L, McCrone P, Tulloch AD, *et al.* Efficacy of a triage system to reduce length of hospital stay. *Br J Psychiatry* 2014;**204**:480–5. https://doi.org/10.1192/bjp.bp.113.141051

180. Tulloch AD, Khondoker MR, Fearon P, David AS. Associations of homelessness and residential mobility with length of stay after acute psychiatric admission. *BMC Psychiatry* 2012;**12**:121. https://doi.org/10.1186/1471-244X-12-121

181. McCrone P. Predicting mental health service use: diagnosis based systems and alternatives. *J Ment Health* 1995;**4**:31–40. https://doi.org/10.1080/09638239550037820

182. Department of Health. *Reference Costs 2009–10 Publication*. London: Department of Health; 2011. URL: www.gov.uk/government/publications/nhs-reference-costs-2009-2010 (accessed 13 July 2017).

183. Department of Health. *Mental Health Act*. London: HMSO; 2007.

184. Csipke E, Williams P, Rose D, Koeser L, McCrone P, Wykes T, Craig T. Following the Francis report: investigating patient experience of mental health in-patient care. *Br J Psychiatry* 2016;**209**:35–9. https://doi.org/10.1192/bjp.bp.115.171124

185. Clegg N. *Making Mental Health a Priority*. London: Cabinet Office; 2014.

186. Wykes T, Csipke E, Williams P, Koese L, Nash S, Rose D, *et al.* Improving patient experiences of mental health inpatient care: a randomised controlled trial. *Psychol Med* 2018;**48**:488–97. https://doi.org/10.1017/S003329171700188X

187. Wykes T. Great expectations for participatory research: what have we achieved in the last ten years? *World Psychiatry* 2014;**13**:24–7. https://doi.org/10.1002/wps.20086

188. Francis R. *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. London: The Stationery Office; 2013.

189. Ulrich RS, Zimring C, Zhu X, DuBose J, Seo HB, Choi YS, *et al.* A review of the research literature on evidence-based healthcare design. *HERD* 2008;**1**:61–125. https://doi.org/10.1177/193758670800100306

190. Papoulias C, Csipke E, Rose D, McKellar S, Wykes T. The psychiatric ward as a therapeutic space: systematic review. *Br J Psychiatry* 2014;**205**:171–6. https://doi.org/10.1192/bjp.bp.114.144873

191. Gould D. *Service Users' Experiences of Recovery Under the Care Programme Approach 2008*. London: Mental Health Foundation and National Service Users Network; 2012.

192. Kendell RE, Chalmers JC, Platz C. Epidemiology of puerperal psychoses. *Br J Psychiatry* 1987;**150**:662–73. https://doi.org/10.1192/bjp.150.5.662

193. Terp IM, Mortensen PB. Post-partum psychoses. Clinical diagnoses and relative risk of admission after parturition. *Br J Psychiatry* 1998;**172**:521–6. https://doi.org/10.1192/bjp.172.6.521

194. Munk-Olsen T, Laursen TM, Pedersen CB, Mors O, Mortensen PB. New parents and mental disorders: a population-based register study. *JAMA* 2006;**296**:2582–9. https://doi.org/10.1001/jama.296.21.2582

195. Riordan D, Appleby L, Faragher B. Mother-infant interaction in post-partum women with schizophrenia and affective disorders. *Psychol Med* 1999;**29**:991–5. https://doi.org/10.1017/S0033291798007727

196. Salmon M, Abel K, Cordingley L, Friedman T, Appleby L. Clinical and parenting skills outcomes following joint mother-baby psychiatric admission. *Aust N Z J Psychiatry* 2003;**37**:556–62. https://doi.org/10.1046/j.1440-1614.2003.01253.x

197. Salmon MP, Abel K, Webb R, Warburton AL, Appleby L. A national audit of joint mother and baby admissions to UK psychiatric hospitals: an overview of findings. *Arch Womens Ment Health* 2004;**7**:65–70. https://doi.org/10.1007/s00737-003-0042-4

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

198. Laker C, Callard F, Flach C, Williams P, Sayer J, Wykes T. The challenge of change in acute mental health services: measuring staff perceptions of barriers to change and their relationship to job status and satisfaction using a new measure (VOCALISE). *Implement Sci* 2014;**9**:23. https://doi.org/10.1186/1748-5908-9-23

199. Kuokkanen L, Suominen T, Härkönen E, Kukkurainen ML, Doran D. Effects of organizational change on work-related empowerment, employee satisfaction, and motivation. *Nurs Adm Q* 2009;**33**:116–24. https://doi.org/10.1097/NAQ.0b013e3181a10c86

200. Burnard P, Morrison P, Phillips C. Job satisfaction amongst nurses in an interim secure forensic unit in Wales. *Aust N Z J Ment Health Nurs* 1999;**8**:9–18. https://doi.org/10.1046/j.1440-0979.1999.00125.x

201. Combs DR, Adams SD, Penn DL, Roberts D, Tiegreen J, Stem P. Social Cognition and Interaction Training (SCIT) for inpatients with schizophrenia spectrum disorders: preliminary findings. *Schizophr Res* 2007;**91**:112–16. https://doi.org/10.1016/j.schres.2006.12.010

202. Combs DR, Penn DL, Wicher M, Waldheter E. The Ambiguous Intentions Hostility Questionnaire (AIHQ): a new measure for evaluating hostile social-cognitive biases in paranoia. *Cogn Neuropsychiatry* 2007;**12**:128–43. https://doi.org/10.1080/13546800600787854

203. Parker S, Foley S, Walker P, Dark F. Improving the social cognitive deficits of schizophrenia: a community trial of Social Cognition and Interaction Training (SCIT). *Australas Psychiatry* 2013;**21**:346–51. https://doi.org/10.1177/1039856213486305

204. Roberts DL, Penn DL, Labate D, Margolis SA, Sterne A. Transportability and feasibility of Social Cognition And Interaction Training (SCIT) in community settings. *Behav Cogn Psychother* 2010;**38**:35–47. https://doi.org/10.1017/S1352465809990464

205. Roberts DL, Kleinlein P, Stevens B. An alternative to generating alternative interpretations in social cognitive therapy for psychosis. *Behav Cogn Psychother* 2012;**40**:491–5. https://doi.org/10.1017/S1352465812000082

206. Kerr SL, Neale JM. Emotion perception in schizophrenia: specific deficit or further evidence of generalized poor performance? *J Abnorm Psychol* 1993;**102**:312–18. https://doi.org/10.1037/0021-843X.102.2.312

207. Corcoran R, Mercer G, Frith CD. Schizophrenia, symptomatology and social inference – investigating theory of mind in people with schizophrenia. *Schizophr Res* 1995;**17**:5–13. https://doi.org/10.1016/0920-9964(95)00024-G

208. Beck AT, Baruch E, Balter JM, Steer RA, Warman DM. A new instrument for measuring insight: the Beck Cognitive Insight Scale. *Schizophr Res* 2004;**68**:319–29. https://doi.org/10.1016/S0920-9964(03)00189-0

209. Gledhill A, Lobban F, Sellwood W. Group CBT for people with schizophrenia: a preliminary evaluation. *Behav Cogn Psychoth* 1998;**26**:63–77. https://doi.org/10.1017/S1352465898000071

210. Taylor R, Cella M, Csipke E, Heriot-Maitland C, Gibbs C, Wykes T. Tackling social cognition in schizophrenia: a randomized feasibility trial. *Behav Cogn Psychother* 2016;**44**:306–17. https://doi.org/10.1017/S1352465815000284

211. Horan WP, Kern RS, Shokat-Fadai K, Sergi MJ, Wynn JK, Green MF. Social cognitive skills training in schizophrenia: An initial efficacy study of stabilized outpatients. *Schizophr Res* 2009;**107**:47–54. https://doi.org/10.1016/j.schres.2008.09.006

212. Mdege ND, Man MS, Taylor Nee Brown CA, Torgerson DJ. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *J Clin Epidemiol* 2011;**64**:936–48. https://doi.org/10.1016/j.jclinepi.2010.12.003

213. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007;**28**:182–91. https://doi.org/10.1016/j.cct.2006.05.007

214. Kotz D, Spigt M, Arts IC, Crutzen R, Viechtbauer W. Use of the stepped wedge design cannot be recommended: a critical appraisal and comparison with the classic cluster randomized controlled trial design. *J Clin Epidemiol* 2012;**65**:1249–52. https://doi.org/10.1016/j.jclinepi.2012.06.004

215. Sabes-Figuera R, McCrone P, Csipke E, Craig TK, Rose D, Sharma B, Wykes T. Predicting psychiatric inpatient costs. *Soc Psychiatry Psychiatr Epidemiol* 2016;**51**:303–8. https://doi.org/10.1007/s00127-015-1152-9

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 1 Design in mind

Background

The contribution of the built environment of health-care facilities to treatment outcomes is increasingly acknowledged.[189] However, the relationship of psychiatric ward design to patient outcomes, as well as to the overall experience of treatment, is a particularly under-researched area. This is rather surprising as, in UK policy, evidence-based design has become a cornerstone for the construction of new facilities under the Private Finance Initiative.[189] Furthermore, current UK policy guidelines for the design of psychiatric wards stipulate that 'service users . . . should have an early and continuing opportunity to contribute to the new service design and participate in the planning process'.[189] However, no study thus far has been based on such a contribution. Although some have elicited service user and front-line staff perceptions of design, no study to date has worked with users to create a robust tool for the evaluation of that environment. The present study grew out of WP1 and applied participatory methods with the aim of eliciting service user and staff priorities concerning psychiatric ward design in order to provide better tools for the future.

Our study had three stages.

1. Systematic review to evaluate the evidence on the impact of the design of psychiatric environments on treatment outcomes and staff well-being (March 2012).
2. Measure development using a model of 'stakeholder involvement' and participatory methodology (April–August 2012 and July–August 2013).
3. Qualitative study of service user experience of ward environment using photo elicitation (June–August 2012).

Stage one: literature review

Evidence on the contribution of psychiatric design to clinical outcomes is limited. A total of 23 studies were identified (20 quantitative and three qualitative). Only three studies measured the impact of facility design on symptom reduction. The majority of the remaining studies explored the relationship between the physical environment, social behaviour and treatment or job satisfaction. Several studies suggested that private spaces and homely features might encourage social interaction in service users. However, staff and service users often had conflicting responses to the same features. In addition, many of the findings are inconclusive, as several studies used small samples, not all involved control groups, and many did not account for confounding variables. Moreover, no study developed a user-generated measure for the assessment of ward design.

Stage two: measure development and pilot study[190]

Setting

The study took place in an inner city NHS mental health trust. Four 18-bed adult acute wards participated. Data collection took place in April–May 2012 (measure development), June–August 2012 and July–August 2013 (pilot study). Inclusion criteria for service user participants were that they had been present on the ward for 3 days prior to data collection and were able to provide informed consent. Inclusion criteria for staff were that they had been working on the ward for at least 4 weeks. Demographic data were collected by self-report from both groups, while clinical data from the service users were collected from NHS records.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Measure generation

Questionnaires were produced through a participatory methodology.[17] Two service user researchers reviewed the literature and produced topic guides which formed the basis of interviews with staff and service users on the participating wards. The interviews were audio-taped, transcribed and thematically analysed by the service user researchers using NVivo 9 (QSR International, Warrington, UK). Key themes identified were then used to develop questionnaire items.

Analysis

The internal reliability of the measure was assessed using Cronbach's alpha[83] with data from the questionnaire phase of the study. Test–retest reliability was carried out with participants who completed the measures on two occasions. Test–retest reliability was estimated for each item using weighted Kappa (quadratic weighting to emphasise the importance of minor differences between time points) and alpha statistics. Based on previous research, we expected that service users who were involuntarily admitted and those from BME groups would show more negative perceptions of the ward environment.

We assessed any differences with one-way ANOVA on views on the ward design.

Analyses were carried out using Stata 12.1 and Statistical Product and Service Solutions (SPSS) version 20 (IBM Corporation, Armonk, NY, USA).

Results

Sample characteristics

Ten service users and 10 nurses were recruited into the measure generation phase of the study. Fifty-three service users and 61 staff were recruited into the piloting phase of the study. A total of 50% of service users in the interview phase and 58.5% in the piloting phase were from BME groups, and the most frequent diagnosis was psychosis spectrum disorder (40% and 46%, respectively). The majority of staff participants were also from BME groups in both the interview and the piloting phases (70% and 56.8%, respectively), and all the participants in the interview phase and 83.7% in the piloting phase were qualified nurses.

Measure generation

Two measures were generated: the service user questionnaire included 20 items, while the staff questionnaire included 22 items. The items were rated via a six-point Likert scale. In addition, free-text spaces were included to provide qualitative data.

Service user measure

Feasibility and acceptability

All 34 participants found the questionnaire easy to understand and complete, but 11% ($n = 4$) felt that it was too short. The Flesch reading ease score was 84.6 (for 8- to 9-year-olds).

Psychometric testing

Following initial reliability checks, it was decided that items Q8 and Q13 should be removed owing to underperformance in terms of kappa. The final measure comprised 18 items. Internal consistency analyses indicated a Cronbach alpha of 0.91 with an overall total scale ICC of 0.67 indicating moderate test–retest reliability.

Exploratory analyses

Service users from a white background were more negative about the ward design than BME service users ($F = 5.30$; $p = 0.026$). Those service users diagnosed with a psychosis spectrum disorder had lower scores; that is, they viewed the ward design more positively than those services users with other diagnoses ($F = 5.20$; $p = 0.009$).

Staff measure

Feasibility and acceptability
All 53 respondents found the questionnaire easy to understand and complete and 96% felt that the length was appropriate. The Flesch readability score was 70.4 (for 12-year-olds).

Psychometric testing
All items were found to be reliable and, therefore, no items were dropped.

The average item–item correlation was 0.42 and the alpha statistic for the overall standardised scale was 0.94.

Exploratory analyses
Following univariate analyses, we found that staff participants from a white background were similar to service users in their more negative assessment of ward design. No other group differences were found.

Stage three: photo elicitation

Sample
Service users on two out of the four acute wards involved in the measure development participated in the study. Inclusion criteria for service user participants were that they had been present on the ward for 3 days prior to data collection and were able to provide informed consent.

Method
Participants were given a point-and-click digital camera and asked to take two photographs: one depicting the best and one depicting the worst part of the ward. Two service user researchers briefly interviewed participants on their views about the aspects of the ward represented on the photographs. Interviews were audio-taped and transcribed verbatim. The service user researchers then thematically analysed both photographs and transcripts using NVivo 9.

Preliminary results
A total of 36 service users participated in the study. Key themes included concerns with hygiene and maintenance, positive evaluation of items representing leisure activities (TV, ping pong table), appreciation of privacy.

Discussion
Both measures showed good or reasonable test–test reliability and internal validity and this should be further strengthened in a larger study. Service users may change their view as they experience the ward for longer periods of time and, therefore, this in particular needs further exploration to detect relationships with other measures (e.g. length of admission, previous admissions, mental state, etc.). The lack of observable differences according to legal status and the positive ratings among BME service users and users with psychosis spectrum disorders were not anticipated and merit further study. UK studies indicate that BME service users are more negative about their experience of mental health services and about inpatient care in particular.[191] However, it is possible that the negative views reported in these studies refer to other aspects of the treatment environment (e.g. relations with staff, continuity of care, violence in the ward). Photo elicitation data suggest that service users prioritise privacy and this may be more important than other design considerations.

Concluding remarks
This study has demonstrated that a robust tool for the evaluation of inpatient psychiatric facilities can be generated through participatory methods, making use of staff and service user perceptions of ward design.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 2 Perinatal VOICE

Background

Women's risk of developing mental health problems is greatest in the postnatal period compared with any other time in her life, particularly so for first time mothers[192,193] and for women with a history of bipolar disorder.[194] Severe mental illness in a mother can have many adverse effects on their ability to physically care for their infants, interact with their infants in a positive way and form meaningful attachments.[195] Although efforts are made to keep treatment community based, some women may need acute inpatient care. In the majority of cases the mother is admitted to regular inpatient services, leading to the separation of mother and baby. However, this separation can be upsetting for both mother and child and may inhibit the formation of their relationship and bonding. It also might induce guilt in the mother who cannot care for her child, which may in turn exacerbate her illness. In an attempt to overcome this, special mother–baby units exist in countries like the UK, France and Australia, and have existed since the 1940s. These units have been found to be beneficial to both mothers and infants and lead to earlier discharge.[196,197]

In the UK, mothers with young babies suffering from acute episodes have three potential types of treatment: in the community, in regular acute wards and on dedicated mother and baby units. Although there is evidence that discharge is earlier on these wards, currently there is no standard method to investigate what the women's perceptions and experiences of the different models of care are. It is also not known if measures assessing perceptions and satisfaction with current measures cover the needs specific to those using acute services in the perinatal period. Mothers with babies might have needs and preferences that are very different from other acute patients, as they must consider what their child's needs and requirements are, as well as those of themselves.

The PERCEIVE project developed the VOICE measure – a measure of perceptions of inpatient care. Colleagues specialising in women's mental health secured funding for a programme grant designed to evaluate treatment efficacy for women in the perinatal period. Specifically, one WP in their programme included comparing the different types of services offered. They approached our team about using the VOICE measure as one of the assessments. After some discussion it was agreed that we would develop a measure tailored for use with women in the perinatal period as we had the expertise within our team and this would be a value added study within the scope of the LIAISE WP. We followed a methodology similar to the VOICE study.

Aim

To determine if mothers (with children aged < 1 year) in need of acute psychiatric treatment have additional or different concerns or value different aspects of mental health services than those who do not and to develop a questionnaire that addresses these needs.

Method

Measure generation phase

Inclusion criteria for our measure generation sample were that participants had been admitted to an acute ward, a mother and baby unit or were deemed to be in an acute episode by their community treatment team and were being treated at home. Furthermore, they were mothers with a child aged < 1 year, or pregnant at the time. Finally, they were able to provide informed consent.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

VOICE itself was used as a basis for the topic guide for the semistructured interview, along with input from the women's mental health team, which also included service user researchers and advisors. Interviews were carried out because focus groups were not feasible. The material from the interviews was audio-taped and subject to thematic analysis.

Quality assurance phase

Inclusion criteria for the feedback group were that they had used perinatal mental health services in the previous 5 years and they were willing to provide informed consent.

The draft of the questionnaire was taken to the feedback group and it was also presented to the women's mental health team and feedback was sought.

Results

Measure generation phase

Sixteen women agreed to be interviewed about their perceptions of their care and needs. Their average age was 34 years (SD 3.6 years) and they were evenly split between having a white or BME background. Two of the women (12%) were pregnant at the time of the interview. Four women were recruited from a standard acute inpatient ward (25%) and the remaining individuals were on mother and baby units. Half of participants had in the past experienced other types of care. No one who was being treated in the community was recruited.

Seven participants were diagnosed with depression (44%). Three individuals (19%) had multiple diagnoses, two (12%) had a form of psychosis, one (6%) had a diagnosis of bipolar disorder and the remaining three had no diagnoses.

Omitted themes

The topic of respect for religious or ethic backgrounds was not raised by any participant. Medication and its potential side effects were also not raised. Medication was only discussed in the context of the staff not taking enough time to provide therapeutic interventions and being seen as the dispensers of medicine.

Shared themes

An analysis of the interviews found that many themes were shared with those involved in the VOICE focus groups. Having a say in their care and treatment, having positive interactions with ward staff and feeling safe were all raised by the women in our add-on study. Keeping in touch with family was also felt to be important.

New themes

Ward facilities were discussed, especially in the context of mother and baby units, particularly their cleanliness, accessibility and equipment. This was entwined with discussions about how the nurses on the ward interact with them and how they care for their babies. The mother and baby units had strict rules about how and where babies are changed, fed or put down to sleep. First time mothers of newborns expressed that sometimes they would appreciate more guidance in how to best care for their babies but not all nurses were able to provide that. They also felt frustrated when nurses provided conflicting advice. Mothers who felt more confident with their babies complained that the strict rules interrupted babies' established schedules or simply made things more complicated or was not the parenting style the mother had chosen to adopt.

The issue of breastfeeding was also raised. Some women were uncomfortable breastfeeding when male staff or visitors could see them and their complaints had gone unnoticed. The lack of privacy arising from staff checking on them several times overnight by looking through the window in their door also came up in this context. Breastfeeding was also raised from another perspective. Some mothers had chosen to do it

exclusively and were not happy when the nurses would give the baby a bottle to quiet them down rather than tell the mother that the baby was crying.

Support from other mothers going through something similar was also raised as something to be valued, especially in the mother and baby units. Focus group members felt that they ought to be encouraged by the staff/ward culture, although they recognised that this was often dependent on who was admitted to the ward at the time and not entirely dependent on nurses. It was suggested that inviting women who had previously been on the ward but were now doing better would be a good way to 'show there is a light at the end of the tunnel'.

For women on the acute wards, not being with their babies was clearly something they were concerned about. The lack of communication about how their baby was doing was particularly important for those women whose babies were in foster care and not with families. Not knowing how their babies were and how, and if, they were being taken care of caused great anxiety. That sometimes staff would not even acknowledge (or not know) that they were separated from their babies was seen in a negative light. This topic was raised on the mother and baby units as well and many stated they would have refused any treatment that meant separation. Social services were seen as helpful in some cases, in arranging housing for example, which expedited discharge planning. On the other hand, most women also worried about their involvement and the unexpected decisions they could make, especially if they had seen an example of that on their ward.

A measure was created out of this analysis and a bank of 38 questions were created.

Measure review and improvement phase

Six participants agreed to feed back on the draft questions. The mean age was 37 years (SD 9.4 years) and all were white. Three reported they had had depression, two reported bipolar and the remaining one did not wish to disclose their diagnosis. All participants had multiple types of care and half of them all three types.

One of the key things raised was the lack of questions regarding medication side effects. It was suggested we add a question despite it not coming up in the interviews and the women's health team agreed with this. They also helped whittle down those questions which were too similar to others and chose the best one. Furthermore, they helped with the phrasing of certain questions.

We also discussed the draft measure with the women's health research team as it was designed to be eventually used by them. Originally conceptualised as consisting of three separate measures for each type of service (mother and baby unit, Acute, Community) they decided that we should try to phrase it in a way that could apply to all settings. After much deliberation, we made only one question optional (which was related to breastfeeding, which not all women may be doing). We currently have a measure of 27 items.

Psychometric testing

Action for Postpartum Psychosis (Derby, UK) have been helping with publicising this psychometric study which is on a SurveyMonkey® (Palo Alto, CA, USA) link in order to provide the participants for this phase which will be carried out within the next months. We hope to collect 40 participants completing the measure twice with a 1-week period. The planned analyses include test–retest reliability and a factor analysis.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 3 VOCALISE: barriers to change[198]

Background

The findings of WP1 (LIAISE) and the challenges of setting up WP3 (DOORWAYS) led to this add-on study. Organisational change in the NHS is often imposed via a top-down approach, which may not take into consideration the views of nursing staff. Indeed, change process issues such as poor involvement in planning, implementation and control of the project have been highlighted as potential barriers to success.[51] The emotional response of staff to proposed changes may therefore be an important barrier to the success of implementation.

We propose a model in which negative staff perceptions of the daily pressures of working on an acute ward and negative staff perceptions of barriers to change lead to burnout and low job satisfaction, and this affects the quality of care delivered. In nursing studies, there has been limited focus on assessing how social, emotional and psychological barriers affect the implementation of new innovations. The relationship between psychological well-being and the acute ward setting in relation to planned changes may, in nursing, be distinct from other professions because nurses remain on the ward for the duration of their shift and are often expected to play a key role in delivering changes. There are several measures developed in health care that focus on the uptake of evidence-based practice in clinical areas by mental health providers and nurses.[64–71] However, there are no measures that focus on general changes in a mental health setting.

What is needed in the literature on acute care is a psychometrically robust, brief, self-report measure reflecting staff experiences of delivering/innovating. A measure of this type would allow a clear assessment of changes to inpatient care following specific interventions to improve the environment and therapy provided. This add-on study was designed to generate such a measure.

Methods

Study design

This study was designed to develop and test this self report measure. The measure was generated through a process of stakeholder (staff) involvement.[19–21] This approach directly involved nursing staff in order to ensure measures that captured an accurate picture of an acute care ward. The process of measure development used for VOCALISE (see *Appendix 10* for full measure) was an adaption of that used for VOTE.

Mapping out the dimensions of inpatient care

In this initial phase, stakeholder groups were held to map out the dimensions of inpatient care as a topic guide. In developing VOCALISE, an interview schedule was refined from the literature and the focus group data as part of a consultation exercise between two nurses, one who was a senior nurse from the trust under study, an expert in qualitative methods/service user researcher and a clinical psychologist.

Measure generation

Sample

Nursing staff from all grades were asked to participate in the development of VOCALISE. The focus groups showed that staff were aware of the nursing hierarchy when discussing sensitive information. Therefore, individual interviews were held to collect data. Finally, two expert panels met to discuss the design of the new measure and to inform the 'instructions for use'. Changes were made to the items and to the layout

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

of the measure on the basis of this feedback. The draft measure was finally presented back to the original reference group for their comments.

Psychometric analysis

The final measure was intended as a self-report tool so studies of *feasibility* and *acceptability* were conducted to evaluate the burden of administering and completing the measure. In the *feasibility study*, 40 participants completed the measure including two additional questions assessing whether the measure was easy to complete and understand. *Acceptability* was assessed using two additional items collected from 115 participants. Test–retest reliability was assessed with participants who completed the measures on two occasions with an interval of 6–10 days. Internal consistency was assessed using Cronbach's alpha.

Whether or not the new measures truly reflected the experiences of those providing care in the acute inpatient services (face validity) and covered the full spectrum of views (content validity) were explored as a result of the participatory methodology during the instrument development phase. We hypothesised that staff with positive perceptions of barriers to change would also have high levels of job satisfaction.[199] This was tested using two groups: staff with positive perceptions of job satisfaction and staff with negative perceptions of job satisfaction.

Exploratory analyses

Do demographic characteristics affect perceptions of barriers to change?

A random-effects regression model was run to provide a clear picture of which demographic variables are true predictors of staff perceptions of barriers to change, controlling for all demographic variables, to test the following hypothesis derived from the literature: staff in managerial roles who (1) are from a younger age group and (2) are degree educated will have more positive perceptions of barriers to change.

Results

Mapping out the dimensions of inpatient care

As the concept of 'perceptions of barriers to change' is abstract, the reference group recommended that the interviewee be asked to consider a scenario in practice in which a significant change to clinical practice had occurred. A topic guide for interviews was created.

Measure generation

Sample

The sample involved in the development of VOTE included inpatient nursing staff from all grades within one London mental health trust. The catchment area of this trust is large and covers four inner and outer London boroughs. In total, 376 individuals were involved at the various stages (*Table 39*).

Interviews

Thematic analysis produced seven domains: 'communication', 'generation of ideas', 'outcomes of changes', 'resistance', 'strategy', 'support and monitoring' and 'team dynamic'. Themes with fewer than 20 references were not included. Items were constructed around perceptions of the team/ward and perceptions of the self in relation to change.

Expert panels

Those who participated in the expert validation process confirmed that the content of the measure was relevant and that the content domain had been widely explored. There were no changes to the measure as a result of the final consultation.

TABLE 39 VOCALISE measure generation group

	VOCALISE measure generation group ($n = 40$)
Staff, *n* (%)	
Health-care assistants	13 (32.5)
Staff nurses	14 (35)
Clinical charge nurses	9 (22.5)
Team leaders	3 (7.5)
Ethnic group, *n* (%)	
White British/other	16 (40)
Black/minority ethnic group	24 (60)
Gender, *n* (%)	
Male	21 (52.5)
Female	19 (47.5)
Age (years), *n* (%)	
Mean (SD)	44.44 (9.07)
Range	28–64

Reproduced from Laker *et al.*[198] © 2014 Laker *et al.*; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

Psychometric assessment

Feasibility and acceptability

The feasibility study (group one, $n = 40$) showed that VOCALISE was easy to complete (94% agreed) and easy to understand (100% agreed). Changes were made to the wording of some items. Those that seemed to be addressing more than one issue were simplified, while ensuring that these simplifications did not lead to the omission of any prevalent themes. Those items with poor or loaded phrasing were rephrased using more neutral language to give staff scope to agree or disagree. After these modifications all 23 items still remained.

The acceptability study showed that 73% thought the length of the questionnaire was about right. 20% enjoyed filling out the questionnaire, and 67.8% had neutral feelings. Finally, 107 (93%) did not find completing the items upsetting.

Total scores were calculated by totalling all items with no missing data. Negatively phrased items were reverse scored so that higher scores indicated a more negative perception of the ward. When comparisons to other measures were made, the same rule was applied to their scoring (i.e. high scores indicate poorer satisfaction).

Reliability

Test–retest reliability was assessed using data from group two ($n = 42$). Four items were unreliable with a kappa maximum < 0.39 and were dropped from the scale, leaving eight items with fair reliability (0.39–0.49), seven items with moderate reliability (0.50–0.56) and three items with substantial reliability (0.61–0.71).

Concordance between the total scores was good (total score, $\rho = 0.76$). However, a paired *t*-test showed that there was a significant difference between the two time points ($t = -2.10$; $p = 0.04$) (mean difference = –2, 95% CI –3.93 to –0.07). Test–retest reliability was therefore assessed according to staff group revealing that staff in direct care roles ($n = 26$) were likely to change their scores ($t = -2.91$;

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

$p = 0.008$) (mean difference = –3.12, 95% CI –5.32 to –0.91). The scores of those in managerial roles were stable ($t = 0.35$; $p = 0.73$) (mean difference = 0.64, 95% CI –3.36 to 4.64; $n = 11$). The test–retest reliability of the 'I' statements and general statements was also assessed showing that over time, the general statements ($t = –1.88$; $p = 0.07$) (mean difference = –1.24, 95% CI –2.59 to 0.10; $n = 37$) were more likely to change than the 'I' statements ($t = –1.50$; $p = 0.27$) (mean difference = 0.14, 95% CI –1.82 to 0.27; $n = 40$).

The final Cronbach's alpha after dropping a single item was 0.75 leaving 18 remaining items.

Face and content validity

A high level of staff involvement throughout the process of measure development ensured good face and content validity. This was achieved because staff participants provided feedback on the content of the themes arising from the qualitative data and on the language used in the item generation phase. Staff agreed that the results did capture what they had reported. The use of a flexible topic guide/interview schedule maximised exploration of the construct under study and minimised omissions in the data set.

Criterion validity

Those with negative perceptions of the barriers to change also had poor job satisfaction after controlling for age and occupational status (coefficient 10.43, SE 1.97; 95% CI 6.58 to 14.30, $n = 101$, eight wards; $p = 0.001$). The predicted mean VOCALISE score in the high job satisfaction group was 56.76 and the predicted mean score in low job satisfaction group was 67.20 (*Table 40*).

TABLE 40 Mean VOCALISE scores

	Mean VOCALISE scores (SD)
Staff	
Whole sample	68.54 (12.77)
Health-care assistants	64.34 (11.35)
Qualified nurses	62.35 (11.34)
Clinical charge nurses	58.26 (11.45)
Team leaders	56 (8.65)
Ethnic group	
White British/other	65.27 (10.17)
BME	61.00 (11.65)
Gender	
Male	62.23 (10.62)
Female	62.25 (12.03)
Age	
≥ 39 years	59.17 (13.37)
< 39 years	64.12 (9.06)
Country of origin	
UK	66.05 (10.92)
NON UK	60.07 (11.29)
Length of employment	
≥ 3 years	61.39 (12.06)
< 3 years	61.37 (10.64)

Reproduced from Laker *et al.*[198] © 2014 Laker *et al.*; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

Do demographic characteristics affect perceptions of barriers to change?

Two demographic factors, age (coefficient –5.67, SE 2.57, 95% CI –10.70 to –0.64; $p = 0.03$) and occupational seniority, significantly affected staff perceptions of barriers to change when assessed using the total score.

Discussion

Using a participatory methodology, we have developed a staff generated, self-report measure of perceptions of barriers to change in acute care – VOCALISE. It demonstrates promising psychometric properties. The internal consistency and criterion validity are high and test–retest data show stability over time. The full involvement of nursing staff throughout the development of the measure has ensured that VOCALISE has good face and content validity and is accessible to the intended recipients.

In this study, staff had slightly more positive perceptions of job satisfaction than staff in other studies.[101–103,200] Staff with negative perceptions of barriers to change also had poor job satisfaction. In previous studies, work stressors have differed between groups including occupational status and gender. Staff with higher organisational status[52,54] view changes more favourably than those in more junior positions. Furthermore, job satisfaction has been linked to staff who express more open views about new changes.[199] These findings were replicated with VOCALISE.

VOCALISE scores were most negative overall in the direct care staff groups (staff nurses/nursing assistants) compared with more positive scores in those who occupied more managerial roles (clinical charge nurses and team leaders). Both nursing assistants and staff nurses spend the most time in direct client contact. Therefore, their more negative perceptions might be one of the inevitable aspects of delivering therapeutic acute care for patients with severe mental illness at their most distressed. Furthermore, staff in leadership roles are more likely to be involved in the planning stages of new changes and, therefore, have an increased sense of control and responsibility over them, while those in direct care roles are more likely to be involved in delivering changes. These important issues, revealed through stakeholder involvement, require further exploration to discover key drivers for these perceptions which might then be subjected to management and other interventions.

The benefits of the participatory model used to develop VOCALISE are visible in both the breadth of the construct investigated and the rich content of the items. VOCALISE captures the complexity of an abstract topic by addressing the interaction between perceptions of the self and the organisation as well as perceptions of the team dynamic in relation to changes. Barriers arising from ward/environmental factors as well as from social factors are addressed. The content of some items is of interest because they exemplify feelings of organisational unfairness, perhaps as a result of poor consultation around new changes which may have been missed using a non-participatory method.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 4 Social cognition and interaction training

Social cognition and interaction training for forensic inpatients: a randomised pilot study

The DOORWAYS project involved SCIT as one of the activities that ward staff were trained to deliver. The effects in DOORWAYS were estimated through the effects of increased activities on service user and staff perceptions of the therapeutic environment. This study is an add-on investigation of the effects of SCIT on the individuals who received it. The intervention is slightly longer than the intervention on acute wards as it is being tested in a forensic setting. The forensic ward population was initially excluded from the study and had not yet benefited from the available interventions. Therefore, the present study aimed to evaluate the feasibility, acceptability and potential benefits of SCIT for the first time in the UK within forensic wards. The SCIT sessions consisted of a modified version of the social cognition groups already running in the existing DOORWAYS project. Forensic wards have the advantage over general acute wards of having a slower patient turnover and SCIT complements the rehabilitation focus of such treatment settings. SCIT has been implemented in the USA in a similar setting and was found to be beneficial in improving social cognition, social relationships, cognitive flexibility and reduced aggression.[201] In different settings SCIT has also been shown to improve emotion perception, theory of mind (ToM), cognitive flexibility, tolerance of ambiguity and need for closure.[129,202–205] However, many of these studies have been carried out within outpatient settings and are limited by small samples and uncontrolled pre–post designs.

We hypothesised that those who completed the intervention would show greater improvement in emotion perception, ToM, cognitive flexibility and less hostile attributional biases for ambiguous situations.[202] Lastly, we expected SCIT would be positively received by participants and that they would meet their social goals or feel they were able to meet these goals as a result of the intervention to provide some indication of the functional significance of the treatment.

Method

Design and procedure

This study is a cluster randomised single-blind controlled trial with crossover design. Four forensic inpatient wards took part with two wards randomised to receive SCIT and two wards randomised to TAU. Following the intervention phase, those randomised to TAU were offered SCIT. All participants were assessed twice: at intake and after the intervention.

Participants

Participants were recruited from adult male forensic psychiatric wards in collaboration with a single inner city London NHS trust. Inclusion criteria were: aged 18–65 years with sufficient English to consent to research, complete assessments as well as the intervention.

Intervention

Therapy consisted of a 16-session group facilitated twice a week with a 45-minute session. Emotion perception was the focus of the first six sessions and the following seven sessions were assigned to ToM, attributional style and underlying processes such as jumping to conclusions and tolerating ambiguity. Integration, which focuses on generalising skills to participants' real-life social situations, was covered in the final two sessions. All sessions end with the setting of homework that is reviewed at the beginning of the following session.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Measures

Feasibility
Participant group attendance and attrition was recorded in addition to feedback from each group facilitator about how the intervention had been received and whether or not there were any challenges in implementation.

Symptoms
The PANSS,[165] as used in WP3.

Social cognition
Emotion perception was measured using the Face Emotion Identification Task[206] in which participants are shown 19 individual photographs of faces and have to identify which basic emotion is being shown.

Theory of mind
The Hinting Task[207] was carried out. Ten individual short stories describing an interaction between two people are read to the participant with each ending with one of the characters dropping a hint. Participants are required to identify what the person really meant. Initial incorrect answers are followed by a single prompt with more information.

Attributional style
The Ambiguous Intentions Hostility Questionnaire (AIHQ[202]) consists of 15 short vignettes describing negative interpersonal situations that vary in intentionality. Only the five items from the ambiguous scale were used as these are more sensitive to social cognitive biases.[202]

Cognitive flexibility and need for closure
These domains were assessed using the Beck Cognitive Insight Scale,[208] a 15-item self-report questionnaire designed to evaluate self-reflectiveness and overconfidence in interpretations of anomalous experiences.

Acceptability
During the first two group sessions, participants were asked to identify one social goal that they would like to achieve or work towards. In their final session participants were asked whether or not they had achieved their goal or felt they were able to use the skills learned from the intervention in order to do so. Participants were also asked to rate whether or not they felt their confidence had increased. All participants completed a satisfaction questionnaire.[209]

Analysis

Feasibility
Attendance of 75% sessions was necessary to receive an adequate treatment dose and those who completed < 12 sessions were excluded from the analyses of intervention benefit.

Intervention benefit
Between-group differences were assessed using univariate analysis of covariance (ANCOVA) for each outcome. For this analysis, post-treatment scores were used as dependent variables, treatment as a fixed factor and baseline scores as covariates together with other baseline scores significantly different between the two groups at baseline. Treatment effects were estimated using partial eta squared (η^2).

The percentages of participants who were able to achieve their social goal or felt they could now achieve this goal and reported an increase in confidence following SCIT were calculated.

Acceptability
All participants completed a satisfaction questionnaire at the end of treatment.

Results

Participants

Of those on the wards, 45 were eligible and 36 gave written informed consent. Fifteen were randomised to TAU and 21 to SCIT (*Table 41*). There were no differences across the four psychiatric wards. The majority of participants were of African Caribbean origin, which is reflective of the population served by the NHS trust (see *Table 41*).

Feasibility

Attrition and attendance

The SCIT participants attended a mean of 14.4 sessions (SD 1.4 sessions) out of a total of 16 sessions, which was an average attendance rate of 90%.

Intervention benefit

Table 42 summarises the descriptive data and the ANCOVA results for the social cognition domains and symptom dimensions considered. Participants in the SCIT group significantly improve on facial affect recognition than participants in the TAU group (F-test = 7.5, $\eta^2 = .24$; $p < 0.001$). No further significant differences were found.

TABLE 41 Demographic and clinical characteristics

	SCIT ($n = 21$), mean (SD)	TAU ($n = 15$), mean (SD)	F (p)
Age (years)	40.7 (10.3)	39.2 (10.6)	0.1 (0.7)
Age first admission	25.2 (4.9)	22.3 (8.1)	1.8 (0.2)
Education (years)	9.8 (2.5)	10.6 (2.1)	0.9 (0.3)
Ethnicity			
Afro-Caribbean, %	81	73	$\chi^2 = 5.3$ (0.32)
Other, %	19	27	
PANSS positive	10.3 (2.8)	10.8 (4.9)	0.1 (0.7)
PANSS negative	9.7 (2)	10.4 (3.1)	0.4 (0.5)
PANSS general	24 (7.1)	23.4 (7.5)	0.4 (0.8)

Source: Taylor R, Cella M, Csipke E, Heriot-Maitland C, Gibbs C, Wykes T, Tackling social cognition in schizophrenia: a randomized feasibility trial, *Behavioural and Cognitive Psychotherapy,* vol. 44, issue 3, pp. 306–17,[210] reproduced with permission.

TABLE 42 The ANCOVA controlling for baseline and PANSS negative *$p < 0.05$; **$p < 0.001$; ***$p < 0.0001$

	Pre treatment		Post treatment		F-test	Effect size
	SCIT ($n = 21$), mean (SD)	TAU ($n = 15$), mean (SD)	SCIT ($n = 16$), mean (SD)	TAU ($n = 11$), mean (SD)	Group effect	Between group (η^2)
AIHQ hostility scale	1.41 (.55)	1.61 (.69)	1.5 (.76)	1.4 (.56)	.41	.06
Facial affect recognition	8.69 (4.5)	9.45 (3.7)	11.06 (3.1)	8.55 (3.3)	7.5**	.24
Hinting task	13.44 (4.1)	13.9 (2.8)	14.3 (2.2)	14.8 (2.6)	0.18	.008
Beck self-certainty scale	5.63 (6.2)	3.6 (4.1)	4.94 (4.9)	5.3 (5.2)	1.3	.053
PANSS positive	9.56 (1.9)	10.4 (4.1)	8.81 (2.2)	10.4 (3.2)	2.9	.11
PANSS negative	8.44 (1.4)	11.64 (3.9)	9.3 (1.6)	10.3 (4.1)	2.3	.089

Source: Taylor R, Cella M, Csipke E, Heriot-Maitland C, Gibbs C, Wykes T, Tackling social cognition in schizophrenia: a randomized feasibility trial, *Behavioural and Cognitive Psychotherapy,* vol. 44, issue 3, pp. 306–17,[210] reproduced with permission.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Acceptability

Participant social goal achievement, confidence and satisfaction

A total of 57% of participants strongly agreed that they had achieved their goal and 36% partly agreed. Similarly, 42% strongly agreed and 36% partly agreed they could now achieve other social goals as a result of skills acquired during SCIT. A total of 71% strongly agreed and 21% partly agreed that their confidence had improved following the intervention.

All participants who completed the intervention stated they had enjoyed the group and found it beneficial, with 83% stating they would like to be involved in a similar group in the future.

Discussion

The modified SCIT intervention seems feasible with 16 out of 21 randomised to SCIT completing at least 12 out of the 16 sessions. Among completers, there was an attendance rate of 90%, which was similar to a previous study implementing SCIT within an inpatient forensic setting.[201] The intervention drop-out rate of 24% is similar to previous community-based SCIT groups[204] but is high in comparison with previous inpatient SCIT groups.[201] The difference may be because of the differing recruitment processes (i.e. Combs *et al.*[201] recruited from a pool of inpatients that had already shown an interest in attending whereas, in the current study, participants were recruited from a pool of eligible inpatients with no prior indication of interest).

We found an improvement in one measure – emotion recognition in the SCIT group compared with TAU, which replicates previous findings.[201,204] Perceiving and understanding emotions is the primary focus of the first six sessions of SCIT and this skill is revisited throughout the programme, which might have produced the positive effect in this domain. A change in ToM or attribution style may require longer therapy and more exposure to social situations. In previous research, there have been mixed findings for ToM, with some SCIT studies showing improvements[204] and others showing none.[211] The non-significant finding in attribution bias might have been influenced by the already low AIHQ ratings at baseline which may be a consequence of the therapeutic focus of these forensic wards.

The therapists identified some challenges with running the groups in this setting, mainly regarding practicalities, staffing arrangements and participant engagement. However, ward staff were enthusiastic about SCIT and after discussing issues were motivated to further support the intervention by being flexible with their activity timetable, prioritising the group over other ward tasks and in encouraging participants to engage and attend. A recent independent study evaluating the feasibility of a 20 session SCIT group within a community mental health setting experienced similar challenges in implementation.[203]

The positive views of service users and the subjective feedback showed that participants not only found SCIT helpful in reaching these personal goals, but also that it was generally enjoyable and beneficial, and the sort of group they would like to be offered again.

In conclusion, this study has replicated previous research in a new setting with an adapted programme that may be more suitable to an inpatient environment. It has demonstrated that SCIT is a feasible and acceptable group-based intervention for forensic inpatients. There are promising indications for the effectiveness of SCIT in improving emotion perception in this setting, which is an important domain of social cognition[210] (see *Table 42*).

Appendix 5 The stepped-wedge design

The use of the stepped-wedge design in randomised control trials

The stepped-wedge design is a type of cluster randomised trial in which it is the timing of the intervention that is randomised. As all clusters eventually receive the intervention, this design is often used in situations when, for ethical reasons, the researchers do not wish to deny the intervention to any particular cluster. This was one of the key considerations for the choice of this design in the PERCEIVE trial (or 'in work package 3 – DOORWAYS'). The use of this design is relatively uncommon[212] and analysis methods are recent.[213] The use and analysis of the stepped-wedge design was concurrently being debated throughout the period of this trial.[214]

Another key reason for selecting this design method was that it allows multiple clusters to be included even when it is not possible to intervene in all clusters simultaneously. In our case, it was not possible for researchers to train all wards in the same 6-month time period. However, using a stepped wedge, with two wards randomised at each time point, overcame this problem of resources. In general, the number of steps (data collection points) needed for the trial is one more than the number of randomisation periods, with the first time point used as a control period for all clusters. In our original design, we had eight wards, with two wards randomised to receive treatment at each time point. Time points were 6 months apart. This meant that we originally planned for four intervention periods and five data collection periods. The trial was then twice extended by a further eight wards (each two intervention periods and three data collection periods). In each case the first data collection period coincided with the final period of the trial as planned before that extension (see *Figure 5*).

Collecting the data for the primary end points (VOTE/VOICE) required researchers to visit the wards and spend time with staff and service users. As the later eight wards joined the trial part way through it was not possible to collect outcome data prior to their entry into the study. However, data from a cross-sectional study of similar design (see *Chapter 5*) was available for Lambeth wards at time points prior to randomisation into the DOORWAYS trial. Furthermore, it was not possible to continue to collect trial data for the original wards beyond the planned time points. Similarly, we did not collect trial data from the first extension wards beyond the planned three time points (see *Figure 5*). It has been shown that the statistical power of the stepped-wedge design increases as the number of 'steps' increases.[213] Thus, the lack of data collection had the consequence of effectively reducing the number of time points in the trial and, thus, reducing the power available to detect any true effects of the intervention.

One further consequence of adding extra wards during the study was that they were not included in the original randomisation. In an ideal stepped-wedge trial all clusters are identified prior to commencement. Each cluster is then randomly assigned the timing of the intervention, with all clusters having an equal chance of receiving the intervention first (or indeed at any particular time point). In this trial wards that joined later could not be randomised to receive the intervention in the first period, as this was in the past by the time they joined the study. Hence, we had three distinct randomisations. The first randomisation consisted of eight wards, the subsequent randomisations each consisted of four wards.

The analysis we have chosen takes account of the features of the design as implemented. Modelling wards as a fixed effect allows the effect of time to be estimated within each ward and then averaged to give an overall time effect.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 6 Costings

TABLE 43 Unit costs of services and activities: care professional

Care professional	Cost per hour of patient contact (2008/9 Great British pounds)
Psychiatrist	316
Psychologist	72
Social worker	89
Pharmacist	63
OT	43
Nurse	43
Counsellor	40
Health-care assistant	23
Advocate, volunteer	23
Activities co-ordinator	20

Reproduced from Sabes-Figuera *et al.*[215] This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

TABLE 44 Unit costs of services and activities: activities

Activities	Cost per attendance (2008/9 Great British pounds)
Community meeting	2.5
Bingo, current affairs group, feeling good group, games/quiz group, information trolley	4.8
Coping with stigma, gentle exercise, hearing voices group, narrative expression, sleep hygiene	7
Community outing, film club or film night, gardening group, healthy breakfast cooking, men's group, sleep hygiene, women's group	7.6
Arts and crafts	9.6
Building a compelling future group, CD DJ mixing, communication group, computer/internet access group, health promotion group, healthy eating group, healthy living group, IT skills group, learning how to cope group, music group, planning your future group, pottery group, tai chi group, textiles group, vocational group, woodwork group, yoga group	10
Belly dancing, clinical exercise group, dance and movement therapy, exercise group, feeling good/reflexology/massage, music therapy group, reflexology group, religious or spiritual group	11.4
Art therapy group	11.6
Group therapy	13.3
Creative writing group, CRT, go to the gym, go to the swimming pool, medication group, remotivation process, staying well, swimming group	14.3
Relaxation group	15.9
Baking group, community visit, cooking group, daily planning meeting, drama therapy, go to the chapel, walking group	16.9
Complaints clinic	38.8
Cooking session	39.7

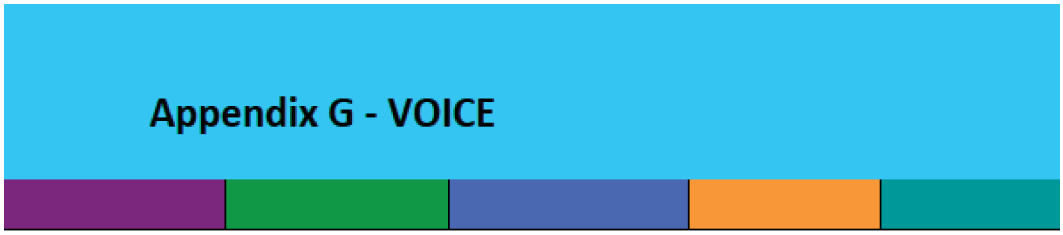
CD, compact disc; DJ, disc jockey; IT, information technology.

Reproduced from Sabes-Figuera *et al.*[215] This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 7 VOICE

Reproduced with permission from King's College London.



Views On Inpatient Care

ServiceUsers'PerceptionsQuestionnaire

Developed through the PERCEIVE programme (NIHR)
www.perceive.iop.kcl.ac.uk

CONFIDENTIAL

Your answers on this questionnaire will be confidential and will not be shown to any of the staff on this ward.

Please give answers based on your experiences on this ward.

We know that individual staff can vary, but try to think about the majority of staff and come up with an average answer.

Please tick one answer per question.

perceive
Patient involvement, improving inpatient care

© 2010 King's College London

© 2010 King's College London

ADMISSION

I was made to feel welcome when I arrived on this ward.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Any Comments

CARE AND TREATMENT

I have a say in my care and treatment.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Ward rounds are useful for me.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Any Comments

MEDICATION

I feel my medication helps me.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I have the opportunity to discuss my medication and side effects.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Staff give me medication instead of talking to me.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Any Comments

STAFF

Staff take an interest in me.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Staff are available to talk to when I need them.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I trust the staff to do a good job.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

© 2010 King's College London

I feel that staff understand how my illness affects me.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I feel that staff treat me with respect.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Any Comments

THERAPYANDACTIVITIES

I think the activities on the ward meet my needs.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I find one-to-one time with staff useful.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Any Comments

ENVIRONMENT

I find it easy to keep in contact with family and friends when I'm on the ward.

O	O	O	O	O	O

Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I feel safe on the ward.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I feel staff respond well when the panic alarm goes off.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I feel staff respond well when I tell them I'm in crisis.

O	O	O	O	O	O
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Any Comments

© 2010 King's College London

Appendix 8 VOTE

Reproduced with permission from King's College London.

Appendix H - VOTE

Views On Therapeutic Environments

StaffPerceptionsQuestionnaire

Developed through the PERCEIVE programme (NIHR)

www.perceive.iop.kcl.ac.uk

CONFIDENTIAL

We understand that the client mix and staff mix changes frequently.

Think about each question over a 4 week period of time on your ward. When you weigh up all the different factors you will be able to arrive at an average answer that can be described using the scale.

Some of these questions will relate to your own practice and some questions will relate to the team generally. If the question relates to the team, try to come up with an average answer.

Please use the comments boxes to 'qualify' your answers. Please circle

one answer per question.

perceive

Patient involvement, improving inpatient care

PATIENT CARE

© 2010 King's College London

It is easy to balance documentation/paperwork and spending time with the patients on my ward.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Patients can feel that there is a sense of 'them and us' on my ward.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Patients are provided with enough information about their medication on my ward.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I feel pressured to complete tasks in my job.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

If I have concerns about patient care I am happy to address it with colleagues.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Comments:

ACTIVITIES

When I ask patients' to join in with activities, they say they are not interested in those on offer.
(See definition 1)

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

TRAINING

Other than mandatory training, staff development opportunities are limited.

(Examples: CBT training, degrees, masters, RMN training for support workers).

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I benefit from regular supervision.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Comments:

MANAGEMENT

I benefit from strong leadership on my ward.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

© 2010 King's College London

I think that the senior managers (above ward managers) understand the current realities of working on acute wards.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Finding enough staff to cover shifts is easy on my ward.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

On my ward there is immense pressure to create bed space.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

When it comes to bed management the clinical perspective of my team is always considered.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

SAFETY

There are enough staff to maintain safety on my ward.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I worry about violence and aggression when at work.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Comments:

TEAM WORKING

When it comes to patient care there are staff in my team who have a 'can't do, won't do' attitude.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I find that communication between the different Multi Disciplinary Team (MDT) professionals is consistently good.
(MDT includes Community Mental Health Team, nursing team and other professionals).

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

Decisions that are made on one shift are changed on the next which makes consistency difficult in my team.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

I'd rather not address relationship issues between teammates because it will create a bad atmosphere.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

There is a strong emphasis on promoting a sense of team spirit on my ward.

1	2	3	4	5	6
Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree

© 2010 King's College London

Any Further Comments?

Thank you very much for completing this questionnaire. We would appreciate any other comments you would like to add.

20 item VOTE scoring card

Total Scores

Total scores are calculated by summing all items with no missing data.

Pro rated Total Scores

Pro rated (mean) total scores are calculated on VOTE questionnaires which were 90% completed or above (i.e. had only one or two missing data). Pro rated scores were calculated as described below:
Pro-rated score = total score/number not missed x total number of items = average of items reported x total number of items.

Reverse Scoring

Negatively phrased items are reverse scored so that higher scores indicated a more negative perception of the ward.

VOTE items	Reverse Scored
1. It is easy to balance documentation/paperwork and spending time with the patients on my ward.	
2. Patients can feel that there is a sense of 'them and us' on my ward.	**Yes**
3. Patients are provided with enough information about their medication on my ward.	
4. I feel pressured to complete tasks in my job.	**Yes**
5. If I have concerns about patient care I am happy to address it with colleagues.	
6. When I ask patients' to join in with activities, they say they are not interested in those on offer.	**Yes**
7. Other than mandatory training, staff development opportunities are limited.	**Yes**
8. I benefit from regular supervision.	
9. I benefit from strong leadership on my ward.	

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

10. I think that the senior managers (above ward managers) understand the current realities of working on acute wards.	
11. Finding enough staff to cover shifts is easy on my ward.	
12. On my ward there is immense pressure to create bed space.	**Yes**
13. When it comes to bed management the clinical perspective of my team is always considered.	
14. There are enough staff to maintain safety on my ward.	
15. I worry about violence and aggression when at work.	**Yes**
16. When it comes to patient care there are staff in my team who have 'can't do, won't do' attitude.	**Yes**
17. I find that communication between the different Multi Disciplinary Team professionals is consistently good.	
18. Decisions that are made on one shift are changed on the next which makes consistency difficult in my team.	**Yes**
19. I'd rather not address relationship issues between teammates because it will create a bad atmosphere.	**Yes**
20. There is a strong emphasis on promoting a sense of team spirit on my ward.	

Appendix 9 CITRINE

Survey of Inpatient Activities and Services Receipt

Ward: _____ Patient ID number: _____

We are interested in finding out what activities patients have been to in the last 7 days. We would also like to find out what staff members patients have seen. Please use a copy of the ward's activity timetable if it would help remind you about activities you have participated in.

Section 1: Group activities on the ward

Please indicate what activities you have taken part in since last_____ by filling out the table below.

Please list the activities you have taken part in:	**How many times since last _____**	**Extra notes** *(researcher use only)*
Activities that take place at the ward		
Community Meeting		

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Activities that take place at the OT resource centre/gym/swimming pool/community/…		

Section 2: Meeting with staff members

Please note any individual meetings you have had with any of the staff members listed below. You do not need to include times where other patients have been involved e.g. in a therapeutic group.

One-on-one time with nurses*:*

Please indicate if you have spent one-on-one time with a nurse since last _____by filling out the table below:

	How many times?	**On average, for about how long ?**
One-on-one time with nurses, other nursing staff or health care assistants		_____minutes

Other professionals:

Please indicate if you have met with the following staff members since last _____by filling out the table below:

Position	**How many times?**	**On average, for about how long ?**
Psychiatrist or consultant		_____minutes
Other doctor		_____minutes
Occupational therapist		_____minutes
Care coordinator (includes telephone conversations)		_____minutes

Please indicate if you have met with any other staff (not nurses) since last_____

e.g. pharmacist, chaplain, psychologist, hairdresser, dentist, physiotherapist, optician, chiropodist, advocate, solicitor

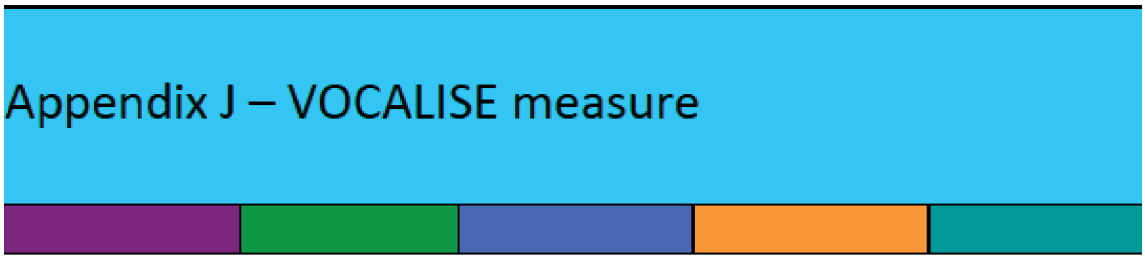
Professional	**How many times?**	**On average, for about how long?**	**Where did you meet?**
		_____minutes	On the ward Off the ward
		_____minutes	On the ward Off the ward

Thank you for taking part in our survey!

© 2010 King's College London

Appendix 10 VOCALISE

Reproduced with permission from King's College London.



Views On Change And Limits in In-patient Settings

A Staff Perceptions Questionnaire

Developed through the PERCEIVE programme (NIHR)
www.perceive.iop.kcl.ac.uk

CONFIDENTIAL

Notes for Completion

1. Within the NHS approaches to practice often change. We would like to learn more about what staff think of this
2. The aim of this questionnaire is to capture staff perceptions of barriers to change
3. It can be useful to pick up on these barriers so we can see how to improve working practices when changes are implemented on wards
4. Please complete the questionnaire based on your experiences of clinical changes that have already happened
5. Some of these questions will relate to your own practice and some questions will relate to the team generally. If the question relates to the team, try to come up with an average answer.
6. Please use the comments boxes to 'qualify' your answers.
7. Please circle one answer per question.

perceive
Patient involvement, improving inpatient care

1

© 2010 King's College London

© 2010 King's College London

1. When it comes to change, information is not circulated effectively on my ward.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
2. I feel confident when delivering new changes.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
3. My whole team is regularly consulted about new ideas for ward practices.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
4. I'm too busy to keep up to date with information about the changes that are happening on my ward.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
5. We can easily fit new changes in with our usual ward practices.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
6. I feel disheartened when others do not want to get involved in changes.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
7. I think that managing risk is more important than delivering new changes.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6

8. Changes just increase my workload and make my life harder.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
9. It is not clear how all changes that we are asked to make will really benefit my ward.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
10. My teammates think that there is no point trying to implement some changes because they won't work.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
11. I find it de-motivating when new changes do not take patients' wishes into account.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
12. I think that some staff would rather let others take the lead in making changes.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
13. When some staff stop engaging with planned changes resistance spreads through my whole team.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
14. I do not really understand how to deliver some of the changes that are suggested by the management.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6

© 2010 King's College London

15. Changes are audited to increase their consistent delivery on my ward.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
16. I always challenge team members who are avoiding delivering new changes.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
17. Inadequate staffing prevents changes being successful on my ward.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6
18. Poor leadership prevents changes happening on my ward.	Strongly Agree	Agree	Slightly Agree	Slightly Disagree	Disagree	Strongly Disagree
	1	2	3	4	5	6

Please rank the top 3 barriers on your ward. You could consider the last 23 barriers mentioned in the questionnaire:

1.

2.

3.

Any Comments?

THANK YOU FOR COMPLETING THIS QUESTIONNAIRE

© 2010 King's College London

Appendix 11 PERCEIVE analysis plans

June 2013

Work package 3 DOORWAYS: service user analysis

Summary

DOORWAYS is a randomised control using a stepped-wedge design. The intervention to increase therapeutic activities is rolled out across wards gradually over time. Different service users were recruited and measured at each time point rather than being monitored over time (see *Tables 45* and *46*).

Objectives

Primary objectives:

1. To measure the impact of increasing therapeutic activities on patient perceptions of care.
2. To determine the cost-effectiveness of increasing therapeutic activities on inpatient wards.

TABLE 45 Stepped-wedge randomisation procedure

	T0 – baseline	T1 – 6 months	T2 – 12 months	T3 – 18 months	T4 – 24 months
Ward 1		Intervention	Intervention	Intervention	Intervention
Ward 2		Intervention	Intervention	Intervention	Intervention
Ward 3		Randomised	Intervention	Intervention	Intervention
Ward 4		Randomised	Intervention	Intervention	Intervention
Ward 5			Randomised	Intervention	Intervention
Ward 6			Randomised	Intervention	Intervention
Ward 7				Randomised	Intervention
Ward 8				Randomised	Intervention

TABLE 46 No cost extension additional wards and time points

	T0 – baseline	T1 – 6 months	T2 – 12 months
Ward 9		Randomised	Intervention
Ward 10		Randomised	Intervention
Ward 11		Randomised	Intervention
Ward 12		Randomised	Intervention
Ward 13			Randomised
Ward 14			Randomised
Ward 15			Randomised
Ward 16			Randomised

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Secondary objectives:

1. To measure the impact of increasing therapeutic activities on –

    i. number of violent episodes on the ward
    ii. patient length of stay
    iii. patient satisfaction
    iv. patient symptoms and functioning.

2. To investigate the sustainability of the positive effects: is this maintained over time or a boost at the initiation of the therapeutic activities which then drops off?

Outcome measures

1. Primary outcomes
2. VOICE: service user perceptions of inpatient care – continuous scale.
3. Secondary outcomes
   Patient measures

    i. SSS-RES: Service Satisfaction Scale – residential, continuous scale, low score (updated 29 November 2013 by Stephen Nash, King's College London, 29 November 2013, personal communication) indicates satisfaction.

   Violent episodes, length of stay.

Additional covariate data

Patient demographics: age, gender, legal status, ethnicity, education, socioeconomic status, accommodation, employment, diagnosis, admission date (hence age at admission and time since admission), time on ward, diagnosis history, medication history, violent incidents, PANSS, GAF, NOSIE, Health of the Nation Outcome Scales (HONOS), activities and one-to-one time.

Ward measures

Number of patients in last 30 days, gender of patients in last 30 days, number of incidents (in 30 days, 7 days and average per day), nurse coverage, average length of stay, activities on the ward, percentage of bed occupancy, ward acuity (average PANSS, GAF or NOSIE).

Descriptive analysis

Summary of patient sample characteristics at each time point.

Summary of ward characteristics at each time point.

Tabulation of patient characteristics by intervention or control arm.

Missing data.

Description of levels of missing data, comparison of missing data between the intervention and control arms. All analyses will account for missing data using weighting or imputations as appropriate.

Analysis of primary outcome

Aim: to measure the impact of increasing therapeutic activities on patient perceptions of care

Descriptive analysis Mean and SD of VOICE in control and intervention arms at each time point and overall, adjusted analysis for calendar effect of time.

Regression analysis Linear regression model of VOICE including an indicator for intervention or control arm accounting for a clustering effect at the ward level.

Adjustment for confounders:

1. Adjust analyses for patient characteristics that are expected to be associated with outcome but are not expected to change with the intervention (e.g. age, legal status, diagnosis, socio-economic status, education, employment, living situation).
2. Adjust analyses for pre-intervention ward level characteristics that are expected to be associated with outcome (e.g. acuity of the ward, patient turnover, average length of stay).
3. Adjust analysis for time since start of study/randomisation.
4. Adjust analysis for number of activities/time spent on activities.

Type of analysis Intention-to-treat analysis will be performed using intervention arm as the covariate of interest.

Analysis of secondary outcomes

Aim: to measure the impact of increasing therapeutic activities on number of violent episodes on the ward, patient length of stay and patient satisfaction

Outcomes: patient satisfaction (SSS-RES) patient length of stay number of violent episodes to date.

Descriptive analysis Means and SDs of outcomes at each time point in intervention and control groups and overall.

Regression analysis Linear regression model of satisfaction and length of stay, Poisson regression of number of violent episodes including an indicator for intervention or control arm accounting for a clustering effect at the ward level.

Adjustment for confounders:

1. Adjust analysis for patient characteristics that are expected to be associated with outcome but are not expected to change with the intervention (e.g. age, legal status, diagnosis, socioeconomic status, education, employment, living situation).
2. Adjust analysis for pre-intervention ward level characteristics that are expected to be associated with outcome (e.g. acuity of the ward, patient turnover, average length of stay).
3. Adjust analysis for time since start of study/randomisation.
4. Restricted analysis to immediate post-intervention comparisons.
5. Adjust analysis for number of activities/time spent on activities.

We intend to include further consideration of the effects of time in this stepped-wedge design (linear trend/non-linear trend, or categorical indicator variables for time points).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Type of analysis
Intention-to-treat analysis will use intervention arm as the covariate of interest.

The following covariates were each tested to see if they confounded the effect of the intervention in WP3.

Service users: gender, age, ethnicity, primary diagnosis, first language, sectioned or voluntary.

June 2013

Work package 3 DOORWAYS: staff analysis

Summary
DOORWAYS is a randomised control using a stepped-wedge design. The intervention to increase therapeutic activities is rolled out across wards gradually over time. Staff members are asked to complete the measures at each time point and so will change from being on a control to an intervention ward over the course of the study (*Tables 47* and *48*).

Objectives
Primary objectives:

1. To measure the impact of increasing therapeutic activities on staff perceptions of ward atmosphere.
2. To determine the cost-effectiveness of increasing therapeutic activities on inpatient wards.

TABLE 47 Stepped-wedge randomisation procedure

	T0 – baseline	T1 – 6 months	T2 – 12 months	T3 – 18 months	T4 – 24 months
Ward 1		Intervention	Intervention	Intervention	Intervention
Ward 2		Intervention	Intervention	Intervention	Intervention
Ward 3		Randomised	Intervention	Intervention	Intervention
Ward 4		Randomised	Intervention	Intervention	Intervention
Ward 5			Randomised	Intervention	Intervention
Ward 6			Randomised	Intervention	Intervention
Ward 7				Randomised	Intervention
Ward 8				Randomised	Intervention

TABLE 48 No cost extension additional wards and time points

	T0 – baseline	T1 – 6 months	T2 – 12 months
Ward 9		Randomised	Intervention
Ward 10		Randomised	Intervention
Ward 11		Randomised	Intervention
Ward 12		Randomised	Intervention
Ward 13			Randomised
Ward 14			Randomised
Ward 15			Randomised
Ward 16			Randomised

Secondary objectives:

1. To investigate the effect (moderating or mediating) of nurses' barriers to change on the impact of increasing therapeutic activities on patient and staff perceptions of the ward.
2. To investigate the sustainability of the positive effects – is this maintained over time or a boost at the initiation of the therapeutic activities which then drops off?

Outcome measures

1. Primary outcome:
   VOTE: staff perceptions of inpatient care – continuous scale.
   Covariates.
   Staff measures
   IWS: Index of Work Satisfaction questionnaire – continuous scale.
   WAS: Ward Atmosphere Scale – continuous scale.
   GMI: Good Milieu Index – continuous scale.
   VOCALISE: Barriers to Change – continuous scale.
   MBI-HSS: Maslach Burnout Inventory-Human Services Survey – continuous scale.

Staff demographics: age, gender, ethnicity, employment band, length of employment, previous wards/trusts, further educational awards.

2. Ward measures:

Number of patients in last 30 days, gender of patients, number of incidents (in 30 days, 7 days and average per day), nurse coverage, average length of stay, activities on the ward, percentage of bed occupancy, ward acuity (average PANSS, GAF or NOSIE).

Descriptive analysis

Summary of staff sample characteristics at each time point.

Summary of ward characteristics at each time point.

Tabulation of staff characteristics between wards randomised to intervention or not at each time point.

Missing data.

Description of levels of missing data, comparison of missing data between the intervention and control arms. All analyses will account for missing data using weighting or imputations as appropriate.

Analysis of primary outcome

Aim: to measure the impact of increasing therapeutic activities on staff perceptions of care

Descriptive analysis Mean and SD of VOTE in control and intervention arms at each time point and overall.

Regression analysis Cross-sectional time series model of VOTE including time and an indicator for intervention or control arm which will change over time accounting for a clustering effect at the ward level.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Adjustment for confounders:

1. Adjust analysis for baseline staff characteristics that are expected to be associated with outcome e.g. band, length of employment, age.
2. Adjust analysis for baseline ward level characteristics that are expected to be associated with outcome e.g. acuity of the ward, patient turnover.
3. Include the staff measure of barriers to change to investigate the impact on and interaction with the treatment effect.

Type of analysis Intention-to-treat analysis will use intervention arm as the covariate of interest.

The following covariates were each tested to see if they confounded the effect of the intervention in WP3.

Staff: gender, age, ethnicity, first language, length of employment.

Work Package 4: BETTER PATHWAYS Service User analysis

Summary

BETTER is a comparison of the two systems of inpatient care (triage care system and routine care). Service users are recruited from wards run under the two models of care. They are recruited four times over a period of 18 months with new participants at every recruitment period.

In addition to BETTER PATHWAYS, there is a full data set of full admissions (ADMIT) during the same time interval of the study.

Objectives

Primary objectives:

1. To investigate differences in patients' satisfaction and patients' perceptions of care between those treated under the two models of care (SSS-RES and VOICE respectively).
2. To investigate differences in ward environment between those treated under the two models of care.
3. To compare cost of activities and staff contacts in the two models of care and their impact of patient outcome (I have not detailed the health economic analyses here as they will be covered by the economics team).

Secondary objectives:

1. To determine the effects on length of stay and re-admission of the two systems (ADMIT).
2. What are the predictors of lengths of stay within the two systems (ADMIT).
3. What are the predictors of patient satisfaction and patient perceptions within the two systems (BETTER PATHWAYS).

Outcomes measures

1. Length of stay – individual patient length of stay until discharged from system.
2. Re-admission.
3. VOICE – Service user perceptions of inpatient care – continuous scale.
4. SSS-RES – Residential Form – continuous scale (total score).

Additional covariate data

Patient demographics: age, gender, marital status, ethnicity, education, socio-economic status, accommodation, employment, diagnosis, admission date (hence age at admission and time since admission), legal status, length of stay, time on ward, diagnosis history, medication history, GAF, HONOS, NOSIE

Ward data
Number of patients in last 30 days, gender of patients, nurse coverage, activities on the ward, average length of stay, number of incidents on ward (in 30 days, 7 days, average per day), acuity of ward (average GAF, NOSIE, HONOS).

Comparison of BETTER PATHWAYS data set to ADMIT
In terms of important variables (socio-demographic and clinical including diagnosis).

Descriptive analysis
Characteristics of the two triage groups in terms of diagnoses, age, gender, employment etc. Test for differences using t-tests and chi-square tests.

Missing data
Description of levels of missing data, comparison of missing data between the systems. Adjustment for significant predictors of missingness, as shown in univariate analyses. Analyses will account for missing data using weighting or imputations if appropriate.

Analysis of outcome

Aim: identify factors that indicate a good model of triage care
We do not know what a good model of care is and so we will investigate this in two stages:

1. Appraise these differences in terms of patient perception and patient satisfaction using the VOICE and SSS-RES respectively (BETTER PATHWAYS).
2. Determine in what ways the two models of care vary e.g. turnover (how many people are admitted every month, length of stay re-admission (ADMIT).

Aim part 1: to investigate differences in patient satisfaction and patient perceptions of care between those treated under the two models of triage care.

Descriptive analysis Mean and SD of VOICE (and its two factors, Interaction and Safety) in the two systems.

Comparisons between VOICE and ward level data (8 wards, 4 time points; 30 data points due to one ward closure); bed occupancy, length of stay, incidents.

Regression analysis Linear regression of VOICE by system, adjustment for time and clustering at ward level.

Further exploratory analyses include adjustment for other predictors:

1. Adjust analysis for patient characteristics that are expected to be associated with outcome but are not expected to change with the intervention e.g. age, diagnosis, legal status, socio-economic status, education, employment, living situation.
3. Adjust analysis for ward level characteristics that are expected to be associated with outcome e.g. acuity of the ward, patient turnover.

Aim part 2: to determine the effects on turnover, length of stay, re-admission and incident rates of the two systems (ADMIT).

Descriptive analysis Mean and SD, median and range if necessary of length of stay and proportions of re-admission in the two triage arms.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Regression analysis Linear regression of length of stay, logistic regression for re-admission rates and Poisson regression for incident rates, by system including an adjustment for clustering at ward level.

Further exploratory analyses include adjustment for other predictors:

1. Adjust analyses for patient characteristics that are expected to be associated with outcome but are not expected to change with the intervention e.g. age, diagnosis, socio-economic status, education, employment, living situation.
2. Adjust analyses forward level characteristics that are expected to be associated with outcome.

Work Package 4: BETTER PATHWAYS staff analysis

Summary

BETTER is a comparison of two systems of care. Staff are recruited from wards run under the two models of care. They are followed up every 6 months and asked to repeat the measures.

Objectives

Primary objectives:

1. To measure differences in staff satisfaction (IWS), staff perceptions of care (VOTE) and levels of burnout (MBI) between those working under the two models of care.
2. To compare cost of activities and staff contacts in the two models of care and their impact of patient outcome (I have not detailed the health economic analyses here as they will be covered by the economics team).

Secondary objectives:

1. What are the predictors of staff satisfaction staff perceptions and staff burnout within the two systems (BETTER PATHWAYS).

Outcomes measures

Primary outcome:

1. VOTE – Staff perceptions of inpatient care – continuous scale.
2. MBI-HSS – Maslach Burnout Inventory-Human Services Survey– continuous scale (Total and three subscales).
3. IWS – IWS.

Secondary outcomes:

1. VOCALISE – Barriers to Change – continuous scale.
2. WAS – WAS – continuous scale.

Additional covariate data

Staff demographics: age, gender, ethnicity, employment band, length of employment, previous wards/ trusts, further educational awards.

Ward data

Number of patients in last 30 days, gender of patients in last 30 days, nurse coverage, activities on the ward, average length of stay, number of incidents on ward (in 30 days, 7 days, average per day), acuity of ward (average GAF, NOSIE, HONOS), proportion of shifts undertaken by temporary staff compared to the total number of shifts excluding student nurses.

Descriptive analysis

Characteristics of the two triage groups in terms of age, gender, employment band, length of employment comparison using t-tests and chi-square as appropriate.

Missing data

Description of levels of missing data, comparison of missing data between the systems. Analyses will account for missing data using weighting or imputations if appropriate.

Analysis of primary outcome

Aim: to measure differences in staff satisfaction between those working under the two models of triage care.

Descriptive analysis

Mean and SD of VOTE in the two triage arms.

Regression analysis

Cross-sectional time series analysis of VOTE by system including time and an adjustment for clustering at two levels, staff level and ward level. Further exploratory analyses include adjustment for other predictors:

1. Adjust analysis for staff characteristics that are expected to be associated with outcome e.g. band, length of employment, age
2. Adjust analysis for ward level characteristics that are expected to be associated with outcome e.g. acuity of the ward, patient turnover.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 12 CONSORT 2010 checklist of information to include when reporting a randomised trial

Section/topic	Item number	Checklist item	Reported on page number
Title and abstract			
	1a	Identification as a randomised trial in the title	n/a
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	xxiii
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	49
	2b	Specific objectives or hypotheses	50
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	51
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	n/a
Participants	4a	Eligibility criteria for participants	52
	4b	Settings and locations where the data were collected	52
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	54
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	52
	6b	Any changes to trial outcomes after the trial commenced, with reasons	n/a
Sample size	7a	How sample size was determined	54
	7b	When applicable, explanation of any interim analyses and stopping guidelines	n/a
Randomisation			
Sequence generation	8a	Method used to generate the random allocation sequence	55
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	55
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	55
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	n/a

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Section/topic	Item number	Checklist item	Reported on page number
Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	n/a
	11b	If relevant, description of the similarity of interventions	n/a
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	55–56
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	55–56
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	n/a
	13b	For each group, losses and exclusions after randomisation, together with reasons	n/a
Recruitment	14a	Dates defining the periods of recruitment and follow-up	n/a
	14b	Why the trial ended or was stopped	n/a
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	58
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	56
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% CI)	61–63
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	n/a
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	n/a
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	67
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	71
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	71
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	71
Other information			
Registration	23	Registration number and name of trial registry	ISRCTN 06545047
Protocol	24	Where the full trial protocol can be accessed, if available	n/a
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	NIHR Programme Grants for Applied Research

Appendix 13 Publications directly arising from the programme grant

Sharac J, McCrone P, Sabes-Figuera R, Csipke E, Wood A, Wykes T. Nurse and patient activities and interaction on psychiatric inpatients wards: a literature review. *Int J Nurs Stud* 2010;**47**:909–17.[15]

Rose D, Evans J, Sweeney A, Wykes T. A model for developing outcome measures from the perspectives of mental health service users. *Int Rev Psychiatry* 2011;**23**:41–6.[17]

Crawford MJ, Robotham D, Thana L, Patterson S, Weaver T, Barber R, *et al.* Selecting outcome measures in mental health: the views of service users. *J Ment Health* 2011;**20**:336–46.[57]

Evans J, Rose D, Flach C, Csipke E, Glossop H, McCrone P, *et al.* VOICE: developing a new measure of service users' perceptions of inpatient care, using a participatory methodology. *J Ment Health* 2012;**21**:57–71.[84]

Laker C, Rose D, Flach C, Csipke E, McCrone P, Craig T, *et al.* Views of the therapeutic environment (VOTE): stakeholder involvement in measuring staff perceptions of acute in-patient care. *Int J Nurs Stud* 2012;**49**:1403–10.[85]

Sabes-Figuera R, McCrone P, Sharac J, Csipke E, Craig T, Rose D, *et al.* Developing a tool for collecting and costing activity data on psychiatric inpatient wards. *Epidemiol Psychiatr Sci* 2012;**21**:393–9.[142]

Csipke E, Flach C, McCrone P, Rose D, Tilley J, Wykes T, Craig T. Inpatient care 50 years after the process of deinstitutionalisation. *Soc Psychiatry Psychiatr Epidemiol* 2014;**49**:665–71.[169]

Wykes T. Great expectations for participatory research: what have we achieved in the last ten years? *World Psychiatry* 2014;**13**:24–7.[187]

Williams P, Csipke E, Rose D, Koeser L, McCrone P, Tulloch AD, *et al.* Efficacy of a triage system to reduce length of hospital stay. *Br J Psychiatry* 2014;**204**:480–5.[179]

Laker C, Callard F, Flach C, Williams P, Sayer J, Wykes T. The challenge of change in acute mental health services: measuring staff perceptions of barriers to change and their relationship to job status and satisfaction using a new measure (VOCALISE). *Implement Sci* 2014;**9**:23.[198]

Papoulias, C, Csipke, E, Rose, D. Mckellar S, Wykes T. The psychiatric ward as a therapeutic space: systematic review. *Br J Psychiatry* 2014;**205**:171–6.[190]

Rose D, Evans J, Laker C, Wykes T. Life in acute mental health settings: experiences and perceptions of service users and nurses. *Epidemiol Psychiatr Sci* 2015;**24**:90–6.[89]

Sabes-Figuera R, McCrone P, Csipke E, Craig T, Rose D, Sharma B, Wykes T. Predicting psychiatric inpatient costs. *Soc Psychiatry Psychiatr Epidemiol* 2016;**51**:303–8.[215]

Csipke E, Williams P, Rose D, Koeser L, McCrone P, Wykes T, Craig T. Following the Francis report: investigating patient experience of mental health in-patient care. *Br J Psychiatry* 2016;**209**:35–9.[184]

Taylor R, Cella M, Csipke E, Heriot-Maitland G, Gibbs G, Wykes T. Tackling social cognition in schizophrenia: a randomized feasibility trial. *Behav Cogn Psychother* 2016;**44**:306–17.[210]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Wykes *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Csipke E, Papoulias C, Vitoratou S, Williams P, Rose DS, Wykes T. Design in mind: eliciting service user and frontline staff perspectives on psychiatric ward design through participatory methods. *J Ment Health* 2016;**25**:114–21.

Wykes T, Csipke E, Williams P, Koeser L, Nash S, Rose D, *et al.* Improving patient experiences of mental health inpatient care: a randomised controlled trial. *Psychol Med* 2018;**48**:488–97.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care

Published by the NIHR Journals Library