

Health Services and Delivery Research

Volume 8 • Issue 28 • June 2020

ISSN 2050-4349

Digital methods to enhance the usefulness of patient experience data in services for long-term conditions: the DEPEND mixed-methods study

Caroline Sanders, Papreen Nahar, Nicola Small, Damian Hodgson, Bie Nio Ong, Azad Dehghan, Charlotte A Sharp, William G Dixon, Shôn Lewis, Evangelos Kontopantelis, Gavin Daker-White, Peter Bower, Linda Davies, Humayun Kayesh, Rebecca Spencer, Aneela McAvoy, Ruth Boaden, Karina Lovell, John Ainsworth, Magdalena Nowakowska, Andrew Shepherd, Patrick Cahoon, Richard Hopkins, Dawn Allen, Annmarie Lewis and Goran Nenadic



Digital methods to enhance the usefulness of patient experience data in services for long-term conditions: the DEPEND mixed-methods study

Caroline Sanders^{1*} Papreen Nahar¹ Nicola Small¹
Damian Hodgson² Bie Nio Ong¹ Azad Dehghan³
Charlotte A Sharp¹ William G Dixon⁴ Shôn Lewis⁵
Evangelos Kontopantelis¹ Gavin Daker-White¹
Peter Bower¹ Linda Davies⁶ Humayun Kayesh³
Rebecca Spencer⁷ Aneela McAvoy^{1,7} Ruth Boaden^{2,7}
Karina Lovell^{1,8} John Ainsworth⁹
Magdalena Nowakowska¹ Andrew Shepherd¹
Patrick Cahoon¹⁰ Richard Hopkins¹⁰ Dawn Allen¹¹
Annmarie Lewis¹¹ and Goran Nenadic³

¹National Institute for Health Research School for Primary Care Research, University of Manchester, Manchester, UK

²Alliance Manchester Business School, University of Manchester, Manchester, UK

³Department of Computer Science, University of Manchester, Manchester, UK

⁴Centre for Epidemiology Versus Arthritis, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK

⁵Division of Psychology and Mental Health, University of Manchester, Manchester, UK

⁶Centre for Health Economics, University of Manchester, Manchester, UK

⁷National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care Greater Manchester, Salford Royal NHS Foundation Trust, Salford, UK

⁸Division of Nursing, Midwifery and Social Work, University of Manchester, Manchester, UK

⁹Centre for Health Informatics, University of Manchester, Manchester, UK

¹⁰Greater Manchester Mental Health NHS Foundation Trust, Manchester, UK

¹¹Patient and public representative

*Corresponding author

Declared competing interests of authors: Caroline Sanders was previously a Director (unpaid) for Affigo CIC (Altrincham, UK) (2016–17), a social enterprise providing digital health products for severe mental illness. Peter Bower reports grants from the National Institute for Health Research (NIHR) during the conduct of the study. Richard Hopkins reports that he is a current director of Affigo CIC, which promotes electronic monitoring of patient symptoms through the use of mobile application, outside the submitted work. Ruth Boaden reports that she was the Director of the NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC) Greater Manchester (2013–19), which was hosted by Salford Royal NHS Foundation Trust where she held an honorary (unpaid) as an Associate Director to fulfil her role as Director of the CLAHRC. She was also a member of the NIHR Dissemination Centre Advisory Group (2015–19) and the Health Services and Delivery Research Funding Committee (2015–19). She was a member of the NIHR Knowledge Mobilisation Research Fellowships Panel (2013–15) and chaired the panel (2016–18). She is a member of the NIHR Advanced Fellowships Panel (2019–present). Azad Dehghan is the Managing Director of DeepCognito Ltd (Manchester, UK) and a Data Analytics Advisor for KMS Solutions Ltd (Manchester, UK). William Dixon receives consultancy fees from Bayer AG (Leverkusen, Germany) and Google Inc. (Mountain View, CA, USA). John Ainsworth reports that he is a Director of Affigo CIC. Shôn Lewis reports that he is a Director for Affigo CIC. Humayun Kayesh reports he is a contract engineer for DeepCognito Ltd. Goran Nenadic reports that he was previously a Scientific Advisor (Non-executive) of DeepCognito Ltd.

Published June 2020

DOI: 10.3310/hsdr08280

This report should be referenced as follows:

Sanders C, Nahar P, Small N, Hodgson D, Ong BN, Dehghan A, *et al.* Digital methods to enhance the usefulness of patient experience data in services for long-term conditions: the DEPEND mixed-methods study. *Health Serv Deliv Res* 2020;**8**(28).

Health Services and Delivery Research

ISSN 2050-4349 (Print)

ISSN 2050-4357 (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 HS&DR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hsdr. 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 *Health Services and Delivery Research* journal

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

HS&DR programme

The HS&DR programme funds research to produce evidence to impact on the quality, accessibility and organisation of health and social care services. This includes evaluations of how the NHS and social care might improve delivery of services.

For more information about the HS&DR programme please visit the website at <https://www.nihr.ac.uk/explore-nihr/funding-programmes/health-services-and-delivery-research.htm>

This report

The research reported in this issue of the journal was funded by the HS&DR programme or one of its preceding programmes as project number 14/156/16. The contractual start date was in April 2016. The final report began editorial review in April 2018 and was accepted for publication in June 2019. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HS&DR 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, NETSCC, the HS&DR programme 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 HS&DR programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2020. This work was produced by Sanders *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).

Editor-in-Chief of *Health Services and Delivery Research* and NIHR Journals Library

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

NIHR Journals Library Editors

Professor John Powell Chair of HTA and EME Editorial Board and Editor-in-Chief of HTA and EME journals. Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK, and Senior Clinical Researcher, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals) and Editor-in-Chief of 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 Director, NIHR Dissemination Centre, 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 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 Ken Stein Professor of Public Health, University of Exeter Medical School, 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

Digital methods to enhance the usefulness of patient experience data in services for long-term conditions: the DEPEND mixed-methods study

Caroline Sanders^{1*} Papreen Nahar¹ Nicola Small¹
Damian Hodgson² Bie Nio Ong¹ Azad Dehghan³
Charlotte A Sharp¹ William G Dixon⁴ Shôn Lewis⁵
Evangelos Kontopantelis¹ Gavin Daker-White¹ Peter Bower¹
Linda Davies⁶ Humayun Kayesh³ Rebecca Spencer⁷
Aneela McAvoy^{1,7} Ruth Boaden^{2,7} Karina Lovell^{1,8}
John Ainsworth⁹ Magdalena Nowakowska¹ Andrew Shepherd¹
Patrick Cahoon¹⁰ Richard Hopkins¹⁰ Dawn Allen¹¹ Annmarie Lewis¹¹
and Goran Nenadic³

¹National Institute for Health Research School for Primary Care Research, University of Manchester, Manchester, UK

²Alliance Manchester Business School, University of Manchester, Manchester, UK

³Department of Computer Science, University of Manchester, Manchester, UK

⁴Centre for Epidemiology Versus Arthritis, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK

⁵Division of Psychology and Mental Health, University of Manchester, Manchester, UK

⁶Centre for Health Economics, University of Manchester, Manchester, UK

⁷National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care Greater Manchester, Salford Royal NHS Foundation Trust, Salford, UK

⁸Division of Nursing, Midwifery and Social Work, University of Manchester, Manchester, UK

⁹Centre for Health Informatics, University of Manchester, Manchester, UK

¹⁰Greater Manchester Mental Health NHS Foundation Trust, Manchester, UK

¹¹Patient and public representative

*Corresponding author caroline.sanders@manchester.ac.uk

Background: Collecting NHS patient experience data is critical to ensure the delivery of high-quality services. Data are obtained from multiple sources, including service-specific surveys and widely used generic surveys. There are concerns about the timeliness of feedback, that some groups of patients and carers do not give feedback and that free-text feedback may be useful but is difficult to analyse.

Objective: To understand how to improve the collection and usefulness of patient experience data in services for people with long-term conditions using digital data capture and improved analysis of comments.

Design: The DEPEND study is a mixed-methods study with four parts: qualitative research to explore the perspectives of patients, carers and staff; use of computer science text-analytics methods to analyse comments; co-design of new tools to improve data collection and usefulness; and implementation and process evaluation to assess use of the tools and any impacts.

Setting: Services for people with severe mental illness and musculoskeletal conditions at four sites as exemplars to reflect both mental health and physical long-term conditions: an acute trust (site A), a mental health trust (site B) and two general practices (sites C1 and C2).

Participants: A total of 100 staff members with diverse roles in patient experience management, clinical practice and information technology; 59 patients and 21 carers participated in the qualitative research components.

Interventions: The tools comprised a digital survey completed using a tablet device (kiosk) or a pen and paper/online version; guidance and information for patients, carers and staff; text-mining programs; reporting templates; and a process for eliciting and recording verbal feedback in community mental health services.

Results: We found a lack of understanding and experience of the process of giving feedback. People wanted more meaningful and informal feedback to suit local contexts. Text mining enabled systematic analysis, although challenges remained, and qualitative analysis provided additional insights. All sites managed to collect feedback digitally; however, there was a perceived need for additional resources, and engagement varied. Observation indicated that patients were apprehensive about using kiosks but often would participate with support. The process for collecting and recording verbal feedback in mental health services made sense to participants, but was not successfully adopted, with staff workload and technical problems often highlighted as barriers. Staff thought that new methods were insightful, but observation did not reveal changes in services during the testing period.

Conclusions: The use of digital methods can produce some improvements in the collection and usefulness of feedback. Context and flexibility are important, and digital methods need to be complemented with alternative methods. Text mining can provide useful analysis for reporting on large data sets within large organisations, but qualitative analysis may be more useful for small data sets and in small organisations.

Limitations: New practices need time and support to be adopted and this study had limited resources and a limited testing time.

Future work: Further research is needed to improve text-analysis methods for routine use in services and to evaluate the impact of methods (digital and non-digital) on service improvement in varied contexts and among diverse patients and carers.

Funding: This project was funded by the NIHR Health Services and Delivery Research programme and will be published in full in *Health Services and Delivery Research*; Vol. 8, No. 28. See the NIHR Journals Library website for further project information.

Contents

List of tables	xiii
List of figures	xv
List of boxes	xvii
List of supplementary material	xix
List of abbreviations	xxi
Plain English summary	xxiii
Scientific summary	xxv
Chapter 1 Background	1
Policy context	1
The concept of patient experience	2
Methods of collecting patient experience data	2
Innovations in the analysis of patient experience data	3
The impact of patient experience data on service improvement	4
Real-time experience	4
Why this research is needed now	4
Choice of long-term conditions and health-care settings	5
Chapter 2 Methodology	7
Summary of the aims and methods for each workstream	7
Workstream 1	7
Workstream 2	7
Workstream 3	8
Workstream 4	8
Exemplar chronic conditions	8
Setting	8
Details of methods for each workstream	10
Workstream 1: perspectives of patients and carers on providing patient experience data and perspectives of staff on the use and usefulness of data	10
Workstream 2: text mining, analysis and presentation of data	13
Workstream 3: co-design of a toolkit	15
Workstream 4: implementation and evaluation	17
Chapter 3 Patient and public involvement	23
Introduction	23
Patient and public involvement contributors	23
Patient participation groups	23
Meetings	24
Some challenges and lessons learnt	24
Participant experiences	25

Results	25
<i>Topics to explore at interview to enhance current feedback collection methods</i>	25
<i>Text mining, analysis and presentation of narrative data</i>	26
<i>Co-design of a suite of tools</i>	26
<i>Development, evaluation, implementation and dissemination of toolkit</i>	28
Chapter 4 Results	31
Workstream 1: perspectives of patients, carers and staff	31
<i>Patients' and carers' views</i>	31
<i>Staff views</i>	32
<i>The need for more meaningful and positive feedback</i>	34
<i>Staff perspectives</i>	35
<i>Methods and tools need to suit the context and informal feedback should be included</i>	37
Workstream 2: text mining, analysis and presentation of data	40
<i>Manual annotation of data sets</i>	40
<i>Automated identification of themes and sentiments by the text-mining methods</i>	41
<i>Report templates to present analysed data</i>	42
Workstream 3: co-design of a toolkit for enhancing the collection, analysis and usefulness of patient feedback	44
<i>Capturing feedback on patient experience: the need for digital and non-digital systems, information and support</i>	45
<i>Analysis and presentation of patient experience data: the value of text mining and perspectives on reporting and presentation of data</i>	49
<i>Eliciting and recording verbal feedback in community mental health services: the importance of enabling feedback through interaction and established ways of working</i>	50
Workstream 4: quantitative analysis of the volume of feedback pre and post introduction of the toolkit	51
Workstream 4: qualitative evaluation	56
<i>Coherence: perceived value of digital tools for collection and analysis</i>	56
<i>Cognitive participation: investment of staff and information and support needs for patients and carers</i>	60
<i>Collective action: organisational and technical work for sustaining new tools</i>	62
<i>Reflexive monitoring: embedding the new intervention</i>	66
Summary and discussion	69
Health economics	70
<i>Costs of developing the toolkit components</i>	70
<i>Costs of implementing the toolkit/kiosk intervention</i>	71
<i>Total costs</i>	71
<i>Costs of the Friends and Family Test</i>	73
Text mining versus qualitative analysis of free-text feedback received in general hospital and mental health service settings: a descriptive comparison of findings	73
<i>Text mining: comparison of service types</i>	73
<i>Qualitative results</i>	74
<i>Comparison between count data obtained using the different analytical methods</i>	74
Chapter 5 Discussion and conclusions	77
Improving the collection and usefulness of patient experience data: perspectives of patients, carers and staff	77
Improving the processing and analysis of narrative data alongside quantitative data	78
Co-design of tools to improve the collection, analysis and presentation of patient experience data for staff to maximise the potential for stimulating service improvement	79

Implementation and process evaluation of new tools to enhance the collection, presentation and use of patient feedback data	80
Implications for health services	81
Implications for future research	81
Acknowledgements	83
References	87
Appendix 1 Tables containing additional information for the text-mining methods	93
Appendix 2 The DEPEND study staff participant information sheet and consent form	97
Appendix 3 The DEPEND study toolkit	103
Appendix 4 Dissemination	111
Appendix 5 The DEPEND study patient and public involvement reflection model	113
Appendix 6 Tables and supplementary material for the text-mining results	115
Appendix 7 Staff Information Sheet	119
Appendix 8 Staff time and costs	125

List of tables

TABLE 1 Existing patient experience data collection methods used in the different sites	10
TABLE 2 Participants by site: WS1	12
TABLE 3 Manually coded data set statistics (the gold standard)	13
TABLE 4 Participants by sites: WS3	16
TABLE 5 Participants by sites: WS4	19
TABLE 6 Number of observation sessions by site: WS4	19
TABLE 7 Theme distribution in the final data set	40
TABLE 8 Inter-rater agreement for the site A data set	41
TABLE 9 Inter-rater agreement for the site B data set	41
TABLE 10 Evaluation results cross-validated on the 'gold standard' data sets for the SBM and CLM	42
TABLE 11 Evaluation of feedback participation rates by method, site and month: WS4	52
TABLE 12 Staff and service user/carer time spent on co-design meetings, focus groups and interviews	70
TABLE 13 Costs of staff time and service user/carer time to develop the toolkit/kiosk (2016–17 prices)	71
TABLE 14 Staff time and costs associated with development of the text-mining tools, interface design and reporting templates and information materials (2016–17 prices)	71
TABLE 15 Costs of implementing the toolkit/kiosks: 9-month evaluation period (2016–17 prices)	72
TABLE 16 Total annual costs per site by length of annuitisation (2016–17 prices)	72
TABLE 17 Text mining: descriptive comparison of findings between mental health and general hospital settings	74
TABLE 18 Count data: text mining versus adapted grounded theory	75
TABLE 19 Topics used for the initial manual coding	93
TABLE 20 Merged themes	93
TABLE 21 Theme distribution in the 'gold standard' data set	93
TABLE 22 Distribution of topics in the 'gold standard' data set	115

TABLE 23 Evaluation results (precision, recall, F ₁ -score) cross-validated on the 'gold standard' data sets for the integrated model	115
TABLE 24 Staff and service user/carer time spent developing the toolkit/kiosk/ data-mining tools	125
TABLE 25 Costs of developing the toolkit/kiosk (2016–17 prices)	126
TABLE 26 Staff time and resources associated with implementing the toolkit/kiosks	126
TABLE 27 Total development and implementation costs per site by allocation of shared costs and length of annuitisation	127

List of figures

FIGURE 1 Summary of the WSs	7
FIGURE 2 Summary of project aims and WSs	9
FIGURE 3 Slide from presentation summarising the links between current context, views of staff and possible tools for change in site A	17
FIGURE 4 Examples of slides used to show participants various ways of capturing the data	17
FIGURE 5 Examples of slides used to show participants potential graphs for displaying analysed data using text mining	18
FIGURE 6 Distribution of positive/negative sentiments in the raw data sets of (a) site A and (b) site B	43
FIGURE 7 Per-class positive and negative aspects in the site A data set (110,854 comments)	43
FIGURE 8 Per-class positive and negative aspects in the site B data set (1653 comments)	44
FIGURE 9 Change in the number of kiosk responses over the study period by collection site	53
FIGURE 10 Change in the number of responses by collection site and response method: (a) site A (acute trust); (b) site B (mental health trust, outpatients); (c) site C1 (primary care practice); and (d) site C2 (primary care practice)	54
FIGURE 11 The SBM workflow	94
FIGURE 12 The CLM workflow	95
FIGURE 13 Automated report to summarise and display the results of the text-mining analysis (example)	116

List of boxes

BOX 1	Illustration of segments of texts labelled by theme and sentiment	14
BOX 2	Technical components and stages for the SBM	94
BOX 3	Technical components and stages for the CLM	95

List of supplementary material

Report Supplementary Material 1 A survey utilising the Friends and Family Test (FFT), with space for free-text comments, to be completed using digital kiosks within study sites, online or using a pen and paper version

Report Supplementary Material 2 Examples of slides used to prompt discussions for co-design focus groups

Report Supplementary Material 3 Additional figures for the responses to the Friends and Family Test questions on kiosks by site

Report Supplementary Material 4 Patient feedback data, organised by sentiment and service setting to facilitate comparison with the text-mining results

Report Supplementary Material 5 Count data from the qualitative analysis for comparison with the text-mining results

Report Supplementary Material 6 Focused descriptive account of exemplar text comments for comparison with the text-mining results

Supplementary material can be found on the NIHR Journals Library report page (<https://doi.org/10.3310/hsdr08280>).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.

List of abbreviations

AGT	adapted grounded theory	NIHR	National Institute for Health Research
CLM	comment-level model	NPT	normalisation process theory
DEPEND	Developing and Enhancing the usefulness of Patient Experience and Narrative Data	PALS	Patient Advice and Liaison Service
FFT	Friends and Family Test	PPG	patient participation group
GP	general practitioner	PPI	patient and public involvement
ID	identifier	SBM	segmentation-based model
IT	information technology	SMI	severe mental illness
MSK	musculoskeletal	SMS	short message service
		WS	workstream

Plain English summary

Background

The NHS collects information about the experience of patients using its services (called 'feedback'), usually in written surveys. Staff do not always find this useful, as surveys may be very general. Patients sometimes write detailed comments, but staff find them hard to analyse.

We conducted a study to make feedback easier to collect and more useful to staff for improving services.

What did we do?

We studied patients with severe mental illness (e.g. schizophrenia) or musculoskeletal conditions (e.g. arthritis) in four sites: a large general hospital outpatient department, a mental health trust and two general practices.

We talked to patients and carers about their views and worked with them to design new ways to collect feedback. We talked to staff about how feedback could be most helpful for service improvement.

We used a computer science method called 'text mining' to summarise detailed patient comments more quickly.

We worked with patients, carers and staff to design new resources (a toolkit) to collect and present feedback.

We talked to patients, carers and staff to see if the toolkit made a difference.

What did we find?

We found a lack of understanding and experience of giving feedback. Participants (patients, carers and staff) wanted more meaningful and positive feedback. They also wanted feedback tools to suit the local context and to include informal feedback (e.g. conversations and thank you cards).

We found that text mining could be used to analyse detailed patient comments, although challenges exist.

The toolkit that was designed with participants included iPad (Apple Inc., Cupertino, CA, USA) kiosks to collect data, guides for patients and staff, text-mining programs and new ways of reporting feedback.

Staff were generally positive about the kiosks. Patients gave more mixed responses and the numbers providing feedback were small. Staff were positive about the feedback reports, but observations showed that no changes were implemented to services during our study.

Scientific summary

Background

Patient experience data are critical to ensure the delivery of high-quality NHS services that are sensitive to population needs. However, there are problems with existing patient experience data. Data are often obtained from structured surveys with low response rates. Front-line professionals are often sceptical about the relevance of data based on generic questions and are concerned that vulnerable respondents are excluded. Many organisations collect additional free-text feedback (e.g. comments, complaints, tweets), which may more accurately reflect patient experiences, but such data are difficult to analyse.

We need more research on the best ways to:

1. collect timely and relevant data
2. analyse existing narrative data routinely and systematically
3. support staff to use multiple forms of data
4. use data to stimulate service improvement.

This study provides evidence on appropriate ways to collect, analyse and use patient experience data, focusing on two groups with long-term conditions: those with severe mental illness and those with musculoskeletal conditions.

Aims and objectives

Our main research question was 'Can the credibility, usefulness and relevance of patient experience data in services for people with long-term conditions be enhanced by using digital data capture and improved analysis of narrative data?'.

To address this, we had four aims:

1. Improve the collection and usefulness of patient experience data by helping people to provide timely, personalised feedback on their experience of services that reflects their priorities and by understanding the needs of staff for effective presentation and use of data. To achieve this we:
 - explored the perspectives of patients, service users and carers on the timing and method of providing feedback
 - investigated the perspectives, needs and current practices of professionals with regard to the collection and use of patient experience data.
2. Improve the processing and analysis of narrative data alongside multiple sources of quantitative data. To achieve this we:
 - analysed existing and new narrative data on patient experiences using text mining
 - explored different ways of presenting the analysed data.

3. Co-design a toolkit with patients, carers and staff to improve resources for enhancing the collection, analysis and presentation of patient experience data to maximise the potential for stimulating service improvements. To achieve this we:
 - co-designed new ways of collecting patient experience data digitally, as well as through alternative methods (e.g. verbal)
 - co-designed better methods for interpreting and summarising narrative patient experience data
 - developed resources and training for staff.
4. Implement the toolkit and conduct a process evaluation to explore implementation, potential mechanisms of effect and the impact of context. To achieve this we:
 - introduced the toolkit to staff
 - investigated staff responses to the toolkit and resulting changes in service provision
 - explored the impact of the toolkit in specific patient groups and service contexts
 - compared text mining with qualitative analysis of feedback
 - compared use of toolkit components with usual feedback mechanisms (including costs).

Methods

Setting

We focused on two exemplar long-term conditions: severe mental illness and musculoskeletal conditions. Both are associated with high levels of service use, complex needs and potential barriers to capturing patient experience. Four sites provided variation in service provision: a rheumatology outpatient department in an acute trust and a community team and outpatient department in a mental health trust; in addition, two general practice sites were selected in the same locality, serving patients with both exemplar conditions.

Design

We conducted four workstreams:

1. *Workstream 1: perspectives of patients and carers and staff.* We conducted semistructured interviews and focus groups with staff ($n = 66$), patients ($n = 41$) and carers ($n = 13$) to understand their views on feedback. A thematic analysis was conducted, drawing on techniques of grounded theory.
2. *Workstream 2: analysis and presentation of patient experience data.* We used text mining to analyse the main themes and sentiments using two data sets of free-text comments from existing patient feedback surveys from the two trusts. Five researchers used a framework of common themes to code data from each trust to assess inter-rater agreement. We developed and validated two machine-learning methodologies and a third system combined those systems using confidence thresholds to improve performance. Reporting templates were created for presenting qualitative and quantitative data together and were used for reporting findings to staff, to explore the value of different presentation methods.
3. *Workstream 3: co-design of a toolkit.* We conducted further interviews and focus groups with participants from workstream 1 and additional participants (total staff, $n = 45$; total patients, $n = 20$). We co-designed new tools, drawing on an experience-based design approach, incorporating findings from workstream 2 and insights from our patient and public involvement group. We created a 'trigger presentation' to stimulate discussion and generate tools for workstream 4.
4. *Workstream 4: implementation and evaluation.* New tools and guidance were introduced in specific staff meetings in the sites. We then conducted a process evaluation using interviews and focus groups with staff ($n = 51$), patients ($n = 24$) and carers ($n = 8$), combined with focused observations (41 sessions), to understand how the new tools were used and how resulting data were used by staff. A thematic analysis was conducted and emerging themes were mapped against core constructs of

normalisation process theory to draw out social practices and interactions associated with implementation of the tools. We analysed the volume of patient experience data before and after the introduction of the toolkit. An economic analysis quantified the time and resources required to use the new tools. We compared text mining of free-text responses with qualitative methods.

Results

Workstream 1: perspectives of patients and carers and staff

The qualitative research revealed three main themes:

1. *Lack of understanding and experience regarding the collection and use of patient experience data.* Many patients and carers did not have a good understanding of why patient feedback was collected and how it was used. Many did not have experience of being asked to give feedback and, even when they did, did not think it was likely to have any influence. Staff often felt distant from feedback mechanisms.
2. *The need for more meaningful and positive feedback.* Staff and patients across all settings thought that there was a need to generate more meaningful data and a need for better ways of analysing and using these data. Narrative comments were viewed as more meaningful. Staff were sceptical of current practices, which were perceived to serve organisational targets rather than being useful for informing delivery of care. Many thought that they would be more engaged if data were more specific to their setting.
3. *Methods and tools need to suit the context, and informal feedback should be included.* Staff and patients were often positive regarding the potential value of digital methods and text mining. They talked about the need to consider context, such as the specific experiences of patients, ways of working among staff (e.g. home visits in mental health) and the organisational environment. Staff and patients in mental health settings were more cautious about the use of digital methods. People with mental health problems said more frequently than those with musculoskeletal conditions that they would be unlikely to use digital methods to give feedback, especially when unwell, and would prefer to give verbal feedback to members of staff who they see regularly. Staff and patients in the mental health context talked about the value of having discussions about patient experiences that would be captured as a more formal record. Staff had a lot of time pressures and any new feedback tools should not add burden. Organisations varied in their capacity to adopt digital tools; for example, primary care staff lacked information technology support and in the trusts there were difficulties identifying who might support new tools.

Workstream 2: analysis and presentation of patient experience data

Initial coding of narratives used 11 categories but, because of overlap and inconsistency between coders, these were merged into five categories: (1) staff attitude, (2) care quality, (3) physical environment, (4) waiting times and (5) 'other'. We merged negative and neutral sentiments, following feedback from clinical staff. Both text-mining systems performed well, with the segment-based model performing marginally better than the comment-level model.

Workstream 3: co-design of a toolkit

The co-designed tools comprised:

1. a survey utilising the Friends and Family Test (FFT), with space for free-text comments to be completed using a digital kiosk, a website or a written version of the survey
2. written guidance for staff, patients and carers to support toolkit use
3. new text-mining programs for analysis
4. new templates for reporting feedback from multiple sources
5. a new process for capturing verbal feedback within community mental health services.

Workstream 4: implementation and evaluation**Quantitative analysis of participation**

The number of participants was relatively low compared with the number of patients using the services. However, the volume of data obtained after introduction of the kiosk was greater than in the previous period in primary care and in the mental health trust. Rates of participation declined over time in all four sites, reflecting the findings related to problems supporting patients and organisational and technical issues.

Qualitative evaluation***Coherence: perceived value of digital tools for collection and analysis***

The construct of coherence refers to the meaning and understanding of new technology and practices. In this case, the initial qualitative research in workstream 1 and the co-design approach in workstream 3 meant that an understanding of 'sense-making' among patients, carers and staff underpinned the development of tools for testing in workstream 4. This helped to maximise a sense of coherence and the tools made sense to staff based on perceived deficits in previous systems. Staff and patients were generally enthusiastic about the kiosks. Staff expected that the kiosks might improve the volume and efficiency of digital data compared with written surveys. However, there was variation between sites, with staff in one primary care site being much more positive. Patients gave more mixed responses and there was also variation among staff in relation to their roles and responsibilities.

Cognitive participation: information and support needs of patients and carers

Cognitive participation refers to relational work to sustain a community of practice for a new intervention. Staff engagement with the new tools varied and observation in the centres indicated that patients were apprehensive to use the kiosks; they mostly would not use them spontaneously but often would participate with support. This made it clear that there was a need for new and specific relational work to sustain the collection of feedback using the digital kiosks; however, staff teams lacked resources and most teams did not have clear motivation to develop a community of support for the tools. Peer support via a 'patient participation group' demonstrated the potential for this in one site.

Collective action: organisational and technical work for sustaining new tools

Collective action refers to the operational work to enact new practices. Low rates of participation highlighted organisational and technical barriers. Staff workload and technical problems were identified as limiting data capture. The level of managerial support for use of the kiosks varied. Collective action varied between the sites, with organisational context being particularly important in the mental health site because a major organisational restructuring took place during the lifetime of the study, which had an impact especially on the ability of staff to operationalise the new verbal feedback process. In the acute trust (site A), some felt that it was ethically wrong to request feedback in case patients felt pressured. The location of the kiosk was highlighted as important, but there were disagreements regarding the optimal location. In one primary care site, introduction of the new tools was more successful. In this site, the senior partner and practice manager were clear champions. In addition, they had a strong patient participation group that provided peer support.

Reflexive monitoring: embedding the new intervention

Reflexive monitoring refers to the work carried out to monitor and appraise new practices. The evaluation period was relatively short. However, staff did reflect on the tools, the data generated and some of the barriers faced in adopting the tools. They were often positive regarding the reports generated to disseminate the feedback and found these to be helpful for stimulating discussion. However, the volume of feedback presented was relatively small and we did not identify any changes to service delivery based on the data during the 9-month evaluation period.

Comparison of qualitative analysis with text mining

The qualitative analysis identified more categories. Estimates of positive and negative sentiments were similar between the two methods in the analysis of mental health data, but different in the data from the acute trust, especially in assessments of 'staff attitudes and professionalism'.

Conclusions

Aim 1: improve the collection and usefulness of patient experience data by helping people to provide timely, personalised feedback on their experience of services that reflects their priorities and by understanding the needs of staff for effective presentation and use of data

Although some have stated that we already collect sufficient data and should shift attention to action in response to data, our study indicates that rates of participation (especially for particular groups) remain a concern.

There was universal acknowledgement that data should be more meaningful than that captured by current brief surveys. Respondents felt that there should be more opportunities to capture verbal feedback, especially in mental health services.

The comparison between different service settings has drawn attention to the importance of context. Although there were some common issues and the potential for some common tools, there is a need to tailor feedback mechanisms. Flexibility, and combining multiple options for feedback, are likely to enhance participation rates.

Aim 2: improve the processing and analysis of narrative data alongside multiple sources of quantitative data

Processing unstructured comments is a challenging task. To solve this, machine-learning classification techniques were used. However, this faced challenges, such as lack of a larger training data set, unbalanced training data, and narratives with complex forms.

Aims 3 and 4: co-design a toolkit to improve resources for enhancing the collection, analysis and presentation of patient experience data and implement the toolkit and conduct a process evaluation

The co-design process led to the development of a model for digital feedback, which was implemented in the four sites. At the start of the study, none of the study teams was routinely collecting patient experience data digitally. We demonstrated that it is possible to enable such routine data capture, with some improvements in data collection, albeit in the short term. However, such tools require additional investment of time and support and there were multiple barriers to adoption, with little evidence of impact over the short period of data collection.

Implications for future research

The findings indicate the need for further research to develop the best ways of enabling more qualitative and informal feedback to be routinely captured and used for service improvement. The study indicates that, when digital methods are used, there is a need for further research into ways of providing guidance and support for people to engage in providing feedback (e.g. through peer support). There is also a need for further research to improve text-mining methodologies for use in service settings and to establish acceptable levels of accuracy and reporting for specific organisational contexts.

Funding

This project was funded by the NIHR Health Services and Delivery Research programme and will be published in full in *Health Services and Delivery Research*; Vol. 8, No. 28. See the NIHR Journals Library website for further project information.

Chapter 1 Background

Policy context

Collecting patient experience data is considered essential for enabling the delivery of high-quality patient-centred care.¹ Better patient care experiences are associated with higher levels of adherence to care, improved clinical outcomes, better patient safety and lower levels of health-care utilisation.² Evidence on what patients value most has been incorporated in the NHS Patient Experience Framework,³ including aspects such as respect and dignity, co-ordination/integration of care, information and communication, physical and emotional care, support for caregivers; and access and continuity. Patient satisfaction surveys are considered integral in the transition to value-based care.⁴

In the UK, patient experience data have routinely been collected for the Care Quality Commission by the Picker Institute through the NHS patient survey programme.⁵ Annual surveys of patient experience, such as the national GP survey and the national inpatient survey, have been conducted retrospectively by mail, with response rates. Recently, the Francis report⁶ and the Berwick review⁷ highlighted the need to collect data that are 'real time', or as near as possible to real time, as a means of enabling safe care.

The Friends and Family Test (FFT) asks whether or not patients would recommend a service to friends and family. The FFT has been used since 2012 as a means of gathering simple and timely patient experience feedback. A new indicator based on this test has also been included within the NHS Outcomes Framework³ to 'enable more "real-time" feedback to be reflected in the framework'.

It is evident that, although policy-makers have highlighted the value of the FFT, there have been critical discussions of its inappropriateness (especially in sensitive circumstances), the limitations of the data generated and the lack of evidence of benefits for service improvement.⁸ It has also been found that services are not very engaged with the FFT, with the purposes of collecting FFT data often unclear to staff. Many believed that the FFT was intended for performance management, resulting in low levels of local 'ownership' of data collection.⁹ Recently, questions have been raised about whether or not use of the FFT should be mandatory if it generates limited insights.¹⁰

Coulter *et al.*¹¹ identified the lack of impact that patient feedback appears to have on change in the NHS. Gleeson *et al.*¹² identified multiple barriers to collecting patient feedback, including the fact that surveys are perceived to lack specificity, that managers and clinicians often lack the skills needed to interpret findings and that reports are not seen to be timely. There are also questions about the impact of feedback (such as complaints) on clinicians and managers.^{13,14}

Facilitators of impact from patient feedback include a supportive culture for change, dedicated time to discuss results and improvements, and more context-specific data. There is also evidence of significant variation between, and within, organisations with regard to the use of, and approach to, patient experience data. For example, the size of a trust, dispersal of sites, history, demographics and corporate culture play a huge part in how well the collection and use of feedback for service improvement is undertaken.¹⁵ In addition, improvement has been evident when there has been specific policy focus and investment, as well as incentives and penalties.¹⁶ Sheard *et al.*¹⁷ propose two conditions that need to be in place for effective use of feedback: first, normative legitimacy, whereby staff believe that listening to patients is worthwhile, and, second, structural legitimacy, which provides teams with adequate autonomy, ownership and resources to enact change.

The concept of patient experience

The policy emphasis on patient experience appears to underline its importance as a vital component of care.¹⁸ The National Quality Board¹⁹ states that ‘experience’ can be understood in the following way: what the person experiences when he or she receives care or treatment. These experiences can be summarised in two ways: (1) the interactions between the person receiving care and the person providing care (i.e. ‘relational’ aspects of experience, including communication) and (2) the processes that the person is involved in or that affect his or her experience, such as booking an appointment (this is known as the ‘functional’ aspects of experience), and how that made him or her feel, for example whether or not he or she felt that he or she was treated with dignity and respect.¹⁹

A literature review²⁰ noted the absence of a common definition of patient experience but described a number of key themes, such as active patient engagement and person-centeredness and its integrative nature. However, questions remain regarding what constitutes patient experience, how it is best captured and how it is used to influence change in specific contexts and for specific groups of patients and carers.

Sociological perspectives on patient experience depict this as a dynamic concept: an exercise in testing versions of reality through ongoing respecification of objects, audiences and identities.²¹ From the patient’s perspective, their experience is shaped by factors such as their social context, past and present health-care encounters, the dynamic between wellness and illness²² and interactions between conditions when living with multiple morbidities.²³ Given the potential variety of patient experiences, an understanding of the mechanisms whereby they are turned into knowledge is important. Mazanderani *et al.*²⁴ analysed the use of the internet to share experiences and highlighted the tension between similarity (such as sharing the same diagnosis or treatment) and difference [such as living with condition(s) in specific contexts]. This tension needs to be negotiated and can lead to a recognition of ‘being differently the same’, which allows for multiple experiences to become a source of knowledge and support.²⁴ Lupton²⁵ offered critical reflections on the ways in which patient experience is shared in digital formats, and commodified forms of usable data, which placed consumerism as a factor in the centre of doctor–patient relationships.

The argument that patient experiences need to be captured in real time is important, yet implicitly accepts that it is possible to describe a dynamic and long-term experience with a single assessment at each point of service use. Vogus and McClland²⁶ argue that the factors that influence patient experience change over time: interaction and communication are more prominent when patients are asked their opinion immediately after their consultation, whereas health status and symptom resolution increase in importance over time.

Methods of collecting patient experience data

A 2012 survey²⁷ found that most trusts report that self-completed paper surveys are the most frequently used data collection method, but that a large proportion also collect digital data (55% with the help of an administrator or volunteer; 42% by patients themselves during an inpatient stay). In addition, 27% of trusts were planning greater use of digital data capture and 23% of trusts stated that the Department of Health and Social Care could best help with data collection by providing better technology to help capture and analyse ‘real-time’ data.²⁷

In recent reports, the use of scores generated from the FFT for measuring and comparing the performance of providers has been viewed as problematic, but it has also been argued that the FFT can be useful for improving services and at least generating discussion about patient experience.^{28,29} With regard to the FFT, patient comments have been perceived to provide the greatest value, but this does then raise questions regarding how to effectively analyse and utilise such comments.

Recent guidance states that qualitative sources of patient experience data are equally as valuable as quantitative surveys and local organisations are currently advised to supplement mandatory survey data with a range of data sources, including patient stories, complaints, Patient Advice and Liaison Service (PALS) data, incident reports and general feedback.³⁰ One study revealed that patients would like to see feedback as a mechanism to express their individual experience in detail and in context.²⁰ There is a perceived failure to 'close the loop' in the sense that people do not see the impact of their feedback in service change, either for individuals or at the organisational level.¹⁸

Recent studies have been performed on how best to measure distinct components of feedback³¹ and the value of the collective judgement of feedback.³² This debate about how to adequately capture the complexity of the patient experience can benefit from recent literature on the use of 'big data'. Authors argue that the perspectives of both patients and staff highlight the limitations of formal forms of quantifiable data and point to the need for more small, informal data sets based on interactive and highly contextual mechanisms in order to facilitate a more detailed understanding of the process and experience of care.¹² Additionally, Ziewitz²¹ argues that mobilising patient experience to influence organisational change needs ethnographic approaches alongside quantitative methods in order to challenge and complement findings from 'big data' analysis.

There has been a growing focus on the value of patient narratives in audiovisual and textual forms and such stories have been drawn on for staff training and service improvement.^{33,34} However, these stories represent limited numbers of individual patients and are time-consuming to produce. One of the tensions around different methods of collecting patient experience data is that between the depth of data and the burden associated with its analysis. Qualitative forms of data (such as narrative descriptions) hold out the promise of rich feedback specific to the particular clinical context, which may be more impactful and provide a better basis for action. However, such data are much more complex and time-consuming to analyse, unless effective ways can be found to automate the process.

Innovations in the analysis of patient experience data

There have been only a limited number of attempts to extract information such as patient feedback automatically, and this was typically for quite focused tasks. For example, Greaves *et al.*³⁵ analysed online patient feedback comments to understand the sentiment about overall care quality, including cleanliness and dignity. The authors applied a machine-learning classification approach that mapped each comment into relevant topics and associated sentiments. Similarly, Cole-Lewis *et al.*³⁶ extracted five predefined topics related to e-cigarettes, experimenting with different classification algorithms. Recently, a similar approach was followed by Gibbons *et al.*³⁷ to extract different aspects from free-text patient comments on doctors' performance.

In contrast to supervised methods, Padmavathy and Leema³⁸ experimented with a patient-opinion-mining system in which unsupervised methods were used to identify key topics. The topics were then clustered for sentiment extraction. Tapi Nzali *et al.*³⁹ applied a similar unsupervised approach to extract topics from social media text and to compare results against questionnaire responses. Wagland *et al.*⁴⁰ similarly compared the outcomes of a machine-learning approach with the outcomes from a manually conducted survey to extract patient perceptions of care.

Brookes and Baker³¹ used mixed methods to qualitatively analyse patient feedback (NHS Choices) after automatically extracting and then manually inspecting a small subset of the most frequently used words. They found that most comments are associated with treatment, communication, interpersonal skills and organisation. However, although they used a large data set, this was mostly manual analysis based on specific words appearing in patient comments.

Given the lexical variability and ungrammaticality of patient-generated comments, it is not surprising that the majority of work in this domain relies on using machine-learning approaches, aiming to learn from data. However, manually providing training data for supervised methods is a challenging issue and, therefore, unsupervised methods have been also explored. In relation to patient feedback, text mining is the processing of comments; a similar process occurs in other settings when customers provide product reviews by commenting on different features of products. Frequently mentioned topics in review comments are often referred to as 'aspects'. Aspect-based sentiment extraction from product reviews has been an active research area for a long time; for example, in addition to unsupervised machine-learning techniques to identify aspects and detect associated sentiment (e.g. Brody and Elhadad⁴¹), there are additional examples of supervised approaches.⁴²⁻⁴⁴ Recently, Hai *et al.*⁴⁵ proposed a method that combines both supervised and unsupervised machine-learning techniques.

The impact of patient experience data on service improvement

Most staff report that their directorate or department collects patient feedback (89.7% in 2017).⁴⁶ The proportion of staff reporting that this feedback is used to inform decision-making is significantly lower, although this has improved over time. In 2015, 55.6% of staff said that feedback was used to inform decision-making in their service area, with this proportion increasing slightly to 56.7% in 2016 (remaining the same for 2017).⁴⁶

Real-time experience

Some trusts have been pioneering the collection and use of real-time feedback using tablet devices for a number of years. For example, in 2011, St George's University Hospitals NHS Foundation Trust initiated such a system and has since introduced tablet computers to almost all patient areas.

Such systems emphasise the importance of making reports instantly available to highlight areas of especially good or especially poor performance, allowing staff to make improvements or share positive feedback as examples of good practice.

Why this research is needed now

Our overall research question was 'Can the credibility, usefulness and relevance of patient experience data in services for people with long-term conditions be enhanced by using digital data capture and improved analysis of narrative data?'

This research is important because NHS organisations are already investing substantial resources in collecting large quantities of data on patient experience. However, as highlighted above, there are major inefficiencies in current methods of collecting, analysing and using such data. To address our overall research question, this study had four aims:

1. Improve the collection and usefulness of patient experience data by helping people to provide timely, personalised feedback on their experience of services that reflects their priorities and by understanding the needs of staff for effective presentation and use of data.
Recent research has shown that professionals are often sceptical of the relevance of patient experience data to local services because they are based on generic questions rather than being tailored for specific service contexts and because it is perceived that vulnerable patients and carers are excluded.^{47,48} It is crucial to understand the needs of staff regarding the feedback of patient experience data if they are to be used to stimulate service improvements. As previously mentioned, NHS organisations already collect a wide range of quantitative and qualitative patient experience

data, but there is a lack of understanding regarding how best these different types of data can be presented and used by staff. Although qualitative data in the form of free-text comments is a large and potentially useful resource, questions remain regarding its representativeness and credibility from the point of view of staff. In addition, individual patient narratives may well be powerful but may not be considered representative of patient experiences more widely. In trusts collaborating with this study, patient stories have been used in the context of board meetings, as in many trusts across the UK; however, such stories are not routinely viewed by teams of front-line staff and the views of staff regarding the relevance and use of such data are unknown. This study sought to use qualitative methods to understand staff perspectives on data requirements.

2. Improve the processing and analysis of narrative data alongside multiple sources of quantitative data.

As highlighted above, patient experience data are often provided in narrative form, yet organisations lack the capacity to utilise these data effectively. Text-analytic techniques enable the automated and systematic analysis of large sets of qualitative data gathered from multiple sources of patient experience feedback. We aimed to explore how automated text-analytic techniques could be utilised⁴⁹ to provide a means for the automated and continuous analysis of relevant textual data that are routinely collected.

3. Co-design a toolkit with patients, carers and staff to improve resources for enhancing the collection, analysis and presentation of patient experience data to maximise the potential for stimulating service improvements.
4. Implement the toolkit and conduct a process evaluation to explore implementation, potential mechanisms of effect, and the impact of context.

The original commissioning brief highlighted the importance of understanding the kind of organisational capacity needed in different settings to interpret and act on patient experience data. We proposed to develop a toolkit for enhancing the collection, analysis and presentation of patient experience data, based on an understanding of the views of patients and staff. We built into the toolkit appropriate text-mining techniques to provide efficient analysis of narrative data.

To explore the implementation and impact of the toolkit in routine NHS settings, we conducted a process evaluation.⁵⁰ The evaluation used qualitative methods to enable a detailed understanding of the needs of distinct organisational teams and how organisational capacity varies according to contextual factors, such as the distinct patient groups served, the size of the teams, management structure and the nature and flow of work. We drew on normalisation process theory (NPT), which has been developed and used to understand the actions and interactions influencing implementation and how new interventions and practices come to be normalised in health-care contexts.⁵¹ There is a dearth of evidence on the relative costs of different methods to collect and use data, and how staff use varied data to inform service changes,¹¹ and we provide costings associated with the toolkit to better understand the costs involved.

Choice of long-term conditions and health-care settings

A 2011 project from The King's Fund⁴⁸ focused on five key pathways (stroke, chronic obstructive pulmonary disease, diabetes mellitus, depression and elective hip replacement) to explore patient experience. This report suggested that, ideally, patient experience should be measured in terms of the journey as experienced by the patient in order to capture transitions in care and continuity issues. It also suggested that generic surveys may need to be supplemented with condition-specific indicators.

In this study we focused on appropriate ways to collect, analyse and use patient experience data in the pathways for severe mental illness (SMI) and musculoskeletal (MSK) conditions.

A SMI, such as psychosis, affects 2% of the UK adult population. Patients with a SMI have a lower life expectancy (25 years less than the general population, mainly because of physical health problems) and are at greater risk of suicide and self-harm.⁵² These are two examples of key issues for which feedback

on patient experience might be used to ensure that services are meeting physical needs, as well as maintaining safety.

In the UK, 14.3% of adults report having a chronic MSK condition.⁵³ This has a major impact on health-care resources: it is one of the most common reasons for a primary care consultation, with one in four adults in Europe being on long-term treatment for a long-standing MSK problem. Common problems occur in managing long-term medications across primary and secondary care and in meeting needs for secondary care that could be reflected in patient experience feedback and could be used to stimulate service improvement.

These were important groups for research on the use of patient experience data because:

- People with both types of condition show high levels of service use in primary and secondary care settings, allowing us to consider the need for patient experience data of multiple clinical teams.
- These long-term conditions invoke common concerns regarding continuity of care and patient safety, often reflected in patient experience narratives.⁵⁴
- Research suggests that there are particular safety concerns in relation to SMI, with aspects of the provision of mental health services able to affect suicide rates.⁵⁵
- There is uncertainty around the applicability of Picker survey frameworks to severe mental health problems, especially because service users may be forced to receive care.
- Both populations commonly have overlapping comorbidities and include those who are under-represented in current methods for capturing data on patient experience: older people with prevalent MSK conditions, vulnerable younger adults with SMI and carers in both cases.

The research focused on the use of patient experience data by staff teams in four sites:

1. site A – an acute trust (focusing on rheumatology outpatients)
2. site B – a mental health trust (focusing on a community mental health team and an outpatient clinic)
3. site C1 – a general practice
4. site C2 – a general practice.

The above focus aligns with the NHS Outcomes Framework,³ which highlights key improvement areas to ensure that patients have positive experiences of care, including patient experiences of community mental health services, patient experiences of outpatient services and access to general practice services.

Chapter 2 Methodology

Our main research question was ‘Can the credibility, usefulness and relevance of patient experience data in services for people with long-term conditions be enhanced by using digital data capture and improved analysis of narrative data?’.

To address this, four aims mapped onto four workstreams (WSs) (*Figure 1*). The methods for each WS are summarised in the following sections, followed by some context for the setting and further methodological details for each WS.

Summary of the aims and methods for each workstream

Workstream 1

Aim

To improve data collection and usefulness by helping people to provide timely, personalised feedback on their experience of services that reflects their priorities and by understanding the needs of staff for effective presentation and use of data.

Methods

We used qualitative methods to explore the perspectives of patients and carers on providing patient experience data. We used the same methods to investigate perspectives and current practices in the use of patient experience data by clinical teams, managers and information technology (IT) staff (when applicable).

Workstream 2

Aim

To improve the processing, analysis and presentation of narrative data.

WS	Research activities
1	<ul style="list-style-type: none"> Qualitative research to explore perspectives of patients, service users and carers on the timing, form and method of providing feedback Qualitative research to investigate perspectives, needs and current practices in the collection and use of patient experience data by professionals
2	<ul style="list-style-type: none"> Analysis of existing and new narrative data on patient experience using text mining Exploration of different ways of presenting analysed data
3	<ul style="list-style-type: none"> Co-designed new ways of collecting patient experience data digitally, as well as using alternative methods (e.g. verbal) Co-designed better methods for interpreting and summarising narrative patient experience data Developed information and guidance for patients, carers and staff
4	<ul style="list-style-type: none"> Introduced the toolkit to staff Investigated staff responses to the toolkit and resulting changes in service provision Explored the impact of the toolkit in specific patient groups and service contexts Compared text mining with qualitative analysis of feedback Compared use of toolkit components with usual feedback mechanism (including costs)

FIGURE 1 Summary of the WSs.

Methods

We used computer science text-analytics methods⁴⁹ to develop programs for routine, automated and systematic analysis of narrative data. We also explored different ways of presenting analysed patient experience data.

Workstream 3

Aim

To co-design a toolkit to improve resources for enhancing the collection and analysis of patient experience data and presentation to staff teams to maximise the potential for stimulating service improvements.

Methods

We used an experience-based design approach,⁵⁶ drawing on the initial qualitative research, the computer science work in WS2 and insights from our patient and public involvement (PPI) group, to co-design ways to enable and support digital data capture, analysis and use of both quantitative and narrative data.

Workstream 4

Aim

To implement the co-designed toolkit and evaluate its impact for improving the collection, analysis and presentation of patient experience data.

Methods

We introduced the toolkit and trained staff in multiple service teams that participated in the initial qualitative research to use the tools. We then conducted a process evaluation⁵⁰ by analysing participation rates to assess whether or not a greater degree of feedback was obtained in the multiple sites using the tools. We also used qualitative methods and drew on NPT⁵¹ to assess the impact of the toolkit for improving the usefulness of data and any influences on service changes. We also compared the text-mining approach to analysing free-text comments with a standard approach and investigated the time spent collecting and analysing data using the new tools to estimate costs.

Figure 2 provides a summary of the project aims and WSs.

Exemplar chronic conditions

The study focused on services for two exemplar long-term conditions, SMI and MSK conditions, with high levels of service use, comorbidity and concerns regarding the continuity and safety of care.

Setting

The study was conducted in four sites:

1. site A – a large acute trust focusing on one outpatient rheumatology department for the qualitative research
2. site B – a mental health and social care trust focusing on one community mental health team and an outpatient clinic for the qualitative research
3. site C1 – a general practice surgery based in an area of high deprivation within the same locality as site A
4. site C2 – a general practice surgery based in an area of high deprivation within the same locality as site B.

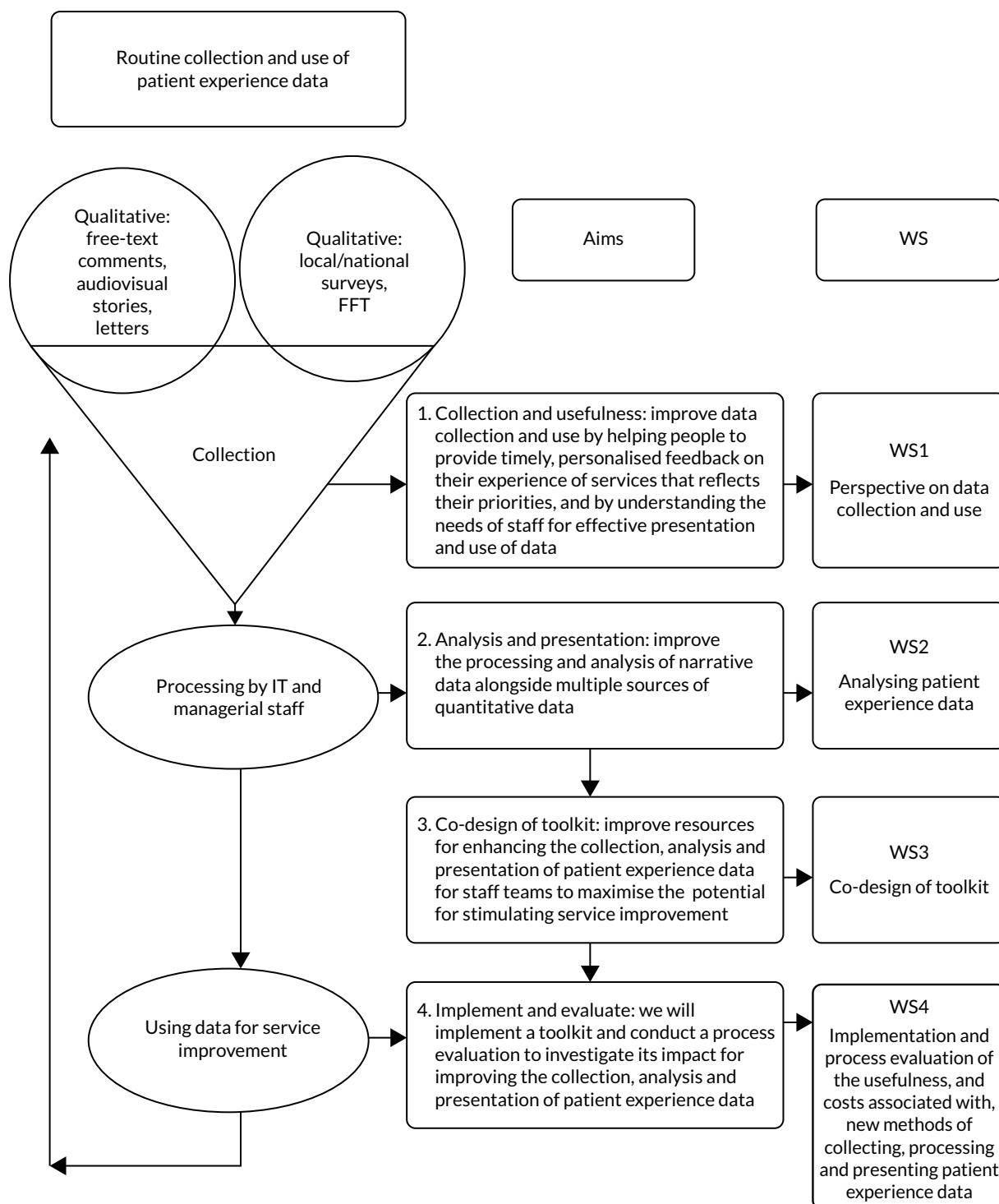


FIGURE 2 Summary of project aims and WSs.

The sites were selected to ensure that diverse organisational contexts and variations in terms of the methods of collecting patient experience feedback were considered. For example, we included a well-resourced tertiary hospital, as well as a main trust providing mental health services, and community health services and small-scale practices.

Site A is an integrated provider of hospital, community and primary care services, including a university teaching trust. The trust provides local services to the city where it is located.

Site B offers a wide spectrum of mental health, social care and well-being services to meet the needs of adults of working age and older adults across a large city. Six community mental health teams based throughout the city provide assessment, care and support for adults of working age and older adults with mental health problems.

The two general practices, site C1 and site C2, serve a range of patients, including those with severe mental health problems and rheumatology patients.

Each setting had different practices for collecting and using narrative data, in addition to standard survey data, including the FFT. Most relied on pen- and paper-based surveys, with low levels of participation and challenges in terms of collecting and processing responses, with sites A and C2 also collecting digital responses by short message service (SMS) text (i.e. message). None of the sites routinely collected digital data on site, for instance in waiting rooms or reception areas in the outpatient departments or at the point of home visits for the community mental health teams. The acute trust collected digital patient experience data only occasionally from selected patients using a handheld digital device. The format of the narrative data varied across the settings, but included free-text comments made in response to an open survey question (e.g. following the FFT). These data also included individual audiovisual stories that were used for organisational activities, such as board meetings or training events, in each large trust (audio-visual stories were carefully produced, with very few patients selected because of some particular issues that may reflect wider problems that trusts were seeking to improve). Additional forms of feedback were collected to varying degrees through the study sites' websites, other websites [e.g. Facebook (Facebook Inc., Menlo Park, USA)], letters to PALS and patient discussion groups. *Table 1* shows examples of the different feedback data collection methods used in the different sites.

Details of methods for each workstream

Workstream 1: perspectives of patients and carers on providing patient experience data and perspectives of staff on the use and usefulness of data

Participants

Staff members were recruited from all four sites and patients and carers were recruited from sites A and B. Although we have used the term 'service users' to refer to those with mental health problems, for simplicity we have sometimes referred to those with mental health problems as 'patients' when referring to both participants with MSK conditions and participants with mental health problems.

TABLE 1 Existing patient experience data collection methods used in the different sites

Data collection methods									
Site	Website	Twitter ^a	SMS	Facebook ^b	Letters to PALS	Audiovisual stories	Patient Voices programme	Dignity walks, community discussion	Pen and paper
A	✓	✓	✓	✓	✓	✓	✗	✗	✓
B	✗	✗	✗	✓	✓	✓	✓	✓	✓
C1	✗	✗	✗	✗	✗	✓	✗	✗	✓
C2	✗	✗	✓	✗	✗	✗	✗	✗	✓

a Twitter, Inc., San Francisco, CA, USA. URL: www.twitter.com
b Facebook Inc., Menlo Park, CA, USA. URL: www.facebook.com

Data collection

Focus groups, individual interviews and observation methods were used for data collection. The use of mixed qualitative methods and mixed participants allowed in-depth exploration of stakeholder views and experiences and also enabled triangulation of the data regarding emerging key themes.

The study was conducted within four WSs to achieve the overarching aim of collecting patients' experience feedback to ensure the delivery of high-quality, effective and safe services that are sensitive to population needs.

The WS1 data collection took place between July 2016 and April 2017. Patients were recruited through staff in the clinical sites.

Perspectives of service providers

Staff participants were invited to take part in either a focus group or a face-to-face individual interview, or both, as determined by participant preference. In total, 66 staff participants took part in the qualitative components of WS1 to understand their perspectives on what data need to be collected to be useful for staff and the current practices in each setting. In relation to staff participants, 21 individual interviews were undertaken (four of these participants chose to take part in both an individual interview and a focus group) and 49 participants took part in one of the staff focus groups in each site.

We sought participants' views on the credibility and usefulness of different types of data, including narrative and textual sources. Views and experiences of clinical teams and managerial and IT staff were sought regarding current practices, organisational capacity for using various data sources and current barriers. We achieved a diverse sample of staff participants with roles in management and patient experience, clinical practice and IT, reflecting the various categories of interest in each organisation.

In site A, there were 20 staff participants (lead IT/governance/patient experience managers, $n = 7$; nurse managers, $n = 4$; consultant specialists, $n = 8$; occupational therapist advanced practitioners, $n = 1$). In site B, there were 22 staff participants (lead governance managers, $n = 1$; team managers, $n = 3$; consultant specialists, $n = 3$; care co-ordinators, $n = 13$; carer support workers, $n = 1$; administrators, $n = 1$). In site C1, there were 13 staff participants [general practitioners (GPs), $n = 3$; nurses, $n = 2$; health-care assistants, $n = 2$; practice managers, $n = 2$; administrators/reception staff, $n = 4$]. In site C2, there were 11 staff participants (GPs, $n = 5$; nurses, $n = 3$; health-care assistants, $n = 1$; practice managers, $n = 1$; finance managers, $n = 1$).

Perspectives of service users, patients and carers

In total, 41 patients and 13 carers in sites A and B took part in the qualitative components of WS1 to provide patient experience data and to define what forms of data they were willing to provide using different methods (Table 2). In total, 20 individual interviews were undertaken with patients with MSK conditions (site A), 10 individual interviews were conducted with service users (site B) and another 11 service users took part in a large focus group (site B). We also conducted five individual interviews with carers of patients with MSK conditions (site A) and eight individual interviews with carers of service users (site B).

Our patient and carer WS1 sample was less diverse than our staff sample. In site A, there were more female than male patient participants (14 women, 6 men); the median age of participants was 60.5 years and 18 participants were of white British ethnicity, one was of mixed ethnicity and one was of Asian/black British ethnicity. All five carer participants from site A were male; the median age of participants was 65.5 years and participants were predominantly of white British ethnicity, with one participant of Asian British ethnicity. Conversely, in site B there were more male than female patient

TABLE 2 Participants by site: WS1

Participants	Site				Total
	A	B	C1	C2	
Staff focus groups	10	15	13	11	49
Staff interviews	12 (2)	9 (2)	0	0	21 (4)
Total staff	20	22	13	11	66 (4)
Patient focus groups	0	11			11
Patient interviews	20	10			30
Total patients	20	21			41
Carer focus groups	0	0			0
Carer interviews	5	8			13
Total carers	5	8			13
Total	45	51	13	11	120 (4)
Note The numbers in parentheses are the numbers of participants who took part in a focus group as well as an interview; data from patients/service users and carers in sites C1 and C2 were not collected for WS1.					

participants (11 men, 10 women); the median age of participants was 19 years and participants were predominantly of white British ethnicity, with one participant of black African/Caribbean/British ethnicity. Five of the eight carer participants in site B were female; the median age of participants was 49 years and all were of white British ethnicity.

Ethics issues and real or potential barriers to participation were also considered.

We investigated perspectives on providing textual data and the structured questions considered most suitable for capturing experiences of the specific service user groups. We also explored views about how to provide feedback to best reflect patients' concerns and needs, for example whether or not people want to provide data digitally and what is the best way to do this – using a handheld or fixed device in service settings, by mobile phone or by computer. We considered specific needs in each group, for example concerns that mental health service users have that are specific to their care pathway, such as involvement in care planning, concerns about enforced inpatient care and how positive and negative experiences of these aspects of care can be conveyed. We also asked for participants' views about providing both experience and outcome data simultaneously and the best ways of capturing these multiple forms of data.

All interviews and focus groups were transcribed, collated and analysed thematically, drawing on techniques of a grounded theory approach⁵⁷ and using NVivo 11 qualitative analysis software (QSR International, Warrington, UK). A process of open coding by the research lead for each site provided an initial framework, which was discussed collectively and refined for consistency and focus by the research team. Through the coding and discussions regarding links and distinctions across cases, we were able to generate a smaller number of selective codes.

Workstream 2: text mining, analysis and presentation of data

The main task in WS2 was to explore the feasibility of automatic analysis of free-text comments to identify themes (also known as topics, aspects or categories) and associated sentiments (positive, negative or neutral). This WS was divided into four steps:

1. manual coding of data with labels to support the design of text-mining methods that are based on searching for specific labels
2. development of text-mining programs to identify themes and sentiments (whether themes are positive or negative)
3. identification of comments that are representative of the labels assigned to them
4. creating report templates to present the analysed data.

To support this WS, we obtained two data sets containing free-text comments extracted from various patient experience surveys (e.g. FFT, Picker survey). The raw data set from site A contained 110,854 comments (2,114,726 words) and the site B data set contained 1653 comments (50,177 words).

Manual annotation of data

The aim of this step was to produce a high-quality, manually analysed data set that provides examples of free-text comments and associated themes and sentiments. After an initial manual inspection of a sample of the available data sets, we focused on nine themes (waiting time, staff attitude and professionalism, care quality, food, process, environment, parking, communication, resource) and two additional classes ('not feedback' and 'other') (see *Appendix 1, Table 19*, for a description of each theme).

A stratified random sample of comments (a roughly equal number of comments across their associated original Likert scales) was extracted and subsequently manually labelled as a 'gold standard' data set with associated themes and sentiment by up to five researchers. *Table 3* provides the basic statistics about the coded data sets.

Three of the researchers carrying out the coding tasks came from a social science background and had extensive qualitative research experience; the other two researchers were computer scientists. All annotators contributed to defining the annotation guidelines based on discussions about the content of the data and discussions with staff during initial meetings in the study sites providing the data (see *Appendix 1, Table 19*).

Initially, the researchers independently coded a small common set of comments, which were subsequently reviewed and the inter-rater agreement assessed (see *Chapter 4*). As a single comment could contain multiple themes, each with potentially different sentiments, the researchers were asked

TABLE 3 Manually coded data set statistics (the gold standard)

Characteristic	Site	
	A	B
Total number of comments	408	727
Total number of words	12,581	26,145
Total number of unique words	1924	2716
Total number of sentences	732	1648
Minimum number of sentences per comment	1	1
Maximum number of sentences per comment	11	18
Shortest sentence length	1	1
Longest sentence length	136	113
Average sentences per comment	1.79	2.27
Average words per comment	30.84	35.96

to identify text fragments that expressed a single theme (referred to as ‘segments’) and assign a theme and sentiment to them (see example in *Box 1*). The process was repeated on new samples until it was noted that there were no significant changes in the level of agreement.

Subsequently, the main coding task of labelling the data commenced. Owing to low numbers of comments attributed to some themes, relatedness across themes (e.g. similarities between parking, food and environment) and lower inter-rater agreement (see *Chapter 4*), we merged a number of themes to provide a final set of five themes (staff attitude, care quality, waiting time, environment, other) (see *Appendix 1, Table 20*). In addition, we merged negative and neutral sentiments following feedback from the clinical staff. The final data set used to develop and validate our text-mining programs included the merged themes (see *Appendix 1, Table 21*, for the final theme distribution in this data set).

Text-mining methods for theme and sentiment identification

We designed, developed and validated two text-mining programs based on the coding labels developed during the initial coding phase (known as supervised machine-learning methods). A third system combined outputs of the programs using confidence thresholds based on which performed the best analysis.

Segmentation-based model

The segmentation-based model (SBM) focused on the main observation that each comment may have several segments that refer to different themes. A model is therefore trained at the segment level and provides predictions at the segment level. This means that the program splits each comment into segments prior to being labelled according to themes and sentiment. Comments were also split into individual sentences before segmentation. *Appendix 1* shows the overall workflow (see *Figure 11*) and defines the technical components and stages (see *Box 2*) for this model.

Comment-level model

The comment-level model (CLM) takes segments as input but the predictions (themes) are made for the whole comment. The system was trained using the one-against-all approach (i.e. one classifier per topic). *Appendix 1* shows the overall CLM workflow (see *Figure 12*). We experimented with several different ways of classifying the comments (see *Appendix 1, Box 3*, for technical details).

Integrated system

The outputs of the SBM and CLM systems were combined using confidence thresholds. Each system has a separate set of confidence thresholds determined by cross-validation (see *Chapter 4*), which corresponds to a given theme (i.e. five thresholds for each system). The assigned themes are combined using a union operation applied to the outputs with confidence values higher than the predefined threshold.

Identification of representative comments

One suggestion from the staff teams was the inclusion of representative comments to provide examples of both negative and positive sentiments to help identify areas of improvement for the former and to highlight what works well in practice for the latter. The confidence values assigned by the classifiers were used to determine a representative set of comments for each theme, consisting of comments with the highest confidence values. This was combined with additional criteria (explained in the following section) to present representative feedback.

BOX 1 Illustration of segments of texts labelled by theme and sentiment

[The service I got was really good.] [The staff was helpful and understanding.]

(Care quality, Positive)

(Staff attitude and professionalism, Positive)

Text in ‘[]’ indicates the segments. The associated theme and sentiment are provided below within ‘()’.

Creating report templates to present analysed data

Following classification with the machine-learning techniques and the association of each comment with a number of themes and associated sentiments (step 2), the data sets were organised for analysis and visualisation in Stata® version 14.1 (StataCorp LP, College Station, TX, USA). We first conducted within-team discussions and meetings with data scientists from site A to inform on the aspects of the data that would be useful and informative but not burdensome for patients, health workers and managers. The biggest challenge appeared to be meeting the disparate needs of these different user groups in a single report, especially considering their various levels of statistical literacy. The consensus was to not compromise on the quality of the information, even if some aspects could be challenging to comprehend. The outputs were primarily directed at managers, although other user groups could find most aspects of the outputs informative, especially if the more advanced aspects were accompanied by explanations. It was also agreed that certain themes, such as 'other', were not informative and should not be reported.

Descriptive statistics were used to evaluate the distribution of the number of topics within comments and sentiments within topics. The focus of the analysis was visualisation, and numerous graphs were generated for inclusion in reports. Each comment was also linked to a self-reported overall satisfaction score. After reviewing these in team discussions, three cross-sectional (i.e. last time point) and two longitudinal graphs were selected for inclusion:

1. a bar chart of volume by topic and sentiment (cross-sectional)
2. a pie chart of volume by topic for negative sentiment only (cross-sectional)
3. a weighted scatterplot (on volume) by satisfaction score and sentiment (cross-sectional)
4. the volume of comments over time by topic (longitudinal)
5. the percentage of negative comments over time by topic (longitudinal).

In addition to the graphs, the team considered the inclusion of representative positive and negative comments. Because of size constraints, priority was given to negative comments. To deliver this, an algorithm was developed that selected comments on certain (editable) criteria and exported them in LaTeX format (*.tex) so that they could be incorporated into reports automatically. The criteria were (1) inclusion of comments in the last 3 months only, (2) ranked on the lowest satisfaction rating so comments with the worst satisfaction scores would be selected, (3) organised by topic but limited to professionalism, care quality and communication because it was felt that other aspects were less likely to be directly relevant to health professionals (e.g. waiting time) and (4) a maximum of five comments per topic, to limit the size of reports to two pages. All of these criteria are easily editable to provide tailored reports.

The final step involved creating a report template in LaTeX that used the exported five graphs and the *.tex comments file, to generate a report automatically. Each step of this process is automated in code (see *Appendix 6*) and requires almost no knowledge of the packages used, that is, Stata and LaTeX (e.g. MiKTeX and TeXnicCenter).

Workstream 3: co-design of a toolkit

For this WS, an experience-based design approach was adopted, drawing on the initial qualitative research (WS1), the computer science work (WS2) and insights from our PPI group, to co-design ways to enable and support digital data capture, analysis and use of both quantitative and narrative data.

These methods have been used to formulate a toolkit comprising:

- a survey utilising the FFT with space for free-text comments to be completed using digital kiosks in study sites, a website or a pen and paper version (see *Report Supplementary Material 1*)
- guidance and information for staff, patients and carers to support use of the new tools
- new text-mining programs for analysing patient feedback data
- new templates for reporting feedback from multiple sources
- a new process for eliciting and recording verbal feedback in community mental health services.

Qualitative data collected from staff, patients with MSK and service users with SMI in WS1 were summarised and discussed in follow-up focus groups in WS3, with 57% of staff and patient/service user participants having previously taken part in the qualitative components in WS1 (Table 4). Of note, no carer participants were recruited in either site A or site B for WS3.

Collectively, 65 participants contributed to WS3, with 37 participants having previously taken part in WS1. We conducted one follow-up focus group with staff participants in each site (site A, $n = 10$; site B, $n = 12$; site C1, $n = 9$; site C2, $n = 14$), individual interviews with patients with MSK ($n = 8$) and one focus group with service user participants ($n = 12$). These discussions enabled us to define priorities for capturing and using patient experience data. Using the new reporting templates, the extracted data were presented to participants to facilitate integration and the contrasting of data with data from other sources.

We achieved a diverse sample of staff participants with roles in management and patient experience, clinical practice and IT, reflecting the various categories of interest in each organisation. In site A, there were 10 staff participants (lead IT/governance/patient experience managers, $n = 2$; specialist nurses, $n = 2$; consultant specialists, $n = 6$). In site B, there were 12 staff participants (team managers, $n = 2$; care co-ordinators, $n = 3$; social workers, $n = 3$; community psychiatric nurses, $n = 3$; support workers, $n = 1$). In site C1, there were nine staff participants (GPs, $n = 2$; health-care assistants, $n = 1$; practice managers, $n = 1$; administrators/reception staff, $n = 5$). In site C2, there were 14 staff participants (GPs, $n = 6$; receptionists, $n = 2$; practice nurses, $n = 3$; pharmacists, $n = 1$; practice managers, $n = 1$; finance managers, $n = 1$).

We developed summaries of findings from WS1 and used Microsoft PowerPoint® 2011 (Microsoft Corporation, Redmond, WA, USA) slides to present these findings and trigger discussions on what tools might be best suited to improving the collection and usefulness of patient experience data for service improvement (see *Report Supplementary Material 2*). Figure 3 provides examples of summaries discussed in staff focus groups.

TABLE 4 Participants by sites: WS3

Participants	Site				Total
	A	B	C1	C2	
Staff focus groups	10 (5)	12 (5)	9 (6)	14 (7)	45 (23)
Staff interviews	0	0	0	0	0
Total staff	10 (5)	12 (5)	9 (6)	14 (7)	45 (23)
Patient focus groups	0	12 (7)			12 (7)
Patient interviews	8 (7)	0			8 (7)
Total patients	8 (7)	12 (7)			20 (14)
Carer focus groups	0	0			0
Carer interviews	0	0			0
Total carers	0	0			0
Total	18	24	9	14	65 (37)
Note					
The numbers of participants who had taken part previously in WS1 are indicated in parentheses; data from patients/service users and carers in sites C1 and C2 were not collected for WS1.					

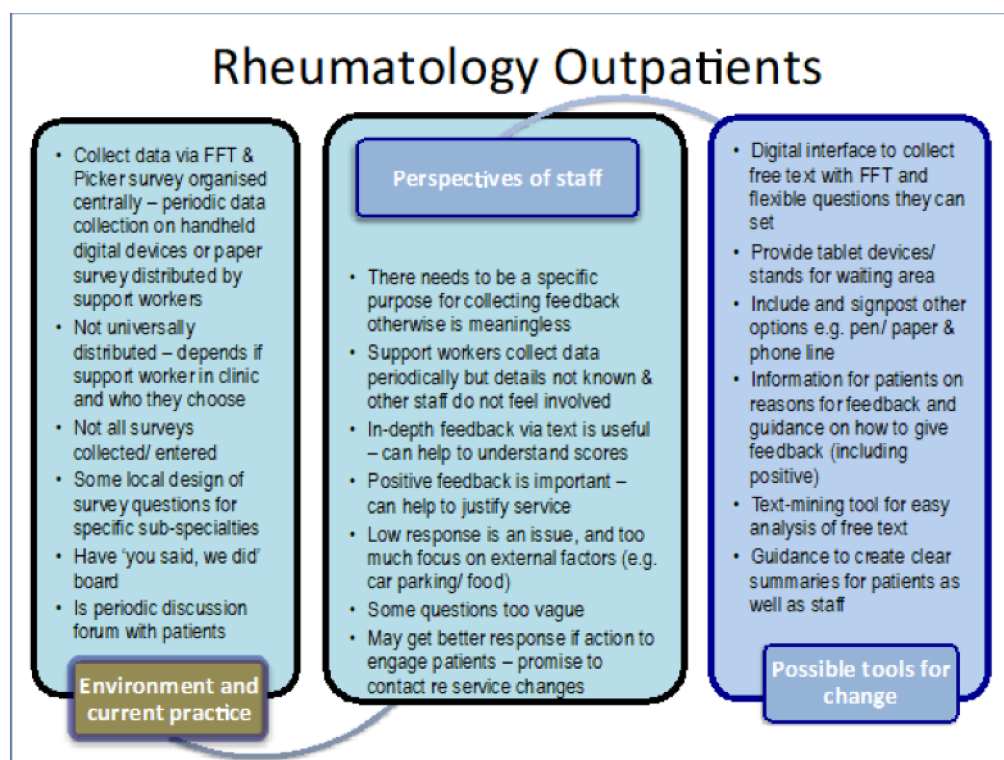


FIGURE 3 Slide from presentation summarising the links between current context, views of staff and possible tools for change in site A.

In addition to summarising the existing processes and the perspectives of staff, we also summarised the perspectives of patients in each of the sites for discussion with staff, alongside various ways of capturing and presenting the data based on analysis undertaken in WS2 (Figures 4 and 5).

Workstream 4: implementation and evaluation

Implementation

To implement the new tools into the four sites, a question-and-answer document was devised for each site and circulated to all staff members. Additionally, members of the research team visited each site and conducted an introductory session to describe the finalised toolkit and describe how the different components would be tested during the evaluation period.

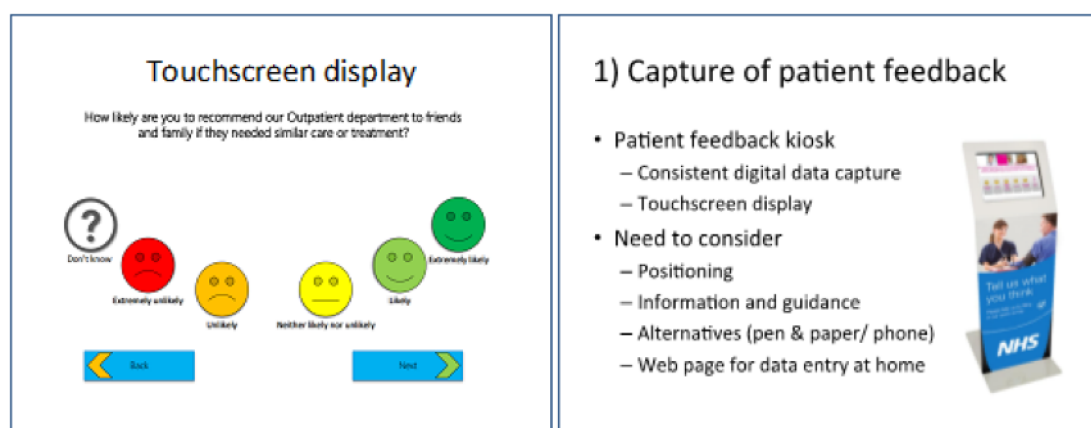


FIGURE 4 Examples of slides used to show participants various ways of capturing the data.

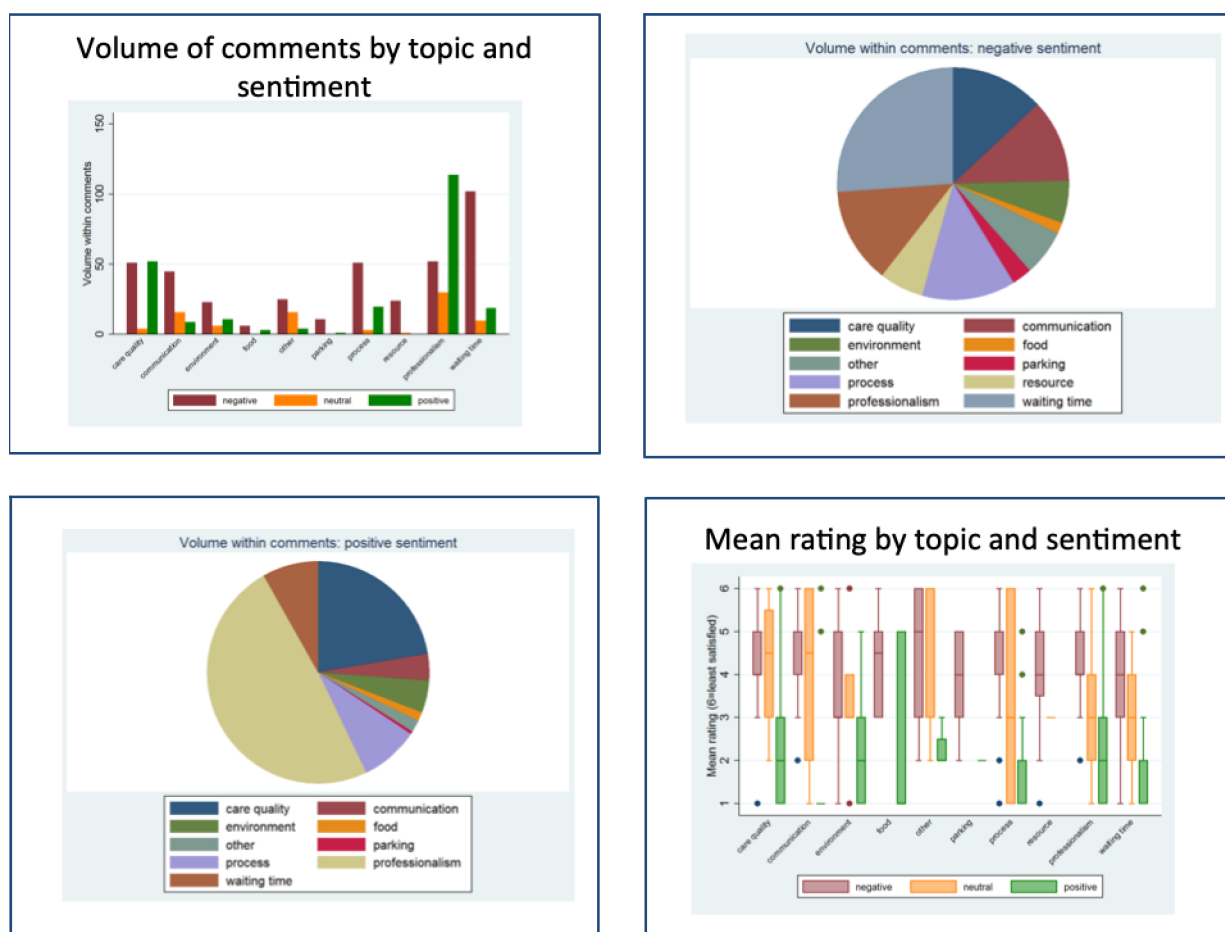


FIGURE 5 Examples of slides used to show participants potential graphs for displaying analysed data using text mining.

Qualitative evaluation

A process evaluation⁵⁰ was conducted using qualitative methods to assess how the tools were used in practice. This included using interviews and focus groups to understand the perspectives of patients and staff on the new tools. Observation sessions were also conducted in each site to determine the degree to which patients, service users and carers automatically approached the self-standing kiosk (or when the kiosk was manned by a volunteer) and typed their feedback using the touchscreen or wrote feedback on a paper survey.

We undertook at least one follow-up focus group with staff in each site (site A, $n = 5$; site B, $n = 19$; site C1, $n = 8$; site C2, $n = 8$) and individual interviews with staff in three sites (site B, $n = 7$; site C1, $n = 2$; site C2, $n = 2$) (total staff, $n = 51$) (Table 5). Of note, 35 WS4 staff participants had taken part in previous components of the study.

We also undertook focus groups with patients in two sites (site B, $n = 13$; site C2, $n = 4$) and individual interviews with patients in two sites (site A, $n = 6$; site C2, $n = 1$) (total, $n = 24$). Additionally, we facilitated focus groups with carers in two sites (site B, $n = 2$; site C2, $n = 1$ – this carer participant took part in a focus group with patient participants) and undertook individual interviews with carers in two sites (site A, $n = 1$; site B, $n = 4$) (total, $n = 8$).

Of note, of the participants included in this phase ($n = 83$), 47 had taken part in the initial focus groups and interviews in WS1 and/or WS3 (see Table 5 for the numbers of participants who had taken part in

TABLE 5 Participants by sites: WS4

Participants	Site				Total
	A	B	C1	C2	
Staff focus groups	5 (5)	19 (8)	8 (5)	8 (6)	40 (24)
Staff interviews	0	7 (7)	2 (2)	2 (2)	11 (11)
Total staff	5 (5)	26 (15)	10 (7)	10 (8)	51 (35)
Patient focus groups	0	13 (6)	0	4	17 (6)
Patient interviews	6 (6)	0	0	1	7 (6)
Total patients	6 (6)	13 (6)	0	5	24 (12)
Carer focus groups	0	2	0	1	3
Carer interviews	1	4	0	0	5
Total carers	1	6	0	1	8
Total	12	45	10	16	83 (47)

Note

The numbers of participants who had taken part in previous components of the study are indicated in parentheses; data from patients/service users and carers in sites C1 and C2 were not collected for WS1.

previous components of the study). Crucially, these discussions enabled us to understand perspectives on how the new tools worked in practice and further explore some of the issues raised in the earlier qualitative research.

Observations were conducted to evaluate the barriers to and opportunities for providing patient experience feedback using the newly implemented methods of the standing kiosks, the website and the pen and paper version. Each observation episode varied between a minimum of 1 hour and a maximum of 3 hours. Staff meetings to discuss feedback during the evaluation period were also observed in the study sites (Table 6).

Normalisation process theory, which has been applied to implementation in varied health-care contexts,⁵¹ was drawn on for analysis. NPT focuses on social practices and interaction and is operationalised via four key constructs: coherence (meaning and understanding of new technology/practices), cognitive participation (relational work to sustain a community of practice for a new intervention), collective action (operational work to enact new practices) and reflexive monitoring (work carried out to monitor and appraise new practices). First, we coded the data using a modified grounded theory approach⁵⁸ and then mapped the emerging themes against NPT constructs.

Quantitative evaluation of pre- and post-participation rates

Response rates for patient experience questionnaires and levels of participation pre and post implementation of the toolkit were compared to investigate the impact of the toolkit for widening participation levels in multiple settings.

TABLE 6 Number of observation sessions by site: WS4

Observation session	Site				Total
	A	B	C1	C2	
Patients	11	8	11	8	38
Staff meetings	1	1	1	1	4

Health economics

The planned health economics work, to develop a decision model to explore the potential costs and benefits of the toolkit/kiosk approach, was modified as the programme progressed. In particular, the uncertain evidence about the benefits of the FFT for generating quality improvements and changes in local trusts, combined with recent challenges about its worth and the results of the toolkit/kiosk evaluations, limited the feasibility and value of such an exploratory model. Accordingly, we focused instead on estimating the costs of developing and implementing the toolkit developed in this programme to inform future development and evaluation of this or similar tools. Our research questions for this part of the study were:

- What are the costs of co-design activities to develop the toolkit components?
- What are the costs of developing the text-mining and reporting elements of the toolkit?
- What are the costs of initial implementation of the toolkit/kiosks in each of the sites?
- What are the costs of analysing and reporting the data generated by the toolkit/kiosks?

The costs of developing and implementing the toolkit were estimated from data on the staff time and resources used in this programme to develop and implement the toolkit in the different sites. The staff time and resource use data were collected from diaries and other records of activities of those involved in developing and implementing the toolkit, as well as from the qualitative interviews and observational studies. The staff time data also included the time spent developing and implementing the analytic approaches used to analyse the data generated by the toolkit and the FFT data available to the team.

National salary scales were used to estimate the costs of university research and NHS staff whereas the payments made as part of the evaluation were used to cost the time of volunteers and non-staff participants in the co-design activities. The costs of equipment and consumables used as components of the toolkit and used to analyse the data were estimated from actual expenditure.

In addition, we reviewed published literature and Department of Health and Social Care policy and guidance to identify the costs of the FFT at the local level. Searches of electronic databases and the Department of Health and Social Care website were conducted for all years up to December 2017, using a simple electronic search strategy. These were updated to March 2018. The electronic databases searched included the Cochrane Database of Systematic Reviews; ACP Journal Club; Database of Abstracts of Reviews of Effects; Cochrane Central Register of Controlled Trials; Cochrane Methodology Register; Health Technology Assessment; NHS Economic Evaluation Database; Allied and Complementary Medicine Database; EMBASE, Health Management Information Consortium (HMIC); Maternity and Infant Care Database (MIDIRS); MEDLINE; and PsycINFO. The search terms included 'friends and family test', 'patient experience' and variants of each term. Initial searches excluding terms related to cost and economics indicated that a low number of studies were available ($n = 36$). Accordingly, the cost-related terms were not included in the electronic search, but were included in the inclusion criteria used to screen full studies. Only studies reported in English were included. The full papers for all of the included studies were obtained. These were reviewed and any resource use or cost data were extracted by one researcher (LD).

Text mining versus qualitative analysis of free-text feedback received in general hospital and mental health service settings: a descriptive comparison of findings

This study set out to compare the findings produced by text mining against those produced by qualitative researchers working on the same data sets. The same data sets were analysed in an independent and blinded fashion by machine-learning algorithms (as described above) and by 'human analysis' using grounded theory coding. A secondary aim was to compare feedback gathered in mental health settings with that gathered in general hospital settings, using both the same and different analytic methods.

The data sets

The data sets used were a subset of those employed in the text-mining work described in WS2. Analysis was conducted on general hospital trust (site A) data for 1 complete calendar month (June 2016, the most recent complete month for which data were available). Using only 1 month of mental health data would have yielded only around 40 comments, so 6 months of data were used for site B.

Data analysis

The qualitative analysts used 'open coding' principles of grounded theory analysis by Corbin and Strauss,⁵⁹ enabled by the use of NVivo 11. The grounded theory approach [hereafter referred to as adapted grounded theory (AGT)] was 'adapted' or 'expanded' in that count data on categories (or 'topics', as in the text-mining analysis) were also collated and described and were subsequently organised by sentiment to facilitate comparison with the text-mining results. The qualitative researchers began by independently coding a sample of around 500 comments received by both health trusts. They then compared results and used this discussion as a basis for drawing up a draft list of around 75 codes in 10 preliminary categories.

The final coding framework consisted of 125 codes (or 'child nodes') in eight nodes (or 'themes'): access process and discharge, communication from and with clinical staff, positive aspects of service, specific complaints, qualified comments, staff attitude, the service made me feel and this service is better than others.

Ethics and consent

Ethics approval was granted by the National Research Ethics Service West Midlands – Black Country Research Ethics Committee (REC) (reference number: 16/WM/0243).

All staff, patient and carer participants were given a participant information sheet explaining the study and provided written consent to take part in a focus group discussion, an individual interview or both, dependent on participant preference (see *Appendix 2* for an example of an information sheet for staff). The researcher(s) stressed the voluntary nature of participation in the study (see *Appendix 2* for an example of a consent form used in the study).

During the study, the research team, staff and patients reflected on a number of ethics issues related to the collection of feedback. Staff in multiple sites reflected on concerns that patients might feel pressured to give positive feedback if they were encouraged to use the digital kiosks and that there may be concerns that the feedback would not be anonymous. In response to this, care was taken to ensure that information given to patients made it clear that giving feedback was their choice. They were also reminded of the multiple other ways by which they could give feedback.

Although the patient experience data obtained from sites A and B were classified as anonymised data, the free text sometimes contained identifiable details (e.g. names, telephone numbers). Care was taken to ensure that data were transferred and stored securely in line with our NHS and Health Research Authority ethics approval and that only anonymised data were used to illustrate themes from analysis.

Within the community mental health team the ethics issues regarding the recording of verbal feedback in the professional–client clinical visits were discussed actively during focus groups and interviews. Many drew links with the conversations about care that were a normal part of care-planning processes, but also identified that the main concern would be to remind service users of the different routes by which they could give formal feedback outside clinical relationships with staff.

Our researchers followed the University of Manchester lone worker policy when visiting both sites and participants' homes.

Chapter 3 Patient and public involvement

Introduction

National policy is increasingly encouraging PPI in research as a means to improve both the relevance and the meaningfulness of applied health research in England. In this chapter, we describe our PPI work for the study. From the outset, we asked our PPI contributors to be actively involved in the conception of the Developing and Enhancing the usefulness of Patient Experience and Narrative Data (DEPEND) study, the investigation and the evaluation and dissemination of the results, bringing insight into each WS because of their experiences of living with a long-term condition and using health services.

Patient and public involvement contributors

We created a core PPI advisory group to represent a range of experiences and preferences concerning feedback among our exemplar physical and mental health long-term conditions. In total, eight members of the public with one of our exemplar conditions joined the core PPI advisory group. Three members had experience of a MSK condition and five members had experience as users of mental health services. Two members represented a dual perspective, as a carer and a service user. All members were paid for their time at INVOLVE rates and claimed additional expenses where applicable.

There were two PPI leads for the core PPI advisory group (AL and NS), who were also co-investigators in the study and had worked with us from the initial planning phase on the design of research questions and plans for the pre-funded proposal. They also had input into the management of the research by attending project management group and study steering committee meetings.

Members had various experience of PPI research work to date. At the start of the study, all of our PPI advisors had experience of working with researchers as PPI advisors. Two members were previous consultants of a PPI advisory group and one member had undertaken PPI research training developed in-house in another department at our institution.

Patient participation groups

We also drew on the perspectives of volunteers of two patient participation groups (PPGs), which are groups set up to enable public participation in primary care organisations. PPGs hold regular meetings to discuss the services provided and how improvements can be made for the benefit of patients and the practice [National Association for Patient Participation; see www.napp.org.uk/ (accessed 14 December 2019)], and attempts should be made to ensure that such groups are representative of the practice population. We found each PPG to be quite different in terms of the group composition, management and levels of engagement with the DEPEND study. The PPG in site C1 is a relatively small group with nine PPG members; four core PPG members worked with us as PPI contributors to the DEPEND study. Virtual PPG participation is enabled at this site through the practice website, but we engaged only with those PPG members who attended face-to-face engagement PPG meetings on site. Of note, this site faced challenges in trying to get a core PPG together. It was eventually decided that a random cross-section of patients would be asked their opinions regarding the practice and to participate as a PPG member by completing an online patient group form. Currently, the assistant practice manager manages the PPG, circulating the minutes from each face-to-face meeting and uploading the minutes to the practice website. The practice manager liaises with members mainly by telephone; only a few members have an active e-mail address.

In comparison, the PPG in site C2 is considerably larger, with 40 members; we worked with seven active core PPG members. Virtual participation is also enabled at this site through the practice website but, again, we engaged only with those PPG members who attended face-to-face PPG meetings in the site C2 reception area. A trainee GP set up this PPG with the practice manager when she was a GP registrar at the practice in 2011.

Meetings

Feedback from the first PPI meeting focused on further clarity about the purpose of our core PPI group and highlighted the importance of establishing ground rules for each meeting. As a result, a 'terms of reference' document was circulated and discussed at the second PPI meeting.

To optimise PPI involvement, separate PPI groups were held, focusing on each long-term condition of interest, SMI or MSK conditions. During the 2-year study, the PPI group members met face-to-face four times within the Centre for Primary Care. We also facilitated a total of five face-to-face PPG meetings per site (C1 and C2) during each WS. The PPG at site C1 met upstairs in the practice manager's room during the early evening after practice hours; the PPG at site C2 met in the seated reception area over a lunchtime when the practice was closed.

We accommodated PPI representatives in terms of access and health. For instance, we enabled virtual participation via Skype™ (Microsoft Corporation) to reach PPI contributors who could not travel. Nicola Small took overall responsibility for the co-ordination and management of the PPI group, with input from the chief investigator (CS). Each PPI face-to-face meeting was chaired by the principal investigator (CS) and co-facilitated by Nicola Small and Papreen Nahar, with refreshments provided at each meeting.

Our PPI group met face-to-face, as well as remotely, for specific input on the content of each WS, the co-design of the toolkit components and accompanying bespoke guidance and the evaluation and dissemination activities. Each PPI meeting was constructed around each WS of the project and comprised an overview of the minutes from the previous group, a PowerPoint presentation consisting of a recruitment and project update with key findings to date and a PPI-led discussion. We made sure that what our PPI contributors had advised in relation to progress to date was always summarised in our slide set. Project documents were provided ahead of each meeting and hard copies were available on the day. Following the meeting, minutes were circulated by the co-ordinator, alongside PowerPoint slides. In between the core PPI meetings, regular communication took place by e-mail to the PPI co-ordinator (NS) and core documents were read and commented on using Dropbox (Dropbox Inc., San Francisco, CA, USA), by e-mail or face-to-face.

Some challenges and lessons learnt

During the course of the project, we faced some major challenges and upsetting events associated with our PPI work. Sadly, two members of our team died unexpectedly: Neal Sinclair and Jane Reid Peters brought their energy, enthusiasm and valuable experiences and expertise to our work. It was a shock to both researchers and fellow PPI group members when they died within a relatively short period of time. This made us reflect on the relationships developed during the course of carrying out PPI work within research and on how to manage difficult situations. Working closely together as researchers and PPI contributors entails sharing a lot of personal experiences and building long-term relationships, which are quite different from those that develop when researchers are carrying out one-off interviews or focus groups. This can result in researchers, and also PPI members, having a sense of responsibility towards fellow team members. When a member of a group dies it may be difficult to know what to do to support other members of the group, who may already feel vulnerable. In our case, we found a way to manage and support each other during an upsetting time, but we found it difficult to find information

about similar circumstances and what might help. We also found that there is little formal support available unless it is explicitly planned for. For example, we thought that it might help other members of the PPI group if they were able to access a counselling service easily should they need it. This is difficult because universities, for example, provide counselling for staff and students but not for PPI members. This has prompted us to work with our Centre for Social Responsibility to build better support for managing difficult situations associated with PPI. Caroline Sanders has been working with the NIHR-funded Primary Care in Manchester Engagement Resource (PRIMER) group and the NIHR-funded Greater Manchester Patient Safety Translational Research Centre (GM PSTRC) to develop this work further. We hosted a workshop on this topic in February 2018 [see <https://gmpstrc.wordpress.com/2018/02/01/managing-difficult-situations-in-patient-and-public-involvement-workshop-event/> (accessed 9 October 2019)].

Our experience has taught us that fellow PPI group members may need additional support when someone in the group dies; researchers may also need support and it is also important to follow up with family members to ensure that they know how much the work carried out by the PPI member has been valued.

Participant experiences

As the PPI co-design meetings progressed, one of our PPI members (DA) became more involved in our recruitment activities. She was able to help recruit carer participants within site B using her existing carer networks. During year 2 of the study she was also able to spend dedicated time working in our centre and was provided with an IT account to enable her to contribute to core project documents. Our PPI co-investigator (AL) also spent dedicated time reviewing components for the toolkit during the co-design phase of the project and we were able to supply her with an iPad mini (Apple Inc., Cupertino, CA, USA) to enable her to complete some of this work remotely.

We had originally planned for both PPI co-investigators to work with us in conducting interviews and focus groups. However, this was not possible because of health problems and other constraints, limiting the time available for the training and governance requirements to enable this.

Results

In this section we present feedback from our PPI group for our four WSs: (1) topics to explore at interview to enhance current feedback collection methods, (2) text mining, analysis and presentation of narrative data, (3) co-design of a suite of tools and (4) development, evaluation, implementation and dissemination of the toolkit. In addition, the recommendations that were developed as a result and how they were acted on are also presented.

Topics to explore at interview to enhance current feedback collection methods

We sought general feedback from our PPI group on current feedback collection methods.

Our PPI SMI group felt that the importance of collecting positive feedback needs to be emphasised and that there should be a range of methods available for giving feedback, either through structured survey questions or narratively. Additionally, the mode of feedback should be considered to encourage wider participation. Privacy may be facilitated by the use of a booth and iPads, video and audio. Simple emoticons could be used for giving feedback, with an option for adding brief text, alongside survey questions. Further, a 'feedback period' might be used to encourage participation in each method and waiting time could be used more effectively for collecting feedback. When feedback is based on previous experience, or expectations, this might be documented in a 'gratitude diary', with examples of positive feedback collated. Finally, feedback co-ordination was emphasised by the group as being essential for 'linking up the services' and ensuring that all professionals are aware of the full patient experience.

It was thought that a 'feedback incentive process', such as the Amazon (Amazon.com, Inc., Bellevue, WA, USA) virtual system of allocating stars to services, might increase interest in prioritising feedback.

The importance of collecting 'relational feedback' from the therapeutic relationship was favoured by our PPI MSK group. PPI contributors also felt that 'medication dosage feedback' should be targeted, as well as having a record of how patients feel. Further, the type of feedback was felt to be important for eliciting more meaningful data. The mode of feedback was seen as key to having access to the feedback data; having online access was a popular choice, but contributors had never tried that mode, having only haphazardly given written feedback. Overall, this group felt that having the option to record personal experiences of NHS services might help staff to tailor care.

Likewise, both PPG groups felt that patients and carers are currently not aware of what patient experience feedback is given, and having the option to capture patients' stories in relation to care experienced might provide more meaningful feedback. Finally, both PPG groups felt that current feedback is far too structured and generic and needs to incorporate real-time feedback methods to invite wider participation.

As a result of this feedback, we recommended that our topic guides at interview should include the aforementioned topics to probe with patients, carers and members of staff to gain perspectives on current feedback collection methods.

Text mining, analysis and presentation of narrative data

In relation to separating feedback by specific topic, our PPI MSK group liked the idea of breaking down the 'staff attitude and professionalism' category further and thought that clinicians should be identified in the feedback report to get the recognition that they deserve (see *Chapter 4*). The idea of using sentiment analysis was viewed as a useful method to pick up common concerns to feedback to staff in order to inform service improvement. Our PPI SMI group thought that the 'word cloud idea' used to visually present all of the positive and negative comments together was a bit like a Google (Google Inc., Mountain View, CA, USA) search, which brings up the keywords associated with a word. PPG members varied in terms of how they preferred feedback data to be presented. Many thought that the pie chart presentation of the results of the aggregated feedback data was confusing and that the theme 'staff attitude' was the same as the theme 'staff professionalism'. Others viewed the pie chart positively, with the segments described as 'simple and useful', although a bar chart 'would also do the same job'. Overall, there was collective agreement that a breakdown of the rates of participation to illustrate who was using which method to feed back, alongside a narrative summary of the core themes of positive and negative experiences of services, would be useful for individual teams.

As a result of this feedback, a prop was co-designed with our PPI group to elicit views on giving summaries of feedback to staff and patients and carers.

Co-design of a suite of tools

The co-design process that we followed is described fully in *Chapter 2*. In brief, we collected PPI insights on which interview props (screenshots of our suite of tools) could be used to make patient and carer participants understand better our proposed tools. The toolkit mock-up can be found in *Appendix 3*. Our ideas to take this feedback forward are discussed in turn.

Digital capture of positive and negative comments via kiosks and tablets and online

Our PPI group felt that being able to record one good experience and one bad experience would be a good way of structuring feedback in relation to the mock-up touchscreen display. Both PPGs liked the idea of using a digital touchscreen to type feedback; the PPG in site C2 was particularly enthusiastic to test this tool as touchscreens were already used to check in for appointments. However, the PPG warned that we should use a short survey as they knew that people had completed short surveys previously. It was also collectively agreed that the FFT question probing for positive and negative

feedback was an innovative idea, but members asked if we had thought of using 'screen 1 of, say, 4' to show patients that it was not a long survey and asking patients to 'press next' if they did not have a comment to make. It was also thought that feedback should be given at the end of the questionnaire, which could be instantaneous: 'if you are ranking the service it would be good at the end if it said, for example, in the last week 85% of people were likely or extremely likely to recommend the practice'.

We received mixed PPI feedback on our proposal to use emoticons to collect feedback via the FFT captured digitally. A few MSK PPI contributors did not like the idea of using emoticons to digitally record an emotional response option as they felt that it would be inappropriate for a patient with physical pain to describe their health experience by pressing an expression button; the alternative traffic light system was preferred. However, the remainder of our PPI group liked the interface using emoticons; the PPG at site C2 particularly liked the concept of emoticons, but wondered if we could offer four response options rather than the five shown, 'as people tended to go for the middle ranking whereas four makes the respondent go one way or the other'.

Having access to digital and non-digital tools

Our PPI group felt that we should not rely on one type or method of feedback as different people respond differently to different methods. The issue of anonymity was a core topic during this discussion, because people might not feel able to give honest feedback in front of a clinician. Using digital methods to collect feedback might allow a 'feedback period' to be implemented, encouraging users to give both pre- and post-consultation feedback. Our PPI group liked the idea of using SMS messages to collect feedback, but the timing of these messages needed thought. One MSK PPI contributor explained how she had received a SMS message feedback alert from site A that came through very late at night, which made her feel anxious. Further, a PPG member said that she had received SMS messages from site A asking for feedback, but found them repetitive, despite liking the option of being able to choose a feedback method.

One PPG liked having the option of using iPads to house short surveys and queried if a short online survey could be added to the Patient Access website (www.patientaccess.com; accessed February 2020) or an iPhone (Apple Inc.) application, as both methods are becoming popular to book appointments and order repeat prescriptions. A further suggestion was that a URL be provided, or text message asking for feedback, routinely to patients after their appointments. It was agreed that no one method would catch everybody and that many people 'didn't have time'. Another PPG contributor queried if anybody would be excluded from giving feedback and how we would manage this. Two MSK PPI contributors liked the idea of having a 'dedicated telephone line' to give feedback, but would prefer a person on the other end of the telephone and not just an answering machine, as this would not encourage the type of feedback sought.

Feedback privacy, confidentiality and location

Digital feedback collected in a booth to remove any confidentiality and privacy barriers was discussed by our PPI group. However, our MSK PPI group did not like the touchscreens that are currently used in general practice surgeries to collect appointment information, as typing in personal information made them 'feel uneasy'; they also agreed that iPads are not private. One PPI contributor felt that it was important to have up-to-date feedback on the noticeboards of clinics/reception areas to encourage an environment of trust and respect so that people felt able to feed back honestly to services. Overall, there was collective agreement from our PPGs that privacy and confidentiality are equally as important to encourage people to feed back to services. One PPG contributor (site C2) liked the idea of having a private cubicle (like a passport photograph cubicle) to give digital feedback. Another PPG contributor (site C1) described taking part in a trial, in which you log on to a website to input your diabetes values, and liked this method; feedback was linked to a unique identifier (ID) and the method was easy to use. Overall, our PPI group advised that the digital kiosk would be best sited in the relevant reception and clinical areas, but may need to be moved around during the testing WS to find the best position.

Toolkit guidance and information

It was agreed by our core PPI group that it would be a good idea for the guidance to indicate why feedback was important for improving services, because people do not know what happens to the feedback that is collected, which can affect people's motivation to give feedback in the first place. Our MSK collaborators felt there should be something on the walls of surgeries that outlines what happens to patient feedback and the outcomes of the feedback. Our PPG at site C2 advised us to have a 'headline grabber' on the guidance above the kiosk to advertise the toolkit, such as 'we need your feedback!'. The PPG at site C1 voiced how there is currently a 'negative feedback ethos' and 'no financial regulation of positive feedback'. Everyone thought that the 'you said, we said' examples were really helpful in the draft guidance to portray this key message, and that hearing what was done with the feedback was an important part of the information to accompany the toolkit.

This feedback resulted in us being able to tailor each set of tools and accompanying toolkit guidance to meet the requirements of each context and organisation.

Development, evaluation, implementation and dissemination of toolkit

At this final stage, our PPI group reviewed the final version of the new toolkit and the core project documents, which were co-designed and implemented across multiple NHS sites to aid the collection, reporting and analysing of feedback. The final toolkit components are presented in *Chapter 4* and *Appendix 3*. As part of the process evaluation we produced several other core documents to aid the implementation phase, which our PPI group reviewed. These are each summarised in turn in the following sections.

New templates for reporting feedback from multiple sources

We asked our PPI group to comment on the content and visual presentation of the feedback reports that we proposed to use, which contained the feedback collected by the new tools being tested at each site, tailored to context. Overall, these were well received and were clear and simple to follow. Our PPI group particularly liked the summary of the quantitative data (i.e. the FFT responses and the demographic data), followed by the narrative feedback split by sentiment and the core themes of patient experience. One PPG member thought that these reports, either handwritten or automated, should be displayed in the reception area to show that feedback was being noted. This PPG felt that feedback reports should be received more often than every 3 months; every 1–2 months may be more acceptable and useful to teams.

Working document and advertising

Our PPI group fed back on the progress of this phase by commenting on a working toolkit document, which outlined any tweaks to the tools made, as well as any technical problems encountered, and helped us to provide solutions in the field. We are particularly grateful to two PPG members who helped us to advertise the new tools at site C2 during the introductory period. To help with this, we co-designed with our PPI group a recruitment flyer for staff, researchers and PPG members to hand out to patients and carers on site to advertise the new tools. PPG members felt that the flyer and poster were equally useful for reminding patients and carers to give feedback using one of the methods available.

In the later months of the evaluation phase, we worked with a carer and a service user volunteer from site B, alongside the patient experience leads, on the co-design of a script for a video to show the different ways of giving feedback. Finally, we used different methods to communicate the key messages of the DEPEND study to lay audiences, including a PPI co-written DEPEND video with corresponding screenplay.

Dissemination

The PPI group worked together to co-produce a script for an animated film and advised on the developing images for the storyboard (see *Appendix 4*) with our commissioned producers of the film (James Munro and Matt Cook).

We have co-authored academic papers, co-presented at academic conferences and disseminated findings at a PPI dissemination workshop in spring 2018 (see *Appendix 4* for details).

Reflections

Soon after the evaluation of the toolkit ended, we conducted a meeting with members of our core PPI group to hear their views on the process and outcome of PPI in the DEPEND study. All of our PPI contributors commented on how much they had enjoyed being part of the team and that the experience had been rewarding. Overall, participating in the DEPEND study had taken various positive forms for our PPI contributors. Some examples are highlighted in the visual representation of the PPI model, developed by one of our lead PPI co-investigators in the DEPEND study, Dawn (see *Appendix 5*). Some barriers to participation in PPI activities included retention of PPI contributors and managing bereavements and difficult situations for both PPI collaborators and researchers, which were discussed during the project. The importance of diversity and inclusion in PPI was considered by our PPI group, as well as measuring fear or the stigma of becoming involved in PPI activities.

Overall, we all felt that both the process and the outcome of PPI in the DEPEND study was successful and that we too had forged strong relationships over the 2-year study. Our PPI collaborators made valuable contributions to, and provided valuable insights into, each WS, ensuring that our research priorities aligned with those of patients, service users and carers and enabling us develop recommendations for delivering future PPI work.

This chapter was co-authored with our PPI collaborators.

Chapter 4 Results

Workstream 1: perspectives of patients, carers and staff

Patients' and carers' views

Lack of understanding and experience regarding the collection and use of patient experience data

Generally, patients and carers did not recall that they had been asked for feedback. The purpose of feedback was not clear to the majority of patients and the perceived bureaucratic nature of feedback did not encourage patients to participate because they often interpreted it as a 'tick-box exercise'. This was then linked to the feeling that their comments would not be used. The idea of feedback leading to action was a persistent theme in the interviews and the following comments were typical:

From the provider's side they should explain more about the feedback and what they are going to do with it and then people will feel encouraged to fill out.

ID 127, site A

It depends on how it's acted upon really and I guess with things like this you've got to have some confidence that it's done for the right reason and not just box ticking.

ID 113, site A

Those patients who were inclined to fill in feedback questionnaires assumed that their feedback would be used to improve service delivery or otherwise would be viewed as 'a waste of money' (ID 109R).

The term 'feedback' is meant to be neutral, but patients perceived there to be a predominant assumption that feedback systems were in place primarily to allow people to provide negative feedback. This might be because of the wider societal context that appears to encourage a consumerist focus on complaints. The following quotation highlights this tendency:

You tend to think of maybe feedback it is a criticism almost ... So if something unexpected happened then people usually start thinking about feedback ... Someone who's not had a very good experience is more willing to want to give you feedback, because they want to tell you why they've had a bad experience.

ID 124, site A

The process of giving feedback was considered equally important and, again, a lack of clarity emerged as a concern. Patients said that they could not identify specific individuals to contact or that the information provided was insufficient. These factors inhibited participation and the following patient explained this:

Comparing with the private service ... whereas if I had a complaint with regards to, I don't know, something in the rheumatology I actually wouldn't quite know where to go. I'd go on to the [site A] website perhaps, and maybe it is all there and it is easy to do and you just start the process, but I don't know, and that if you could be autonomous about it perhaps that would be ... would make me more inclined to do that as well, so.

ID 128, site A

RESULTS

The implications of giving feedback were discussed and a number of concerns emerged, including some issues around anonymity. Although a number of people worried that their critical feedback might be linked to them directly, others indicated the need for transparency to enable complaints (or potential complaints) to be resolved:

I think people are a little bit worried that if they give feedback that doesn't, you know, naturally say that the National Health [Service] is brilliant and they might have something that's a problem that they've dealt with themselves, they might feel a little bit intimidated that they might not get the same treatment when they go back again maybe.

ID 121, site A

I think it's very important that they acknowledge patient feedback, because any issues that are brought up by patients can then be resolved, and you move forward from that.

ID 120R, site A

The direct link between feedback and action was emphasised to motivate people to put in the effort of providing feedback. The idea of 'you said, we did' was emphasised repeatedly:

If I see that it's worked, and feedback always works anyhow. It would be wonderful if I said at feedback, such and such could make it better, and if I came again and saw that it was, it would be fantastic, you know, you would see the fruits of your labour, sort of thing, yeah.

ID 129, site A

Overall, many patients and carers did not have a good understanding of why patient feedback was collected and how such feedback was used in health services. The experience of being asked to give feedback appeared to be patchy. Those patients who recollected that they were asked for feedback were doubtful that it would directly influence service delivery.

Staff views

In the acute trust, many staff members were largely unaware of the system for collecting and analysing feedback and consequently reported a sense of disengagement from this:

I'm not sure, but I know in outpatients the nurses are often filling in, [and] the support workers, some questionnaires that we give to patients. But I don't know what happens to the results of those.

Consultant, site A

I realise that I don't quite know the full range of what's collected and how from our patients. And in addition to what's collected in the waiting area last year there was a 1-day meeting that was organised to collect patients views. Where that information then went I'm also not too ... sure.

IT consultant, site A

General practitioners considered the feedback system to be a performance-monitoring tool, and one that may not be an effective reflection of the quality of performance in general practice. They also expressed concerns about the value of collecting anonymous feedback when this makes it difficult to consider the clinical appropriateness of treatment and the validity of any complaints:

People can be rude as they like, completely unjustifiably ... when we get a complaint and no one puts a name to it ... we're going to throw it in the bin because what does it mean? It doesn't mean anything ... if somebody just says, 'I'm so unhappy now, I'll never go and see that doctor again' and you don't put a name to it, how can I reply? Because I don't know whether it's somebody being malicious or, I've actually made a genuine error. I don't know the clinical circumstances so I can't look at my notes and see whether I was right or wrong. So anonymous feedback, I don't have much value in.

ID 141, GP, site C

My friends who are GPs are very worried about that, because you can get a very low score by refusing to give a patient the opioid medication they want. And what you're doing is clinically correct and it's the right thing to do, but if the patient doesn't like it, you get a low score just by doing your job correctly.

ID 104, GP, site C1

General practitioners overall were in favour of the use of meaningful and more detailed narrative feedback focused on their individual practice. Their main concerns with the formal ways of collecting feedback were that these data were too generic and selective to be meaningful. Several GPs thought that to make the feedback system meaningful there needed to be opportunities for patients and carers to provide narrative accounts of their experience instead of just 'box ticking':

There's stuff we get fulfils our tick-box approach, so it's fine for our ... appraisals, 'cause we get 40 questionnaires back that say we're wonderful. ... in many ways it's not desperately useful feedback. If we could get feedback that was constructive ... and we could use that in our ... developing the service, let's say the flu clinic ... you know, the new services we're offering, then that would be useful ... what we currently collect is box ticking.

ID 303, GP, site C2

Well it's getting enough information to actually make ... it to be meaningful. It's the quality of it as well to be meaningful and it's having the time, as you say, to look at it and analyse it and rely on it because otherwise you'd constantly be just jumping from one supposed problem to another supposed problem but you haven't got the right information to make any reasonable decision about it.

ID 104, GP, site C1

The level of understanding of and involvement in the system of collecting and using feedback varied considerably by role in each site. For example, in the hospital trust at site A, the IT staff and lead nurses were more actively engaged with collecting and using feedback. In particular, several nurse managers had overall responsibility for patient experience data and their use for quality improvement. This gave them extensive knowledge and understanding, as well as expertise in analysing and interpreting the data. Conversely, in the hospital trust at site B, the service user and carer engagement lead for the trust was solely responsible for the analysis, documentation and recording of thousands of entry and exit questionnaires across the trust, and this was carried out manually by him and his personal assistant:

Sometimes they've [entry and exit questionnaires] come in with a, with a stamp on them ... so we know that it's a particular service. If it's got a yellow stamp on it we know that it's the X team ... If it's on this coloured paper here ... we know it's come from Y team ... and people sometimes write on them ... sometimes people are a little bit naughty and that service user put their name on it, which you're not supposed to 'cause it's anonymous.

ID 201, clinical lead/manager, site B

The quality improvement lead in site B described having the opportunity to use 'trigger films' (patient-recorded storytelling) at the board meetings to illustrate key moments of interaction between the system and users in which quality could be improved by working together to implement agreed action points:

So what we do is we show these stories and we say 'what do you think of the issues that come out of that story? What would you have done differently? How would you make sure that you led the team and the privacy and dignity and respect and all of these kind of things we're providing to the service user?' And then we have kind of conversations with them about some of these issues. And it's a really powerful way of testing out the sort of softer issues that really mask some of the ... So if I asked you, you know,

do you have compassion? You'd probably really struggle to answer that question. But if I showed you one of the digital stories, I could tell right away if you were a compassionate person, because you'd probably be moved by what you saw. So it's a really simple way of, of testing that out.

ID 201, clinical lead/manager, site B

It was also acknowledged how other hospital trusts make use of other verbal feedback, such as discussions to accompany the completion of user-led patient-reported outcome measures in care planning and quality circles, to understand service user and carer experiences.

In both primary care sites (sites C1 and C2), the practice manager collated the feedback using a spreadsheet as there was very little feedback to analyse:

We had friends and family but we get very, very little feedback on that. We're lucky if we get one filled in a month.

ID 104, GP, site C1

So when they've got an appointment, we [receptionists] will then say 'Would you answer these friends and family questions?' ... So what I was going to say, when we have the little pieces of paper [FFT postcards], and the little box when it first came ... so the kids would draw on the pieces of paper, there's nothing ever when I look in that box downstairs, they never fill them in.

ID 333, practice manager, site C2

The need for more meaningful and positive feedback

Patients' and carers' perspectives

The format of feedback was a point of debate. Many patients felt that giving feedback was often too restrictive because of the structured approach and lack of space for open-ended comments. Many people felt that, alongside any criticism, they would like to offer positive feedback but were uncertain about the best way in which to do this. They voiced a need to allow more personalised and narrative feedback as this would be able to reflect their individual experience in a more nuanced manner. The following patient also saw this as offering more choice:

The forms that I saw at my GP clinic, they were just a feedback form and there was no real area to write anything for the feedback. It was just tick boxes. So that narrows your opinions down to ...

ID 11, site A

The benefits to staff that positive feedback could bring in terms of improving staff morale were explained as follows:

Well I think from the hospital's point of view they need to know when they're doing something well and that, you know, it's for morale in the hospital ... If they know it's something works well then perhaps they can be a bit of a champion to help other hospitals in the group ... we know it works well because we have this feedback from this proportion of patients and they tell us why it is. So, you can, hopefully, eradicate the bad and roll out the good.

ID 128, site A

A further point made about positive feedback was that it could serve as additional information for patients:

Absolutely, yeah. It's reassuring to people as well, because you don't always know what to expect. If somebody can say it's a good clinic, you know, you must go, it's not negative it's positive.

ID 107, site A

The underlying feeling of the majority of patients and carers was that dominant survey-based mechanisms for collecting feedback were 'straightjacketing responses', leading to predominantly negative feedback. In contrast, patients and carers indicated that they adopted more informal and innovative ways of expressing their positive feedback. The following example is typical:

I did [give positive feedback] when I was in the heart care unit because I could not fault them and the nurses and everything and when I went home and came back I bought them a big box of chocolates to say thank you.

ID 124, site A

Staff perspectives

Staff members presented diverse views about the meaning of feedback. However, there was broad consensus between groups that meaningful feedback would be the type that would give the respondent an opportunity to explain their experience in narrative form rather than ticking boxes to complete a rating scale:

I think it's got to appeal to the person, to want to do it, to engage with it. So I would think that there's a number of clients, that if the question was such, how . . . what have you gained from this? Let us know something that we've done to help you. Or, what's good about our service? And also, what has helped you less? So if it's black and white like that, maybe they don't have to think about it too much. People don't want to look at loads of questions and things, and gobbledegook. They'll just see that as gobbledegook. And so if there's something very, very simple like, 'what's been good about our service?' As simple as that. 'What's been bad?' Then they can actually put down something personal, something that's really affected them uniquely. I really like my care co-ordinator.

ID 217, care co-ordinator, site B

So if questionnaires could be better designed, and they could be much more open and give people an opportunity to tell those stories without it being too long.

ID 303, GP, site C2

Of note, staff also thought that feedback was really meaningful only when it could be seen to have an impact on service improvement:

I suppose you want nice constructive feedback, so that would . . . fulfil both angles, wouldn't it, if patients came up with suggestions that were not critical but were improving the service. That would be the biggest win for everyone. Nobody likes negative feedback.

ID 101, nurse, site C1

I think so because I think it would make people feel . . . like understand that their experiences are being listened to, and I think people, whether they've had a negative experience or a positive experience, having that recognised by someone can be a really important thing. Even if there aren't changes that can be made, I think having someone listen to that would be really beneficial, especially for the people that we work with who can be quite marginalised in society and feel that their voices aren't being heard. I think being more aware of how they can engage with the process would only be a positive thing.

ID 218, care co-ordinator, site B

Consequently, positive feedback was considered to be less valuable because it did not highlight aspects of service requiring improvement, and staff commonly thought that patients would also be more likely to want to give feedback in the existing surveys if they had negative experiences to report.

However, many staff (across all sites) did talk about the value and importance of positive feedback for boosting staff morale and confidence:

If you really only have feedback where there's been a problem, that can be somewhat demoralising, and yet we're all working hard every day and it's important to get that positive feedback.

Rheumatology consultant, site A

I have it as an item on the agenda for supervision, so as a supervisor, I ask have they had any compliments, and I don't know whether to say or not, if I feel that somebody's gone above and beyond then I might consider sharing that wider in the trust, or putting more focus on it because I think that certainly when you have negative feedback given to you, maybe that we don't always highlight positive feedback.

Assistant team manager/care co-ordinator, site B

Such positive feedback was commonly experienced by staff, but was received by more informal mechanisms, such as in face-to-face discussions with clients:

I suppose you get feedback through doing things like, you know, during your regularly routine, care co-ordinated visits. You get feedback about all kinds of things, you know, into obviously indications about how clients are finding the service. Now, it's not . . . doesn't necessarily other than getting written up in my daily . . . in your nursing records there, so that might not necessarily go any further really than that. So unless you were to trawl through hundreds of entries you might not see if there is any feedback from clients in that respect.

ID 204, community psychiatric nurse, site B

Some staff reported receiving thank you cards and gifts (sites A and C1), as well as e-mails from patients who just wanted to thank them (all sites). Field notes from observations also reflected the value of positive feedback given during informal discussions:

A lady saw me standing next to the kiosk with my UoM [University of Manchester] lanyard on and started to tell me verbally about her positive experiences at the practice. I asked her would she consider using the kiosk, she said she preferred to tell me in person rather than tell a machine and wouldn't use it even with the offer of guidance from a volunteer.

Observation note, site C2

An interesting interaction took place. I saw, a young girl (maybe teenage) came with a box of chocolate and gave it to the receptionist. Couldn't really follow the conversation between the girl and the receptionist. Not sure if that was a form of saying thank you!

Observation note, site C1

Staff perceived such feedback to be meaningful and valuable and thought that it was important that such feedback should be captured and recognised formally when summarising and reflecting on feedback about their service.

In summary, staff and patients across all settings thought that there was a need to generate more meaningful data, and narrative comments were viewed as being more insightful. It was recognised that analysis of this material might be complex but the results would provide better insights into what might be improved. Staff working in larger organisations were often sceptical of current practices used to collect feedback, which were perceived to serve the purpose of meeting organisational targets, rather than being useful for informing service delivery. Many thought that they would be more engaged in encouraging collection and using data to inform practice if the data were more specific to their clinical area.

Methods and tools need to suit the context and informal feedback should be included

Patient and carer perspectives

The final theme to emerge from the data analysis for this WS focused on the flexibility and adaptability of feedback methods and tools. The use of digital devices was generally welcomed by patients, carers and staff, but with some key caveats. One related to the ability to use digital devices and the effect on representativeness and equity:

You'll have, I don't know, a third who might struggle either with the technology or with the literacy. So, I think it would be really hard to avoid having some sort of support available because the difficulty will be some of your most disadvantaged patients will be some of the people who might struggle the most with access to some of the smart technology or with the literacy. So, you could get a very skewed sample in terms of the sort of feedback that you got.

ID 126, site B

Another issue pertained to having choice, namely some people would continue to prefer non-digital approaches:

Yeah, and you can use it in the future, I'm sorry to sound like an old Luddite, but that's the way it is. No, I would not be interested.

ID 118, site A

Suggestions were made that one-to-one interviews would yield more useful data, and group interviews were also considered to be a good option to elicit more detailed, experiential data. Most of the participants supported the survey format with a mix of closed- and open-ended questions. Although they wanted to supply in-depth feedback, participants were concerned that the process itself should be simple and easy and not time-consuming.

These findings indicate that the 'one size fits all' approach would not work, which was summarised well by this person:

So, if you're going to go down the route of digitalising the majority of the feedback, go down that route. It'll work for the majority of people. I think for the others you need to think about more traditional routes and to be able to do it in ways that are tailored to the needs of that population. So, inevitably I think there's something about increasingly smart segmentation in the population for the more challenged parts of the community.

ID 126, site A

In relation to people with mental health problems, the issue was raised that eliciting feedback required considerable sensitivity, given the complexity that most of them lived with. People with mental health problems more frequently said that they would be unlikely to use digital methods to give feedback, especially when unwell, as they often felt unable to write:

I don't like writing ... I hate writing. I can't put my mind down on paper.

ID 234, site B

There was also a feeling of mistrust of digital methods described by some service users:

I don't like putting things in writing ... that goes across the board, putting things in writing ... I'm not prepared to do anything like that online because I'm always suspicious of anything online ... No, I wouldn't want to be filmed either ... It's all part of my way of thinking, I suppose.

ID 221, site B

Others also openly described being less computer savvy and admitted feeling anxious and uncertain about using digital methods to give feedback:

She could do it one day and maybe not another. It's all to do with her mental health ... Well, if you're computer minded ...

ID 231, site B

I only call my brother or [mental health support worker's name], or taxis, and that's it ... I can't text, I don't know how to text. But I don't want to text because I'm scared of the messages I'll get back, so I don't want to text.

ID 226, site B

Rather, they talked about the value of having individual discussions about their experiences as a better way of receiving and recording feedback:

I have a CPN [community psychiatric nurse] nurse and she comes to see me every 2-3 weeks about my tablets and talks to me ... then you can discuss things, can't you?

ID 208, site B

I had nowhere to go to talk to anybody even though, like, I had a CPN [community psychiatric nurse]. [CPN's name] has been my CPN for a long time now, but I had nowhere else to go and that's when I started coming more frequent to the drop-in. I met new friends, some of us we decided, you know, we have a little talk on our medication so it's not just as though you're the only one that's on medication, you're actually speaking to others and explaining how you felt and things like that.

ID 318, site B

In conclusion, a range of factors were mentioned that emphasised the importance of context, such as individuals' social and cultural background, their personal preferences, the type of condition(s) they were living with, the stage of their long-term condition, the type of service accessed and the timing of asking for feedback. Consequently, flexible methods and tools were seen as a necessity if a high level of meaningful participation was to be achieved.

Context was referred to in terms of condition-specific experiences for patients, ways of working for staff and the organisational environment.

Staff perspectives

General practitioners and practice staff talked about the pressures on their service and having limited capacity to cope with a higher volume of feedback in the current system of collecting feedback:

Well ... at the moment it would be useless because we wouldn't have any system for dealing with it so, you know, if we ... if we got 500 free texts back, it would be the lowest part of the ... you know, the last thing we would do is look at that. There's many more priorities. So at the moment, it would be useless. So all we want to hear at the moment is, yes, you're brilliant, which is a ... you know, how many per cent of people thought you were good? That's the level it is now.

ID 303, GP, site C2

Previous processes using pen and paper had not worked; patients did not complete them and no one had time to enter or process these few bits of data:

... well, in previous years, we've done big surveys, we've done mailshots, we've handed them out over the desk and the response is actually quite poor, especially if it's a mailshot, we don't get a lot back.

ID 139, practice manager, site C1

Furthermore, GPs described their limited resources in terms of staff and IT expertise as a key reason why simple digital systems with automated analysis could be valuable to manage data. However, one of the GP sites was more keen than the other to engage with digital methods. During the interviews, this site described the benefits of offering a suite of digital and non-digital tools to increase and widen participation in feedback:

R1: ... the ones who want to write reams have got the paper, they've got the paper option if they did, they won't but they have the option if they want to write reams.

R6: the ones who could use the keypad will be the ones who are in a rush.

R4: ... if it's the younger population, that'll be the cohort that will be presumably going to work or ... it'll be interesting to see the take-up!

ID 235, GP, site C2

... now we've got ... I think the response to the text is ... I don't know. I'm going to have to look at the figures, but [not] massively ... we get much more response via the text now than we did on paper in the waiting room.

ID 333, practice manager, site C2

This site expressed a preference for receiving feedback relevant to particular issues that they were facing as a practice, such as receiving feedback experiences from the flu clinic, but with the caveat that the analysis and reporting was automated alongside the digital collection:

But then there is potentially other things that you'd want to investigate each year which ... let's say the flu clinic. And if you have a system whereby the university analysed your free-text data, so it collated all the people who said 'prefer 10 o'clock on Saturday morning', so that was a bit that it identified out, we could then get more people in to the 10 o'clock bit and less people into the other bits ... I'm paraphrasing. But the actual university system would collate the free text. It's not us reading each individual free-text thing.

ID 303, GP, site C2

Staff in a community health team and service users/carers talked about the value of face-to-face discussion about experiences of mental health services. This was viewed to be something that was happening anyway during visits and conversations but was not always documented in detail or systematically. Capturing this discussion formally as feedback might provide a more sensitive and inclusive process, be more consistent with the recovery model of care that they worked from and not add to current heavy workloads:

... it should really be something we'd be doing, and then ... because it's digging really for what is the therapeutic purpose that we're actually engaging with, because one of the early thoughts they'd had was about hanging it on the CPN [community psychiatric nurse], so that when we go and do the visit, we've got the CPN with us and we have the discussion and then we ask for the feedback, and I'm just thinking the complimentary stuff could be got through that value question. Equally, the hostility would come also with that same question because if somebody's not happy with what we've said to them, has this been of value? And you've got it, because there's no inhibition about that, is there? So I'm just thinking if we build that into each and every visit, that's systematic.

ID 349, care co-ordinator, site B

The community mental health team assistant team lead concurred with this and thought that this new approach of capturing the discussion formally as feedback would enable service users to participate who would not give feedback using the current formal ways of collecting feedback:

I think as well with that it's difficult when it's on paper because some of the service users we cover, not generally, but can't read or can't write, so we do have some people that either their first language isn't English or they just can't read or write. So if there was some other way of providing feedback ... I know I've supported a number of service users who can't read so generally wouldn't even open the letter, so you wouldn't get a response from them.

ID 208, care co-ordinator in leadership role, site B

It was felt that having discussions about experiences of services formalised as feedback and analysed in an aggregated way would allow common issues to be identified and reflected on at team meetings and enable these to inform best practice. It might also help to make feedback more meaningful and useful to individual staff, the team locally and the wider trust.

Workstream 2: text mining, analysis and presentation of data

Manual annotation of data sets

The annotators manually labelled 684 and 1004 different segments from 408 and 727 free-text comments from site A and site B, respectively. *Appendix 6, Table 22*, provides the distribution of topics (at the segment level) for the initial set of 11 themes. In the site B data set, the highest number of segments refers to care quality (36.6%), whereas, in the site A data set, staff attitude and professionalism was the most common segment category (24.9%). It is worth noting that 12% of comments from site A (general hospital) were considered to be 'not feedback', compared with only 0.2% from site B (mental health trust).

As explained in *Chapter 3*, the initial themes were merged into four topics (staff attitude, care quality, waiting time and environment) and the 'other' category (including the original 'other' and 'not feedback' comments). *Table 7* presents the theme distribution in the final (gold standard) data set.

Given that multiple researchers carry out the coding, we estimated the inter-rater agreement on random subsets of the coded data using the average F_1 score (averaging the overlap between the raters) and the average Cohen's kappa value (difference between the observed and expected agreement by chance). In the site A data set, the average F_1 score was 73% and the average Cohen's kappa was 0.338. The average F_1 score in the site B data set was 81% and the average Cohen's kappa score was 0.662. Cohen's kappa values of ≤ 0.40 are typically considered to indicate low agreement, whereas values of > 0.40 to ≤ 0.60 and > 0.60 to ≤ 0.80 indicate moderate and substantial agreement respectively

TABLE 7 Theme distribution in the final data set

Theme	Site A data set		Site B data set	
	Examples	Percentage	Examples	Percentage
Staff attitude	224	32.75	218	21.71
Care quality	156	22.81	520	51.79
Waiting time	103	15.06	98	9.76
Environment	45	6.58	60	5.98
Other	156	22.81	108	10.76
Total	684	100.00	1004	100.00

(with values of > 0.80 to ≤ 1.00 indicating almost perfect agreement). The metrics indicate the low to moderate inter-rater agreement, in particular for rare themes, suggesting that the identification of theme and sentiment pairs is challenging, even for human users (Tables 8 and 9).

Automated identification of themes and sentiments by the text-mining methods

The 'gold standard' data set described in the previous section was used to train and validate the machine-learning methods described in Chapter 3. We used a standard cross-validation approach in which the data sets were split randomly into a number of subsets (called folds) and the training and testing processes were repeated a number of times, each time using a different fold(s) for training and testing to reduce the variability. The results were then averaged across these training/test rounds. Specifically, we used 5×5 -fold cross-validation, performing five training/test rounds and splitting each data set into five groups, each time using four subsets for training and one for evaluation. For evaluation, we used the standard precision (P; the percentage of predicted themes for segments that are correct), recall (R; the percentage of segments with a given theme that are correctly predicted) and F_1 score (a harmonic mean of P and R). In addition to evaluating performance on individual segments (the segment level), we also investigated how well the models were able to predict all themes assigned to a given comment (the comment level). For example, if a comment in the 'gold standard' data set is coded with themes (t_1 , t_2), a correct prediction for the comment-level evaluation would be recorded only if both t_1 and t_2 are predicted by a system. We averaged the scores both at the segment level and at the comment level (micro averages).

TABLE 8 Inter-rater agreement for the site A data set

Aspect-sentiment pairs	Label count	F_1 (%)	Cohen's kappa
Care quality (negative)	8	50.00	0.2973
Care quality (positive)	1	0.00	0.0000
Staff attitude (negative)	7	85.71	0.8060
Staff attitude (positive)	12	100.00	1.0000
Waiting time (negative)	12	100.00	1.0000
Environment (negative)	2	0.00	0.0000
Environment (positive)	3	66.67	0.6286

TABLE 9 Inter-rater agreement for the site B data set

Aspect-sentiment pairs	Label count	F_1 (%)	Cohen's kappa
Care quality (negative)	22	72.73	0.6492
Care quality (positive)	43	93.02	0.8757
Staff attitude (negative)	16	75.00	0.7021
Staff attitude (positive)	13	76.92	0.7361
Waiting time (negative)	13	76.92	0.7342
Environment (positive)	2	100.00	1.0000

Table 10 presents the results for the two text-mining models developed (SBM and CLM). The F_1 scores varied widely across the different themes (between 30% and 85%), with some themes (e.g. waiting time, staff attitude and care quality) predicted with a reasonable level of accuracy. The environment theme proved to be specifically challenging. There were no significant differences between the segment-level performance and the comment-level performance. As expected, the SBM performed generally better at the segment level. The CLM performed better at the comment-level, but only on the site A data set (general hospital); on the site B data set the two systems were comparable.

As discussed in Chapter 2, we integrated the results of the two models to optimise the precision, recall and F_1 score (see Appendix 6, Table 23). This led to significant improvements across most of the themes, comparable to the inter-rater agreement (although still $\approx 10\%$ lower). Specifically, the optimised precision values at the segment level were 71% and 85% for sites A and B, respectively, whereas the optimised F_1 scores were 61% and 71%, respectively, for the two data sets.

Following a similar approach, associated sentiment was evaluated at the segment level using accuracy (the percentage of segments that have been attached to the correct sentiment). The models predicted 92% and 88% of segments correctly in the site A and site B data sets, respectively.

To illustrate the potential for processing large-scale data sets, the integrated text-mining model was applied to the full data sets from site A (110,854 comments) and site B (1653 comments). Figures 6–8 suggest that both of the data sets contained a similar proportion of the positive and negative aspects, with the number of negative aspects being higher than the number of positive aspects. Similar patterns were also seen regarding the distribution of topics, as noted in the ‘gold standard’ data sets.

Report templates to present analysed data

Following development and analysis using the text-mining programs, the data (classified into the key themes) were transferred for processing to generate an automated report file for staff (see Appendix 6, Figure 13, for an example of the automated report).

TABLE 10 Evaluation results cross-validated on the ‘gold standard’ data sets for the SBM and CLM

Theme	Site A data set						Site B data set					
	SBM			CLM			SBM			CLM		
	P (%)	R (%)	F_1 (%)	P (%)	R (%)	F_1 (%)	P (%)	R (%)	F_1 (%)	P (%)	R (%)	F_1 (%)
Care quality	51.58	55.12	53.20	46.86	32.43	38.26	71.73	94.37	84.13	78.35	68.86	73.08
Staff attitude	64.04	83.91	72.61	81.25	60.54	69.36	61.87	47.03	51.77	64.65	36.43	46.59
Waiting time	61.63	87.61	72.22	78.78	56.40	65.70	75.35	86.66	81.35	92.30	79.75	85.52
Environment	61.84	21.12	31.02	51.18	25.75	34.20	78.38	40.65	50.33	31.48	32.07	31.33
Other	45.89	60.83	51.96	49.29	45.32	47.10	28.60	12.37	16.59	21.46	28.45	24.36
Micro average (segment) ^a	59.29	67.64	61.41	63.03	47.18	53.96	68.58	70.61	69.57	65.05	56.00	60.18
Micro average (comment) ^a	36.76	75.71	49.36	47.51	59.50	52.81	56.60	80.93	66.58	53.17	69.61	60.27

P, precision; R, recall.

^a Micro average (segment) and micro average (comment) refer to evaluation on individual segments and the whole comment respectively.

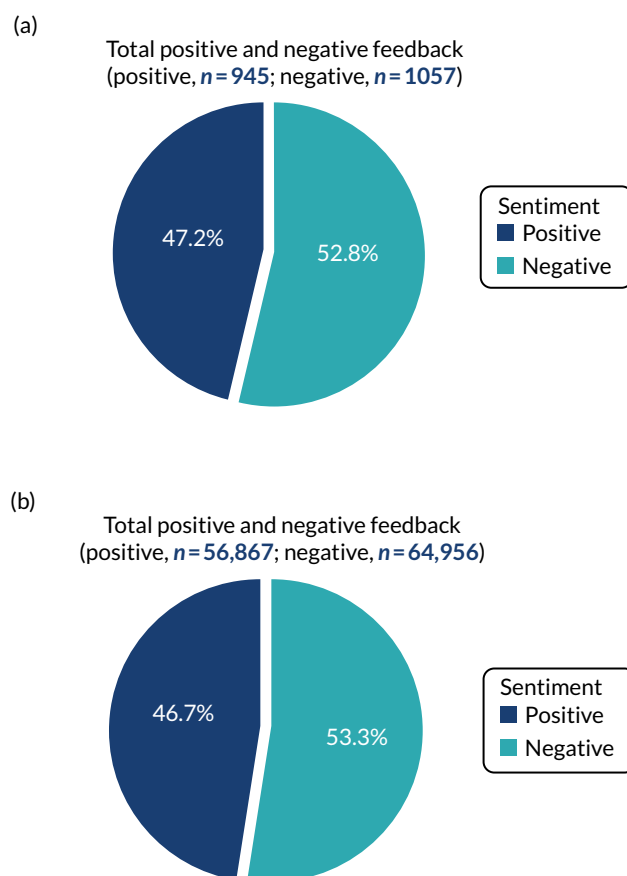


FIGURE 6 Distribution of positive/negative sentiments in the raw data sets of (a) site A and (b) site B.

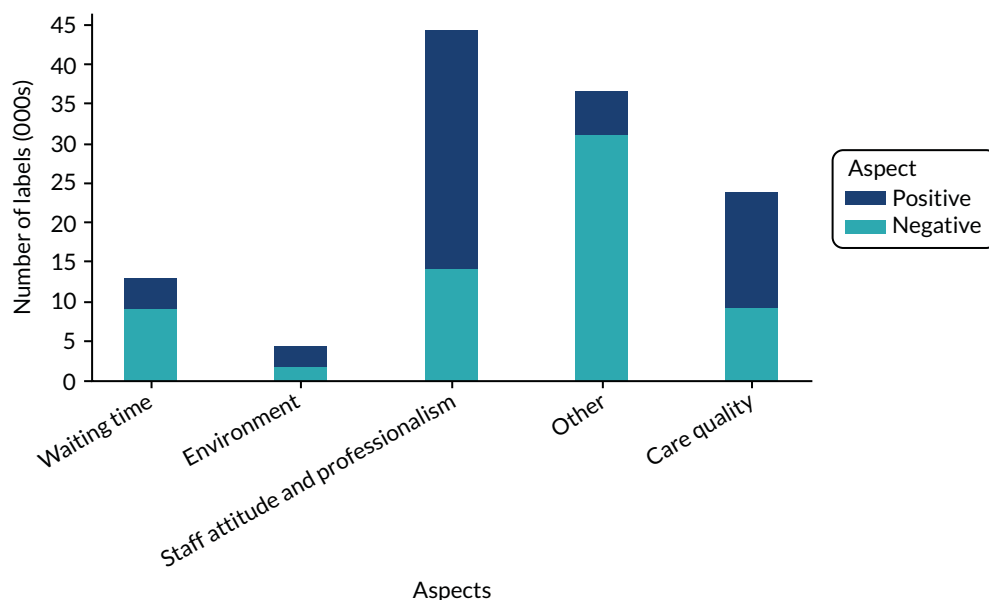


FIGURE 7 Per-class positive and negative aspects in the site A data set (110,854 comments). Sentiments by aspect categories.

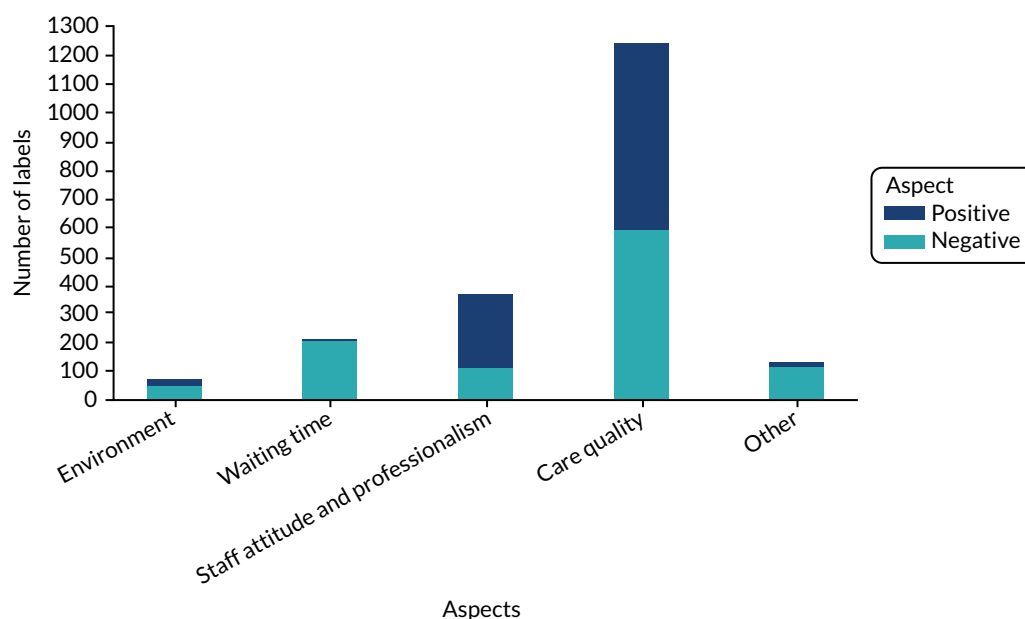


FIGURE 8 Per-class positive and negative aspects in the site B data set (1653 comments). Sentiments by aspect categories.

Workstream 3: co-design of a toolkit for enhancing the collection, analysis and usefulness of patient feedback

Interviews and focus groups in this phase confirmed many of the perspectives from WS1. This WS focused on discussions about practical solutions that could be tested out as part of the toolkit in WS4. Participants were also asked for their views on what would need to be in place to support implementation and adoption of the new tools. As illustrated in *Chapter 2*, a number of slides were developed and presented as a means of summarising some of the research conducted in WS1 and WS2 and to prompt discussion around possible tools to make clear decisions on what could be implemented and evaluated in WS4.

The discussions ultimately led to the development of a toolkit comprising the following items:

1. A survey utilising the FFT, with space for free-text comments, to be completed using a digital kiosk within study sites, online or using a pen and paper version (see *Appendix 3* for the DEPEND toolkit). At the beginning of the survey there were screening questions to ensure that only the patients or carers of that service completed the survey; at the end of the survey there were optional questions on demographics (age, sex and ethnicity). Throughout the implementation period the surveys were adapted for each individual site to improve the response obtained.
The online survey in all sites was almost identical to the kiosk survey. However, there were some minor formatting differences in the way that it was displayed on screen, for example emoticons were not available in the online survey.
Different pen and paper surveys were already available in each site prior to the study. These ranged from a two-page outpatient survey at site A to postcards including the basic FFT at site C2.
2. Guidance and information for staff, patients and carers to support use of the new tools.
An information sheet was distributed to staff, patients and carers at all sites to answer commonly asked questions about the new digital kiosk prior to it being installed. The following guidance was also available for patients and carers:
 - an information slip given to them by their clinician at the end of their appointment to prompt them to give feedback using the available methods at that site
 - a poster highlighting the reasons why feedback is important and the different methods available to give feedback

- a leaflet highlighting the reasons why feedback is important and the different methods available to give feedback
 - one-to-one support for using the digital kiosk from a researcher or volunteer (when this was available at the study site).
3. New text-mining program for analysing patient feedback data.
- The digital survey textual data captured using the kiosk and website versions of the survey were analysed using new text-mining algorithms to capture the volume of comments made over a range of topics and the sentiments in each. The open-source code for the text-mining analysis is provided online [<http://gnteam.cs.manchester.ac.uk/depend/> (accessed 10 October 2019)]. An instruction manual can be downloaded from the same website (see also *Appendix 6*).
- Although this is a component of the toolkit created during the project, it was not used to analyse monthly reports because of the small volumes of data collected during the 1-month periods. It was decided to run the text-mining analysis only on the large data set collected over a longer period of time.
4. New templates for reporting feedback from multiple sources.
- Templates were created to provide a summary of quantitative data from the survey, followed by examples of comments collected to illustrate key themes, with a summary of key themes from all of the comments collected over a 1-month period.
5. A new process for eliciting and recording verbal feedback in community mental health services.
- During the focus group discussions and interviews with staff in the community mental health team, participants discussed a process for eliciting and recording verbal feedback in community mental health services that they envisaged working because it would easily fit with their usual work practices. This formed the basis for a guidance document to support implementation of this process.

See *Appendix 3* for examples and resources for each component of the DEPEND toolkit bespoke to each site.

The findings presented in this section summarise the discussions in focus groups and interviews in WS3 and how these relate to decisions made and activities carried out to develop the toolkit components. The views of staff, patients and carers are provided in three separate sections related to the toolkit components described above:

1. capturing feedback on patient experience – the need for digital and non-digital systems, information and support (toolkit components 1 and 2)
2. analysis and presentation of patient experience data – the value of text mining and perspectives on reporting and presentation of data (toolkit components 3 and 4)
3. eliciting and recording verbal feedback in community mental health services – the importance of enabling feedback through interaction and established ways of working (toolkit component 5).

Capturing feedback on patient experience: the need for digital and non-digital systems, information and support

Ways of capturing digital and non-digital feedback

Following the presentation summarising the findings from WS1 and WS2, and some illustrations of potential ways of addressing the issues raised, there was some discussion about the possible use of a digital kiosk to collect feedback in the waiting areas in each site. There were many overlaps between the discussions with staff and the discussions with patients and carers. The discussions led to the decision to install digital kiosks in the waiting areas (outpatient departments and general practices) in each site.

Patients mentioned a number of advantages of using the digital devices and expressed their views about the value of collecting feedback in this way, such as improving speed, efficiency and accuracy. As one respondent said:

Well, I think the idea of the digital ... yeah, pressing the box is quite a good one, because it's relatively quick, it's straightforward, it's easy to deal with.

ID 115, patient interview, site A

I think the fact that you've got ... I mean, like the idea, for example, the touchscreen ... I mean, I'm just thinking again of like the GP surgery where you automatically go in and you book in ... I think that's quite a good way of doing it, because it's quite visual, it's automatically recorded.

ID 115, patient, site A

In terms of being able to handle data more accurately or probably more speedily, if you've got, say, paper input then in some way if you're going to analyse it you've got to process that paper input. It's far easier to process if the information is already in digital form. And so it benefits everybody to have a faster, more economical and more accurate process. So that's how I see the benefits of digital.

ID 113, patient, site A

Most staff members also talked about the value of having easy-to-use digital tools in waiting areas, with these becoming increasingly familiar to patients. However, following on from issues discussed in WS1, there were also many suggestions regarding how to ensure the availability of alternative mechanisms for giving feedback and how to ensure flexibility:

... you can have your screen but you could actually have pen and paper next to it.

ID 141, GP, site C1

Service users and carers with experience of mental health problems again discussed the need for alternative methods of collecting feedback, and these views overlapped with the views of staff (discussed below). Some older respondents with experience of MSK conditions expressed a hesitance to use digital tools, but also that they might use a digital device if given support. There was also a concern that some people with painful or swollen hands or mobility restrictions may find the kiosk difficult to use:

The only thing I was worried about ... was that people with arthritic hands might not be able to type.

ID 107, patient, site A

Let's look at the other side of the coin. People in wheelchairs, are they going to find it as convenient as an able-bodied person that can stand up and operate a screen?

ID 225, patient, site A

I would have thought the idea is to move as many people as possible onto a digital system. But then if for any reason you're unable to comply ... if they've got problems with their hands, other than a verbal interview I can't see how they're going to give feedback anyway, if they're not going to be able to write. I suppose the idea is you wouldn't want to leave anybody out if they wanted to leave feedback.

ID 113, patient, site A

One respondent highlighted that there should be options for responding in other languages or in other ways as an alternative to typing:

See, the other thing that springs to mind, which you've probably already thought about, is the various languages that you get in the outpatient department ... The world we live in or the country we live in,

there's got to be different nationalities and languages. To me, if you want feedback from a whole cross-section of the patients that you're getting in that department, you're going to get Polish people, you're going to get ...

ID 225, patient, site A

Discussions highlighted the need to ensure simplicity and maximise the usefulness of the digital process, which led to the decision to use the FFT question. In addition, all sites wanted to enable the collection of free-text comments, with an explicit prompt for positive feedback, as well as things that could be improved, and examples to illustrate the types of comments that could be usefully provided.

Staff discussed the benefits and limitations of using a generic question such as the FFT. Although some thought that there was a need for questions to be more specific and directly relevant to their area of service provision, there was also a view among staff that it would be most appropriate to start with the FFT because all services were obliged to collect these data anyway and it seemed to make most sense to explore if the usefulness of these data could be improved by collecting them digitally, as well as by collecting free-text comments, in waiting areas (outpatient departments and general practice waiting rooms), as this had not been routinely carried out across the sites:

Perhaps it'll be useful to have like a general, if possible, because people come to us all the time, don't they ... what's been your overall experience type questions and then, specifically, how was it today ...

ID 134, practice manager, site C1

Actually trying to get at rather than making it so generic, actually asking them the specific questions, I think is a much better idea.

ID 141, GP, site C1

There was some discussion about the use of emoticons as a means of making the interface simple and friendly to use. PPI contributors also favoured the use of emoticons in the interface because they used them regularly anyway. However, staff working in mental health services (site B) talked about the sensitivities of using emoticons in the context of mental health:

I don't know whether necessarily in mental health they're always the best tool for measuring satisfaction, you know, for somebody who is suffering with an episode of depression, there's going to be nothing that makes them smile from ear to ear, so to see it as an emoticon to measure the most satisfied with a big smiley face, it's not something that's right at that time.

ID 349, care co-ordinator, FG, site B

These diverse views led to a decision to use emoticons on the interface at three sites (sites A, C1 and C2), but not at the mental health trust (site B).

There was some discussion on this theme about the issue of anonymity and related ethics concerns over patients and staff being identified in the feedback given. However, there were mixed views. Some recognised that people might feel concerned about the consequences of giving negative feedback, worrying that it might have a detrimental impact on their care, whereas others highlighted the importance of enabling feedback to be given more naturally in the context of health-care interactions and enabling staff to reflect more specifically on their own practice. Similarly, there was a recognition that patients and carers might want a response to any issues raised:

And ultimately what the feedback's for, if it's just a general ... if it's specifically about a professional or people he had seen then you've obviously got to put a name to it, haven't you?

GP, FG, site C1

... the 360 appraisals are a complete waste of time. They don't provide that level ... You're not going to say something nasty about a colleague. And also the person who you nominate, you're not going to nominate a person who's going to give you negative feedback. So you're not going to get it from your colleagues.

ID 6, consultant, site A

I suppose the other thing is, looking through all these comments, I know it's good to be anonymous but at the end of the day maybe I'm the one that everyone's saying was slightly rude and difficult or not. And we're then trying to second guess who that rude doctor was or who that really nice ... do you know what I mean? And I've no idea ... But it's how useful that is to know as a team, there are one or two people that are often behaving rudely or not, and how do we then deal with that?

Consultant, FG, site A

Information and support

The views of staff expressed during follow-up interviews expanded on the initial views expressed in WS1 in highlighting the lack of awareness of feedback among patients and the need to inform patients about the value of describing what is good about a service when giving positive feedback:

Many a times the way the services are set up in any sector, it's easy to complain ... this is a pure guess, most if not a lot of consultations are actually fine or good. And you never get feedback. So you don't know what good things you are already doing subconsciously. And there is no easy mechanism of getting that ...

ID 2, consultant, site A

Patients emphasised the value of both negative and positive feedback and that both kinds of feedback should be encouraged:

The negative feedback is clearly you want to know where things could be improved. The positive feedback is where you don't want to make unnecessary changes when things are working well, but you need to identify those areas where the patient thinks things are working.

ID 107, patient, site A

One of the patients talked about the potential for using information and signposting to motivate people to give feedback, and this idea was ultimately adopted to design a large poster, in conjunction with the study PPI group, to support use of the kiosk:

I don't know if you'll remember this, but in the first and second world war the Americans had this sign of Uncle Sam, he needs you, and it was like an old guy with a top hat pointing his finger at people, big signs and saying that we need you ... They needed recruits for the army and things like that. But I can't think of what kind of picture you could have really trying to get people's attention as to the importance of feedback. But it's something for you to play around with isn't it? I don't know.

ID 225, site A

This gave rise to discussions about simple ways of providing this information and led to the decision to provide examples of positive comments on posters advertising the kiosk, as well including a specific prompt asking 'What did we do well?'.

Some suggested developing a video clip channel to popularise feedback:

I always think in pictures so what might be an idea, and it's only an idea, the kiosk ... has to have a little 20-second video that shows all this information in cartoon form or something like that so it's very, very short ... just so it's more entertaining rather than like it's boring. People are more inclined to look at it and think 'oh, OK, I can do something like that'.

ID 117, patient, site A

Some people talked about the need to provide hands-on support, especially for older people or others who might be less confident in using digital devices:

I think it just depends on people's ... well, partly the age, isn't it, and their abilities. I mean, obviously some older people will struggle. I mean, I do intend to get computer literate again, but I've not been able to use my computer, and I mean, my skills are quite basic ... If there's like some support available, so for example when you think of the supermarket when you use the ... self-serve, and there's usually someone there, and it comes up on screen and so if there's someone there if you're stuck or something flags up.

ID 115, patient, site A

These issues gave rise to discussions about possible ways of providing support for patients and carers to use the kiosk, including providing support from volunteers in the trusts and members of the PPG groups in primary care.

Analysis and presentation of patient experience data: the value of text mining and perspectives on reporting and presentation of data

Most of the staff liked the principle of having a system using text analytics and automated reporting, which would provide them with a simple tool for analysis of the collected data. In terms of the results of the feedback, some would like to receive all of the results before modification in the form of a summary report. For example, an assistant manager said:

No, I feel we want to see all of it, we'd like to see the comments, we'd like to see percentages, yeah, if we're analysing it I think that's what we want.

ID 141, GP, site C1

In terms of the presentation of the results, most staff members preferred a more traditional method of presentation, such as a bar graph or a pie chart. Most thought that a bar graph would be useful:

I think the bar graph is what we're used to anyway, every time we get practice feedback it tends to be the bar charts.

ID 140, GP, R2

Staff members talked about the importance of ensuring that displayed data are easy to understand and familiar to staff. Many respondents did not like the word 'cloud' as a way of summarising language used in comments. As one participant said, 'that, I've tried to read it, I still don't know what that is ... because that is just confusing me' (ID 1, site A).

In terms of reporting feedback, staff members thought that it was always nice to start with positive feedback, but that they need to receive both positive and negative feedback:

Because you'd come out of review thinking you're doing rubbish if you just view the negatives ... You do need both but I'd start with the good news.

ID 140, GP, R2

An issue that arose in WS1 for all of the sites was the importance of asking very specific questions that relate to specific practices. Almost all staff members agreed with this.

Some staff members talked about the importance of enabling analysis that would be useful for managers in assessing issues across the wider organisation:

I know what the problems are. It's getting management to do something about it and showing them that this change is the only way I think. Or benchmarking them against someone else, saying we're doing worse than others. So it's some way to get some change to benchmark against other units, other departments. I think it would influence management.

ID 132, consultant, FG, site A

Staff members talked about the importance of reporting results to patients and the specific requirements for this:

... yeah, that would be more useful to us to analyse lots of things but not for patients, it's great whereas those two [pie and graph charts], if we're going to put something out on the board downstairs they're much clearer.

ID 141, GP, site c1

Staff felt strongly that patients should receive feedback on their feedback:

Actually I'd want to feed back the feedback to the patients and that is how I would do it.

ID 141, GP, site C1

Many staff members thought that the ability to look at changes in patient experience over time using graphs on the reporting templates would be very useful because they could assess such changes alongside other data and in relation to other factors that could potentially have had an influence on patient experience:

I think the time element is really important to this feedback because if you're seeing an increase in ... I know we were discussing before about waiting times, but if you're seeing an increase in the number of negative comments about waiting times it could be related to sickness and absence levels or something. So I think the time element is quite important.

ID 133, manager, site A

Eliciting and recording verbal feedback in community mental health services: the importance of enabling feedback through interaction and established ways of working

Staff in site B expressed a preference for a complementary way of working, in line with the recovery model of care and care planning approach:

I guess that is kind of getting feedback on what's working and what isn't within someone's care plan. So if there was a way that that could be done, because then that information could just be pulled off and taken somewhere and wouldn't mean any more work for us or for the service users, which could be nice. And I think sometimes something you might have set up with someone who you're working with on your caseload and done for them could then be something that other people could see and think 'oh, that's worked really well, that sounds like a really good option'.

ID 218, care co-ordinator, interview, site B

Staff recognised that sometimes feedback is given naturally in discussions with service users, and sometimes people ask specifically about this at the end of a visit, but this varies according to circumstances.

It was hoped that capturing discussions formally as feedback would provide a more sensitive and inclusive process and might also help to make feedback more meaningful and useful to individual staff, the team locally and the wider trust:

I think interviewing individually will possibly capture more of those people that we work with who, not wanting to label but the middle-ground kind of people who are happy but either are socially anxious or won't come to a group or struggle with motivation because of moods so won't fill out questionnaires or anything like that, so may engage more with that kind of one to one, I'm thinking? Because it's those people that I think generally, and it's an assumption, we would miss out and we miss their views because they're the people that can't or don't, for whatever reason, fill out forms or attend groups, so ... I think that's quite a high majority as well.

ID 208, care co-ordinator in leadership role, site B

Staff in site B felt that they needed to ensure that service users knew that they were not under any pressure to give feedback using the new method and that they could also give feedback using the mandated anonymous surveys in the trust.

Similarly to other sites, staff at site B (focus group) talked about the importance of reporting the results to the team and liked the concept of having access to feedback reports each month, but that these would need to be simple and structured with regard to their content:

ID 341, care co-ordinator in leadership role, site B: I think if they [the reports] specify what they think we're doing well, what we're doing badly or could improve on, rather than just general, oh yes, they're really great, just specify what parts of our service are good or bad and identifying needs.

ID 342, care co-ordinator in leadership role, site B: And just maybe some ideas for improvement or things we could ... more of a summary and some bullet points, nobody's going to have the time to read through transcripts.

Workstream 4: quantitative analysis of the volume of feedback pre and post introduction of the toolkit

Table 11 provides a comparison of the feedback participation rates for the two feedback periods [January 2017–31 May 2017 (pre DEPEND study) and June 2017–February 2018 (DEPEND study)] by method and site.

Overall, the quantitative evaluation indicates low levels of participation across sites over the two time periods, but a steady increase in the volume of feedback in both primary care sites, which had previously collected relatively small amounts of feedback data in comparison to the large acute trusts.

The kiosks enabled collection of a greater volume of FFT data and richer narrative comments (see the example summary report for site A in *Appendix 7*). However, the results were mixed. The figures for the large trusts (site A and site B) are based on one department in each trust. Furthermore, site B was already collecting digital kiosk data across a number of locations (site B was one of 30 locations) before the DEPEND tools were implemented in the sites in the summer months of 2017.

To inspect the pattern of kiosk responses, we plotted the number of kiosk responses over the study period (*Figure 9*). Overall, this figure highlights the downwards trend in kiosk use throughout the study period, which supports our qualitative data, in which patients and carers expressed reluctance to give feedback. Of interest, the total number of responses peaked during the 2-week introductory phase for each site when a volunteer, staff member or researcher actively promoted use of the kiosk. For example:

I approached a young male to use the kiosk and he agreed to use it. He explained he was autistic before he did. He completed it fine and quickly, but did not type very much free text.

Outpatients, observational note, site B

We present more data extracts to illustrate the level of support needed to sustain kiosk use across all sites in *Workstream 4: qualitative evaluation*. It was unfortunate that we were unable to sustain this level of support throughout the testing period. In site A there were 294 kiosk users. Of these, 235 responded to a question asking about their age category and 27% of those responding were aged between 45 and 54 years. This reflects the general population using this site (MSK conditions can occur at any age but the peak incidence occurs within the fourth and fifth decades of life).

TABLE 11 Evaluation of feedback participation rates by method, site and month: WS4

Site	Method	Month (n)													
		January 2017 (pre DEPEND study)	February	March	April	May	June (start of DEPEND study)	July	August	September	October	November	December	January 2018	February (end of DEPEND study)
A	Kiosk						86		42	38	76	21	6	16	9
	Online								1	1	0	2	0	0	0
	SMS	52	43	42	23	42	Turned off during DEPEND study								
	Telephone	23	27	23	23	18	13	20	14	20	20	17			
	Paper survey	75	70	65	46	60	54	86							
B	Kiosk	45					15	10	13	11	13	34	1	1	3
	Postcard	21									13	3	0	1	0
	Discussion							0	0		6	0	0	0	
C1	Kiosk						41		22	42	16	10	50	0	0
	Online								3	0	0	2	0	0	1
	Paper survey	1	0	2	0	0	1		3	0	0	0	0	0	0
C2	Kiosk							109	96	53	70	54	28	28	25
	Online							0	0	1	0	0	0	2	2
	SMS	30	70	75	93	56	68	56	65	66	100	35	20	25	32
	Postcard						0	0	3	0	0	0	0	0	0

Blank spaces in the table indicate where data were unavailable/missing.

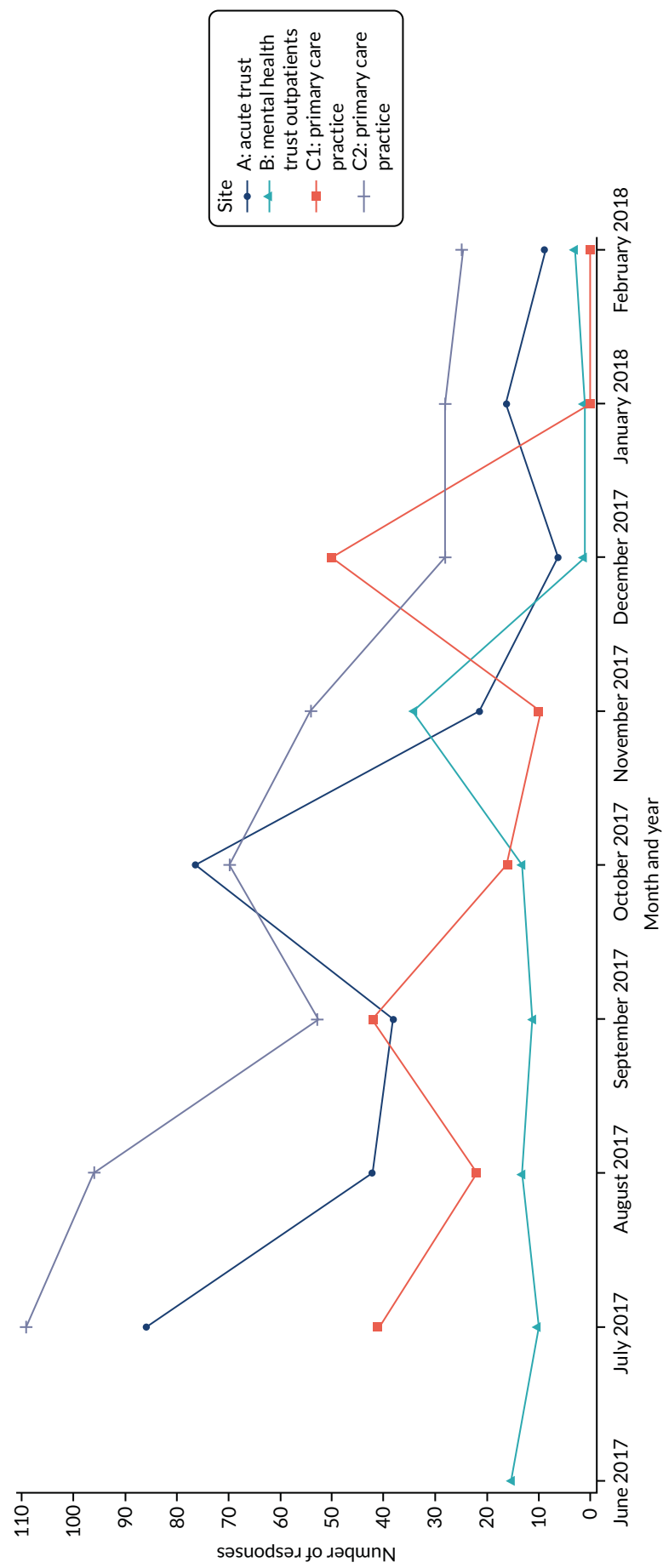


FIGURE 9 Change in the number of kiosk responses over the study period by collection site.

RESULTS

In site B, there were 111 kiosk users during the study period. Less than half of the users ($n = 48$) responded to the question asking about their age category and these respondents were more evenly spread across the age groups ($n = 5$, 16–24 years; $n = 7$, 25–34 years; $n = 10$, 35–44 years; $n = 10$, 45–54 years; $n = 11$, 55–64 years; $n = 1$, 65–74 years; $n = 3$, 75–84 years) (see *Report Supplementary Material 3, Figure 2b*). The majority of kiosk users ($n = 48$) were aged 55–64 years, closely followed by those aged 35–54 years and those aged 16–34 years, which is consistent with the wide spectrum of mental health service users across age groups in Greater Manchester.

At all sites most patients were either likely or very likely to recommend the service to their friends and family (see *Report Supplementary Material 3, Figure 1*).

The majority of kiosk users from the two primary care sites ($n = 644$) were also evenly divided between the different age categories, indicating that digital methods of collecting feedback data are acceptable for different age groups.

Kiosk users across all sites were predominantly female (62% of those responding to the question asking about gender) (see *Report Supplementary Material 3, Figure 2c*). Ethnicity data were collected in site B outpatients only and showed that almost all kiosk users who responded to this question were white (white, $n = 40$; black or black British, $n = 1$; Asian, $n = 1$; mixed, $n = 4$; other, $n = 0$; prefer not to say, $n = 2$; missing, $n = 25$).

The line graphs in *Figure 10* clearly represent the change in volume of feedback obtained by response method. This can help us to interpret the core issues that we experienced in attempting to enhance current feedback collection methods throughout the testing period.

Site A suffered a downwards trend for all feedback response methods and so participation was worse after the tools were introduced; we hypothesise explanations for this in *Workstream 4: qualitative evaluation*, but it is worth noting that this site suffered from many technical and practical issues from the onset. There were no comparative data for collection of feedback using other methods for site B, as FFT feedback responses were not routinely collected on site by this team. Site C1 had an extremely low participation rate for FFT feedback using the pen and paper survey before the new tools were tested, and so the participation rate in this site increased, even though it was still relatively low. Finally, some FFT feedback data were routinely collected at site C2 via SMS messaging, as well as a small number of FFT postcard responses; however, the rate of participation still continued to deteriorate over time, even though this site had the support of volunteers and an enthusiastic staff team, who championed digital data collection.

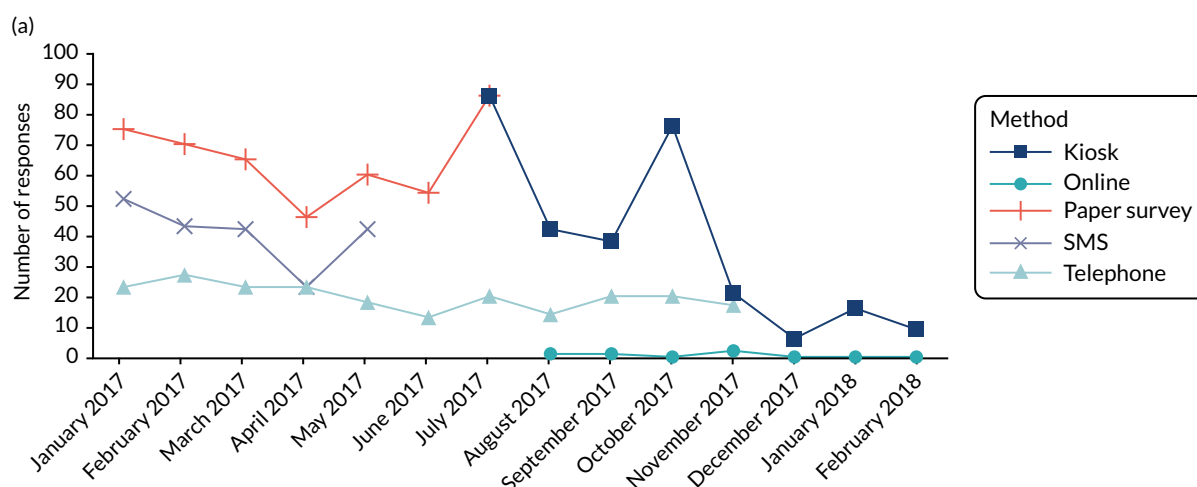


FIGURE 10 Change in the number of responses by collection site and response method: (a) site A (acute trust); (b) site B (mental health trust, outpatients); (c) site C1 (primary care practice); and (d) site C2 (primary care practice). (continued)

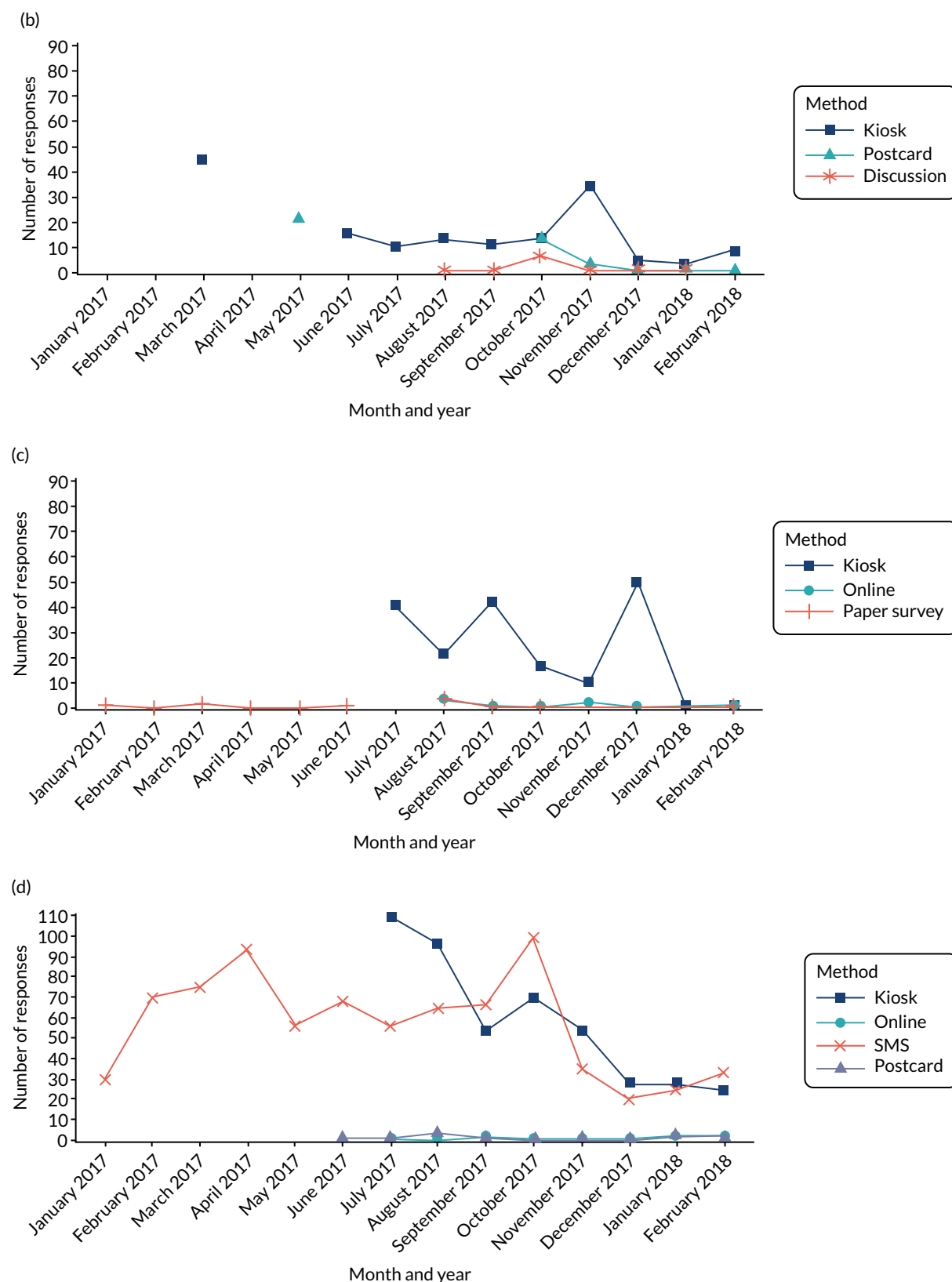


FIGURE 10 Change in the number of responses by collection site and response method: (a) site A (acute trust); (b) site B (mental health trust, outpatients); (c) site C1 (primary care practice); and (d) site C2 (primary care practice).

We found that the new FFT online survey (the same FFT survey as in the kiosk) was used sparingly across all three testing sites (sites A, C1 and C2) and this seemed to be because it was not advertised sufficiently [(the Quick Response (QR) code and short URL were available only on the poster and recruitment flyer at all sites and on the side of a prescription sheet at one site)].

The new approach at site B of collecting feedback through discussion with a care co-ordinator had the lowest participation rate. The challenges and barriers to collecting and recording feedback using this new approach were observed from the outset in the team and at the organisational level. For instance, this trust underwent a merger and acquisition in April 2017 to create a bigger trust, which impacted on participation rates and the capacity/ability to be involved in the DEPEND study. The lessons learned are further elaborated on in *Workstream 4: qualitative evaluation*.

Feedback collected via SMS message was enabled in one site only (site C2), where, interestingly, the volume of data remained consistent (see *Figure 10d*).

Participation rates for non-digital feedback methods were consistently low throughout the study, apart from in site A, where the pen and paper survey had comparable feedback rates to the kiosk (see *Figure 10a*). This might be because this was the most popular feedback method at this site for older people (age 45–54 years) attending the outpatient department pre DEPEND study.

Workstream 4: qualitative evaluation

The findings in this section are also reported and extended in Ong *et al.*⁶⁰ Technological innovations in health care have moved apace and it has become increasingly apparent that the organisational context in which new interventions are situated needs to be better understood. A wide range of conceptual models have been developed to analyse the sociocultural processes required to introduce, action and embed complex interventions. In this study we make use of NPT because it is a medium-range sociological theory that explains the work that people do to implement a new practice. It consists of four core constructs: coherence, cognitive participation, collective action and reflexive monitoring.⁵¹ These constructs are explained in turn in the following sections and applied to our study.

Coherence: perceived value of digital tools for collection and analysis

Coherence refers to the sense-making carried out individually and collectively when new technologies and practices are introduced. When digital methods are introduced in different settings, participants (staff and patients) need to be able to see the new approach as distinct from current practice. Moreover, the process of sense-making needs an understanding of the purpose of an intervention and what it requires participants to do. Finally, the value of digital tools for collection and analysis has to be recognised by individuals and teams.

The initial qualitative research in WS1 and the co-design approach used in WS3 meant that an understanding of 'sense-making' among patients, carers and staff underpinned the development of tools for testing in WS4. This helped to maximise a sense of coherence from the outset of the implementation and evaluation phase in WS4. Staff members across all sites were generally enthusiastic about the introduction of new tools for collecting feedback digitally on site, including the kiosk. They felt that the digital collection of feedback might improve the volume and efficiency of routine patient experience data collection, because the previous reliance on pen and paper surveys had been ineffective. However, as demonstrated in the findings from earlier WSs, there were variations between sites in terms of the degree of optimism about how well this might work and the perceived potential limitations of digital data capture. There were also variations among staff in relation to their roles and responsibilities and how these were associated with the perceived need for digital tools at the level of collection, analysis and use of data in specific contexts.

Site A: acute trust

In site A, although most of the respondents, including staff and patients, were in favour of using the kiosk over other methods, they advocated for multiple methods of collecting feedback. One patient said:

Well, I suppose I could. But again, I know people, elderly people, you know, my brother-in-law, my sister . . . has no idea how to use a computer, and to get online. And there are lots of people like that. So online would be less practical for them as well. And musculoskeletal problems, things like arthritis and so on, rheumatologically problems, they often affect elderly people, more than younger people often, so it may be a problem for them as well. There should be a range of options.

ID 145, interview, site A

Another patient stated:

. . . a lot of people don't, and a lot of people don't have computers. A lot of people don't use e-mail, so they don't know how to approach starting with it, to answer that, because not everybody likes that kind of thing, is there, you've got to think of the other person that prefers maybe just getting a bit of paper and saying 'oh, I was happy today', fold it and throw it in a box.

ID 223, interview, site A

A number of respondents viewed the best option for giving feedback as patients and staff having a face-to-face discussion, 'because sometimes people can talk about it more than write' (site A, ID223, interview, patient).

Other staff stated that:

I think the best way is face-to-face, talk to people, and ask them. And in my view it would be, 'what did you think of the service?'

ID 140, interview, site A

I mean, we obviously do the, where we, sort of, do patient surveys, that sort of thing, but that's usually a mixture of paper and online. But actually, a one to one then no, we haven't, but I think that's probably . . . quite a good way of getting feedback because it's probably the most accurate feedback you'd get really.

FG, site A

Site B: mental health trust

In site B, it was clear that the new tools made sense at the level of senior management. A senior manager with responsibility for patient experience talked about the value of the study and the potential value of new digital innovations that could improve the capacity to make data more useful for service improvement:

. . . what always really, really frustrated me was that lack of ability that we had to then funnel that back down into services. So [team lead name] actually was usually one of the ones that would always come to me and say, 'come on, now, we've sent you all of this data, all of this information about our team, our service users, what is that telling us?' . . . At one point I was doing all of these bespoke reports . . . It became very apparent very quickly that I just could not do that . . . And I think what this research always I suppose promised, or had the potential, was the ability to do that bit that was missing. So it all comes into the system, and then the system then tells the services what the experience of service users is like in the context of that service.

ID 201, interview, site B

The same manager also stated that the new process for recording verbal feedback in the community mental health team made a lot of sense in the context of staff members' routine day-to-day practice:

I think the idea itself is really, really good. And it's always something that I really, really supported. And I think what makes it hugely valuable is it creates that sort of ability for the teams to hear about the experiences in terms of some of that really fundamental care-planning work that they do on a day-to-day basis with service users. And to get some sort of really valuable feedback from them around how that's going, how that's happening, what would work better, how they could do it differently. Some of the things that they really value, some of the things that drive them potty.

ID 201, interview, site B

Use of the kiosk in site B: outpatients

Some staff were positive about the kiosk and thought that it made sense as a simple way of collecting feedback routinely on site:

I think it gives people an opportunity to say how they feel about the service they're receiving here, and quite often I imagine in the day-to-day contacts with the care co-ordinators the time is so absorbed with their illness or what's been going on in their lives that they don't have that time to even ask that question, or it might not be appropriate to ask that question. Whereas this is a choice, isn't it? I can choose now to say how I feel about our service in a very simplistic way, I'm happy about it or it could be improved or just a general overview. And I think it's quick and easy and that's what we need it to be.

ID 243, outpatients, interview, site B

Because it's different, isn't it, when people are giving feedback within the run of their practice as normal rather than making a complaint or compliment about a specific situation. It's more generalised, isn't it?

ID 254, FG, site B

However, there was also recognition of the low rate of participation and some concerns that this could be a problematic way of giving feedback that would not make sense for many service users and carers:

Obviously, you know, when it first came in we all kind of had a go at it and had a look at ... you know, had a play with to see what it was about, and what have you, but personally I haven't seen anybody, you know, go up to it at all.

ID 242, outpatients, interview, site B

It was observed that, when people were asked to provide feedback using the kiosk, the majority of people preferred to use pen and paper postcards rather than the digital kiosk or politely declined. Similar to people at other sites, people at this site seemed more willing to use the kiosk when a volunteer was present.

In interviews, patients and carers spoke favourably about having both digital and non-digital feedback options available, welcoming the choice of different methods for providing feedback. Once they realised the rationale for and the importance of giving feedback, and that the kiosk was there for them to use, some patients and carers said that they would be willing to try giving digital feedback using the kiosk. A number of people stated that they would complete a simple tick box survey but would not write any text, and others talked about the difficulties of engaging with digital technology in general:

I don't like writing and stuff, tends to, like, be a bit difficult in terms of ... I prefer just, like, a tick box ... Just ask on the visits I think. It would be easy 'cause then ... when you get it in the post, you just tend to put them in the bin, don't you? ... You don't really ... so it's better just to do it there and then in the moment.

ID 262, outpatients, interview, site B

I've got a computer at home, but the thing is, I don't like the digital age 'cause I'm only learning a bit more on the computer. I had to go for a course on it actually, a beginner course, you see? 'Cause when I used to work, we used to work on stock systems you see, that was years ago, in the 90s, but then I haven't been able to work since about 2001, you see? Cause of this illness, you know? But the thing is, I'm not so used to digital stuff, like whizz kids are, like these youngsters today. I'm going to learn a bit more ...

ID 247, outpatients, FG, site B

There was also recognition of the marked variations in uses and preferences regarding digital technology, in that it may make sense for those routinely engaging with it, but could be very difficult for others:

I think a lot of people these days are familiar with using that kind of technology and so ... and the majority of people are comfortable and like to do it in that way. I know that if it was me I would rather, you know, work on something like that than paper and writing things down ... on the flipside of it, you know ... it scares some people ... I'm just sort of going off really sort of personal experience of people that I know in my ... home life who are experiencing some difficulties, you know, with mental health or age related.

ID 242, outpatients, administrator, site B

Some patients and a number of carers described seeing the kiosk in the reception area but dismissed it as they thought that it was intended for staff use only. One carer assumed that she would not be entitled to use the kiosk:

I would think it was for a patient. Only because my doctor has got something similar where you don't have to queue up, if you've got an appointment you just type in your date of birth. And do that, so if I was to see that, I think that's why I was interested in it because it was my sister's appointment, she is a patient, so I didn't know I was allowed to touch that ... I think people just walk in and presume that it's just for patients to let their doctors know that they are there or something like that.

ID 263, outpatients, interview, site B

In addition to the above-mentioned reservations, staff suggested other reasons for the low participation rates, such as a lack of privacy in one site, where the feedback needed to be given via a digital screen located on the reception desk (rather than via a kiosk stand), and patients feeling rushed because they were unsure how long they would be waiting before their consultation.

Verbal feedback in site B: community mental health team

Members of the community mental health team talked about having a sense of uncertainty about the new process, which had stemmed from discussions of many different ideas throughout the co-design process, as well as the lack of differentiation, because of the technical constraints of the existing electronic care record:

I think the ambiguity of purpose has also been a factor because whenever people have ... walked away from these workshops [focus groups] because they've been purposely trying to get people's co-operation and facilitate as many ideas as possible I think, in a sense, then we've not come away with a really clearly defined purpose. So, I think all of the factors we've discussed, plus that lack of clear definition has undermined, really, our ability to do it.

ID 341, interview

Sites C1 and C2: primary care

Staff at one of the primary care sites (site C2) appeared to have a clearer understanding than staff at the other primary care site (site C1) of the purpose and value of the intervention. One of the reasons for this might be that the senior partner and the practice manager at site C2 were both highly engaged from the start.

Patients gave more mixed responses, ranging from enthusiasm to outright rejection. Embracing the idea of the kiosk was related by some respondents to wider societal changes in communication. One person put this clearly:

... you can't avoid this digital world ... it controls everything ... Well, everything's going ... also have the information of that, you've got it in or anything. Information about this, whether it's going in the post from here to the consultant, from the consultant to the powers that be. You know, it's wonderful, it's immediate, and yes, and also the prescription, sometimes you have to wait weeks for prescriptions, you know, have they made a mistake and had to come back. Now everything is digital.

ID 249, FG, patient, site C2

Providing feedback on services was seen as a logical extension of the advantages of the digital world. However, not everyone felt that digital feedback methods made sense as they would not help them overcome their reservations about answering survey questions. One patient expressed this as follows:

I just don't want to go and have a look at it because I don't know the answers to most of the questions. And I feel that I'll be giving a wrong impression.

ID 249, FG, patient, site C2

In certain cases the kiosk made this more difficult because people had to make a conscious choice to enter and felt that this was a public statement. It was therefore not seen as an easier alternative than filling in a paper questionnaire, which could be done in private.

Cognitive participation: investment of staff and information and support needs for patients and carers

The NPT construct of cognitive participation focuses on whether or not participants drive the intervention forward and whether or not they see themselves as the right people to be involved. Cognitive participation stimulates them to buy into and sustain the intervention. One of the key elements in our study was how staff members indicated a sense of investment in the intervention as opposed to detachment from it, as well as how the organisational teams responded to the information and support needs of patients and carers.

Site A: acute trust

In site A, a reasonable level of coherence emerged and there was recognition that testing digital tools to support data capture made sense at the outset. However, during the testing period a number of staff members indicated a degree of detachment from the tools being tested, including the kiosk. They did not think that it was appropriate for them to be seen to be 'promoting' or driving the sustained use of the kiosk:

To be honest with you it just sits there and I don't have any impression about it, it's been installed because of the study hasn't it and I don't encourage or discourage people to fill it out. If they see it and wish to fill it out, they fill it out. But I don't really have any ... I don't ... do you know what I mean?

ID 141, FG, site A

Site B: mental health trust

In site B, a number of comments made by staff members in the community mental health team indicated that there were barriers to sustaining a community of practice to drive the new process forward for recording verbal feedback:

I've heard people speak about it, you know, and say . . . I know one person said 'I don't think I'm the best person to actually collect the information' . . . from a personal point of view, I also think this . . . if we're collecting the data and it's not being recorded, then I wonder if that's because negative things have been said. Maybe care co-ordinators don't want that reflection. So don't want to put it in the notes. I don't know, it's just a thought.

ID 216, interview, site B

Staff members also talked about how, during the testing period, the new system for recording feedback seemed to slip from their agenda, and there was a lack of relational work to stay activated in collectively defining and sustaining the practice:

I think a lot of us have kind of forgotten about it, to be honest with you. It's a good job that you're here today, it's a reminder . . . it's interesting, but it's also a reminder.

ID 341, FG, site B

I've got two clients that I see. What we have agreed, and is what we agreed with staff, is that we should be asking the question about how a person has felt the encounter to be and to record it in 'plan' [referring to specific field in the electronic record]. What we've also acknowledged is that individuals will ask that at whatever frequency they feel is appropriate, but what we encourage them to do is to put in 'plan' 'question not asked', if they don't . . . Now I've been asking staff in supervision and they've forgotten and, candidly, for my own part with my two patients, I have as well.

ID 341, site B

Sites C1 and C2: primary care

In the general practices there were differences in terms of the level of investment of staff to drive forward the use of the kiosk. For example, in site C1, staff members talked about their reluctance to encourage patients to use the kiosk, mainly because of the existing burden of tasks for patients and the existing workload of staff:

You know, so I couldn't even sort of say I'll ask the receptionists to encourage patients, because at the moment we're already asking them to do a lot, remember a lot. So, you know, another job on top of that I think would probably tip them over the [edge].

ID 134, FG, site C1

We'd start off with good intentions, but I think the problem is at the moment, we're doing that many different things that you're trying to remember, because we're doing like QOF [Quality and Outcomes framework], [local] standards, we're trying to get . . . by the time you've done your consultation and gone through all the various bits, then we might remember. But it would be very hit and miss if I'm honest, because you're usually quite pleased if you've managed to make all the boxes go away, as well as do a consultation. And I would imagine, certainly at my age, I'd forget.

ID 141, FG, site C1

In site C2, the staff with a managerial role helped to advertise the kiosk and to involve the PPG in the core phases of the project from the outset. As a result, new ways of advertising the study were discussed and agreed with the practice team and PPG to encourage patients and carers to use the digital kiosk. The PPG saw this active involvement as being appropriate for their role and function.

Field notes made during a PPG meeting in site C2 demonstrated that the group worked proactively with the practice manager to plan ways to help sustain and further enhance the routine use of the digital kiosk among patients and carers in the practice:

- Practice manager [PM] will upload project information onto the practice Facebook page as a new method of advertising the toolkit to get feedback on specific clinics.
- Advertisement and the next steps (including testing out bespoke questions to gather useful data that will help with staff revalidation) will be discussed in next week's staff meeting.
- The PM and project researcher will pick this up at the next PPG meeting.
- The PM to place a project advert in both the practice and PPG newsletter.

Observation note, site C2

The kiosk at site C1 was viewed by one staff member as a good alternative that might stimulate people to provide feedback:

I think the screen is an attraction because it's standing there, and people might think it's quicker than hanging about if they've not got a pen, if they make a mistake on the form, whereas if they make a mistake, they can go . . . if they know how to go back on it, to change it.

ID 134, interview, site C1

Collecting patient and carer perspectives was considered imperative for service delivery, and making it easy for people to provide feedback was one of the kiosk's benefits:

I think it's really good to get patient feedback as we do . . . well, we do Friends and Family [test], and the more information, the more we can find out how the patients feel. So having 'something' in the waiting room is beneficial to the practice.

ID 333, staff interview, site C2

Observation in the centres indicated that patients were apprehensive about using the kiosks; most patients would not use the kiosk spontaneously by themselves. Various members of staff noted this hesitation:

I find a lot of people are quite reticent about things as well, you know. They might sort of look at it and think 'oh I wonder what that is', but then not get up and have a look at it.

ID 137, FG, site C1

Participation was enhanced if support was given by the researchers, staff or volunteers, when on site. Most patients found the kiosk easy to use, but they also highlighted the need for alternative ways to provide feedback, as digital methods may not be suitable for everyone.

Collective action: organisational and technical work for sustaining new tools

A third NPT construct, collective action, reflects the fact that work has to be maintained to operationalise a new intervention so that it has a chance of becoming embedded. Therefore, participants need to consistently fulfil the tasks required and maintain confidence in the new approach. For this to happen, the right people need to be involved and adequate support has to be available. In this study, we focused in particular on internal team communications with regard to the feedback process and whether or not any actions resulted from discussions.

Site A: acute trust

In site A, objections centred on difficulties related to the working of the kiosk, including several operational issues, such as ongoing screen-freezing issues, no Wi-Fi connectivity [manual uploads were made via an encrypted universal serial bus (USB) drive] and the location of the kiosk, which was not ideal for maximum visibility:

I recently had a rheumatology appointment and I couldn't actually see the machine. I had to ask somebody where it was and where it was positioned, I think was a bit awkward, it's not really visible ... After I'd asked somebody where it was, I then went over to try and record my feelings and it wasn't working, it said it had logged out or there was an error or something like that.

ID 107, patient, site A

The kiosk at this site was placed in a very busy outpatient clinic. Patients and carers were frequently called away while using the kiosk and it was not possible to save typed comments and work on these again when free:

One patient was willing to give feedback after getting consultation from the doctor while she was still waiting for blood test to be done. When she was typing on the kiosk, the nurse called her for blood collection. She then left the kiosk immediately and gone for blood test. The page that she was typing disappeared when she returned from the blood room. She had to type everything again.

Observational note, site A

Staff members were unsure who would take on the role of promoting the kiosk, highly pressurised with their current roles and either unable or unwilling to take on this responsibility. The timing of feedback collection was also a barrier to use of the kiosk; some clinicians thought that that it was 'unfair' to ask people to provide feedback directly after a consultation, which was described by one consultant as 'forced feedback' (ID 3211).

Other responses included:

I don't think many doctors are mentioning their patients to provide feedback ... I had a brief chat with Dr [name], she thinks it's better if I tell the patients to give feedback instead of doctors telling them, otherwise it might influence the patients' feedback.

Observational note, site A

So what we personally feel like if the consultation has been good, it becomes a bit unfair to ask them go and give us feedback.

ID 3211, FG, site A

Clinicians also admitted that they had not seen the flyers advertising the new tools for providing feedback, despite the research team and lead giving these out during observational sessions and at team meetings:

These little flyers here, I'm ashamed to say I've not seen these before. Where were they?

ID 3211, FG, site A

Site B: mental health trust

In site B, the staff in the outpatient department raised concerns about the ongoing impact of the new technologies on aspects of their role and workload:

... if I'm having to come away from my desk to talk them through it, even through the glass, that is going to take me away from the phones, it's going to take me away from what I'm supposed to be doing.

ID 262, outpatients, FG, site B

This staff member clearly distinguishes their primary function from what they consider to be a distraction, in this case having to explain how to operate the kiosk. More generally, the importance of detailed and specific feedback that takes account of clinical concerns was highlighted, and many staff felt that the ideal way to collect feedback was through face-to-face discussions. Constructive team discussions led to the development of a new process for giving and using feedback generated through interactions and, in the case of the community mental health team, during home visits. However, although there was enthusiasm among team members for collecting feedback in these ways, adoption of this method was slow and very few data were recorded during the evaluation period.

In site B, the team experienced major changes in the structure and leadership of the organisation, and in the community mental health team there were many problems with long-term sickness and a high staff turnover rate, such that the team seemed unable to collectively build a shared accountability for the new process for recording verbal feedback:

Because of staff absence, people are having to pick up other people's work to some extent or another, on top of their own work, and they've already got really high caseloads.

ID 216, interview, site B

It was clear that members of the community mental health team faced barriers to operationalising the new process for recording verbal feedback, in terms of allocating the work within the structure of the team, but also in terms of building confidence in the new practice, as well as adapting the existing technology of the care record to try to make this work:

Well, the, I mean, the area we'd identified was, needing to put the narrative response in ... the box ... And I think the problem ... was about a question ... And at the time we sort of did it as an open exercise, to practice ... then it became an issue of, well, which type of question ... perhaps the ambiguity as well of how work is conducted can make it difficult because if there is an ambiguity of purpose then it's harder to ask a question, and, equally, what kind of the answer are you going to get if the person on the other side equally shares that ambiguity?

ID 341, interview, site B

This also illustrates some of the overlaps with concerns raised regarding coherence:

... we did agree, though, that we didn't have to ask on every single visit ... because then it gets a bit too repetitive, and ... uncomfortable.

ID 341, interview, site B

Sites C1 and C2: primary care

When discussing responsibilities and roles in site C1, one of the staff members said:

... as far as I'm concerned it's a research project, I don't know what our responsibilities are ... I wouldn't touch it with a bargepole because I wouldn't know what I'm doing.

ID 141, FG, site C1

The viability of the intervention here had been affected by a recurring problem with the screen becoming frozen, and it sometimes took some time before this was noticed; the team had also experienced difficulties in arranging for the external suppliers to attend to the technical problems (site A also suffered such operational issues, but our other two sites did not).

Also in C1, efforts to encourage interaction with the kiosk appeared to have subsided over time. Staff had originally decided that one useful way of informing patients that they could give feedback via the kiosk was by advertising this on patients' prescriptions. However, when this was subsequently

discussed, there was lack of clarity about who had responsibility for this and whether or not it had been carried out:

... but what we haven't discussed is how as a team we can promote it. At the moment, it's just been really sort of feeding back. It is on the prescriptions, but I have noticed today it's very hidden, it's not that noticeable. It's small print and if there's other messages on there, I mean, the one I looked at today, it was sort of in between two other messages, one about the flu, another one about reordering your prescriptions, and then the DEPEND was in the middle of that, it's very tiny.

ID 134, FG, site C1

In contrast, the handwritten summary reports outlining the volume and quality of free-text comments were well received by the team in site C2. The positive and negative feedback prompted a lengthy discussion in the staff meeting on the core areas of concern, or what caused expressions of gratitude. Most of the monthly negative feedback in general practice was concerned with access and waiting times. Staff emphasised the importance of patients having enough information about how the general practice system works to help ease the flow of concerns:

We could have a noticeboard to say positives and room for improvement, and then say these are the ways we're thinking of working on these issues, if you have any suggestions please provide further comments, and then we're getting a bit of feedback rather than just from the patient participation group, aren't we?

ID 333, FG, site C2

All staff and the PPG volunteers reported that the majority of patients would not use the kiosk unless they were asked to do so. Signposting was changed and use increased once the larger colourful laminated poster with guidance notes was in place above the kiosk:

The PPG volunteer was active in promoting the kiosk in the reception area. Five people walked up to the poster (this has never happened before with the smaller A4 landscape unlaminated poster), read it for a few minutes, then used the kiosk without being prompted by either of us.

Observation note, site C2

Work was carried out by the PPG in site C2 from the outset, and in particular throughout this WS. During the introductory testing period, two PPG members provided peer guidance during the busiest clinics on how to use the new kiosk. Having this component of peer support in place allowed data capture on the acceptability of and continued engagement with digital feedback by patients and carers:

The PPG volunteer noticed that younger people seemed to be a lot quicker at typing their experiences, whereas the middle-aged people took their time, and the elderly wanted to try the kiosk with the help from a volunteer.

Observation note, site C2

The presence of a PPG member or researcher meant that support could be offered to those who had difficulty using the kiosk. PPG members saw this as an appropriate task and confidence in the intervention could be maintained. The following examples show how the use of the kiosk was supported:

A middle-aged man needed help from the PPG volunteer to type as he had visibly shaky hands; the PPG volunteer ended up typing on his behalf.

Observation note, site C2

One lady had painful rheumatoid arthritis in her painful fingers. She explained if we provided a touchscreen stylus as an alternative to a biro she would certainly have a go, or voice dictation software.

Observation note, site C2

Despite the physical assistance at the kiosk, many patients still expressed reservations:

One lady saw me standing next to the kiosk with my university lanyard on and started to tell me verbally about her experiences at the practice. When I asked her would she consider using the kiosk, she said she preferred to tell me in person rather than tell a machine and wouldn't use it, even with the offer of guidance from a volunteer.

Observation note, site C2

The low rates of participation during WS4 and the views of staff and patients drew attention to a number of organisational and technical barriers. The workload of staff was often highlighted, which was perceived to negatively affect the capacity to take responsibility for the kiosks and recording of data. Staff did not feel that they had the time to support and motivate patients to give feedback. Some felt that it was ethically wrong to ask patients for feedback in case they felt pressurised to give positive feedback. This could have adverse implications for the therapeutic relationship.

The spatial position of the kiosk was highlighted in all sites as an issue but no agreement was reached about the optimal position. Technical problems encountered included the machine becoming frozen and uncertainty around allocating responsibility for maintenance (including hygiene).

In site C2 the lead GP and staff with a managerial role took on the main tasks of ensuring that the kiosk was used as planned, whereas in site C1 the intervention was not supported to the same extent.

Reflexive monitoring: embedding the new intervention

The final NPT construct, reflexive monitoring, was concerned with participants' perspectives on the effects of the new tools and whether or not they judged these to be worthwhile, both individually and collectively. It was also important to ascertain if individuals and/or teams made changes to their work as a result so that the new practice could be embedded in routine practice.

Although staff were positive regarding the feedback reports generated during the project and found these to be helpful for discussing issues raised, we did not observe that they influenced changes to service delivery during the evaluation period. The follow-up period was too short to state whether or not the new approach has become embedded, but in two sites medium-term change seems to be happening in that the kiosks have been retained.

Site A: acute trust

The monthly reports (see *Appendix 7* for an example) summarising feedback collected via the kiosk were found by staff to be useful to present a potted summary of monthly feedback at team meetings:

I think this will be very useful. I mean we've just had a governance meeting so presenting this every month would be very useful.

ID 131, FG, site A

Having the feedback analysed and presented in this simple format elicited discussions around potential changes that could be made to clinical practice; this was often not the case for the generic feedback currently collected. The team also expressed a preference for tailored questions over generic FFT

questions to elicit specific and detailed free-text comments that would be meaningful for service improvement:

'... there is nothing in this report quite specific which needs to be changed in order to improve our Dept'. She continued, 'some feedback will go to nowhere ... you can be friendly and nice and useless at the same time'. What is the use of getting this feedback? The question can be asked differently, ask them specific questions, like access, or blood ... Only then we may be able to use the report to improve this department.

Feedback meeting, observational note, site A

Similar to the other sites, the team described any negative feedback received in the reports as system constraints, which they described as being 'beyond their control'.

... we actually cannot take anything forward from this report, except two things, which are two complaints ... here we have picked up some structural data, not really about the Rheumatology Unit (e.g. access and waiting time).

Observational note

Reflective discussions on the monthly reports focused on some of the limitations of the data and particularly the need to set the data in the context of the wider organisation and being able to obtain data that would stimulate change via management:

... it's interesting because it's the hospital systems around the specific consultation wait and we know the problems. It's a problem with car parking, it's a problem with no nurses, it's a problem with ... I mean I know what the problems are. It's getting management to do something about it and showing them that this change is the only way I think. Or benchmarking them against someone else, saying we're doing worse than others. So it's some way to get some change to benchmark against other units, other departments, I think it would influence management. Perhaps not the clinical stuff that we would control, for which this text is fantastic. I mean I think that's very useful and the type of outcomes, what you described. But I think I would like something that might influence management in a more structured way to say we need something done would you sort it out.

FG, site A

Site B: mental health trust

One of the senior managers in site B reflected on the value of the new digital tools at the level of the trust as a whole organisation:

I think as an organisation it's been a hugely helpful experience really. I think what it's obviously told us is we really want to generate the quantity and create something that's smart and it's easy then to filter back down and attribute to individual wards and all the rest of it.

ID 201, interview, site B

However, staff with a managerial role also reflected on the barriers and ongoing challenges that the organisation faces that created problems with regard to the adoption of new practices for collecting verbal feedback in the community mental health team:

I think generally all of that change as we went through as an organisation, if you think about that in the context of [manager's name]'s team, I think from a timing point of view it was probably really unhelpful for the research. Because I think what was happening, certainly in my case there was change, there was change in circumstances, there was new priorities. There were new things that we needed to focus on as an organisation. And I think sadly what that meant was there were a lot of things that we just couldn't do anymore, or that became really difficult to do because there were new things, other things that we had to worry about.

ID 201, interview, site B

Other staff also cited changes in the trust as reasons for the lack of adoption of the new process:

A year ago we were taken over by a new trust, reached a consultation period of proposed changes. A lot of staff aren't happy with some of the changes. A number of staff have left. We've also had long-term sickness on the team. So it's led to no consistency for service users and carers. In terms of carer support, we've also almost simultaneously to the takeover of the new trust, we've had changes in processes as well. A new care assessment was developed that's taken time to learn ... new processes for processing the paperwork. Takes a lot more time now.

ID 216, interview, site B

Sites C1 and C2: primary care

The staff in both site C1 and site C2 thought that the handwritten summary reports were useful and provided timely feedback that could be discussed face-to-face at staff meetings rather than by e-mail. In one observation of a practice meeting (site C2) at which the monthly report was being discussed, frequent negative feedback comments on 'access' and 'waiting time' were rationalised as staff not having the capacity to do anything about this within the wider constraints of general practice. Examples of comments made on this theme included 'phoning takes ages', 'told to attend hospital', 'very difficult to get appointments', 'never appointments' and 'had a wait but not usually that long'. In the meeting, staff discussed this feedback but this was restricted to a shared sense of frustration that these problems were all part of wider pressures faced in primary care.

In the same meeting, staff reflected on some of the positive feedback that they had received and the value of this for making staff feel positive. Examples of comments made on the theme of 'staff attitude and professionalism' were 'you have always supported my family through bad times and still are. I have always felt very lucky to have such good family doctors that includes [name]', 'always helpful doctors and reception staff', 'staff always helpful', 'front desk service polite informative and welcoming', 'lovely friendly doctor very informative' and 'listened'. At this site, staff expressed a preference for feedback to identify the relevant members of staff so that the feedback could be used for revalidation purposes. At this same site, staff were keen to optimise PPG engagement in the current study by involving the PPG in future feedback staff meetings, in addition to holding separate PPG meetings. It was agreed that the PPG would stay involved and should see the provision of support to patients and carers as part of their remit. The GP who set up the remit of the PPG in site C2 as a research project (see Chapter 2) is keen to sustain the digital kiosk at this site as a future case study after hearing about the successful PPG engagement. The physical presence of the kiosk and structural change in the PPG role combined may prove important factors that facilitate the adoption of digital feedback as routine:

So I think the kiosk has been positively viewed ... I can't think of any negative comments about the kiosk. There may have been, but I'm not aware of them. It's generally been positively viewed. It seems to have been used by quite a number of people. I think it's still got potential to have ... to be used more for different and innovative things, but as a first trial it seems to have worked very well and people seem to be engaging with it.

ID 335, interview, site C2

Patients and carers here appeared to be very positive about giving feedback through digital platforms when discussing this with researchers. The majority of people using the kiosk found it easy to operate, whereas those requiring assistance felt that they could complete the survey without difficulty. However, despite the positive response to the digital methods, people stated that they would like this to be offered alongside other, more personalised options such as face-to-face and narrative approaches. The resource implications of providing a menu of options will need to be more fully understood before this can be implemented as routine practice.

Summary and discussion

The importance of context has been demonstrated in this study and, in particular, the comparison between the four sites allows a deeper understanding of the processes at play. Furthermore, patients' and carers' perspectives differ from those of staff and differentiation occurs between staff groups because of their different organisational roles. First, the co-design approach facilitated sense-making and initially staff and patients expressed positive thoughts about digital methods. Patients emphasised, however, that these methods should be offered as part of a wider menu of feedback options, including face-to-face approaches. This was felt to be particularly relevant given the huge variation in digital awareness among the general population. Senior staff saw the efficient collection and analysis of digital data as the main advantage and front-line staff thought that it would make it easier for them to routinely collect data.

Second, in terms of participation, variation between the sites became apparent and, where clear 'product champions' were present, participation rates were highest. Such champions ensured that staff were clear about the purpose of the process of feedback and what was required from them. In the primary care site where the PPG was closely involved, peer support enhanced participation. In the acute trust, mental health trust and one of the primary care sites, a gap between espoused enthusiasm about the concept of digital feedback and actual practice emerged. Various reasons were offered such as the pressure of work for staff or perceived incompatibility with existing systems and patients feeling unsure whether or not the kiosk was for their use. The site where participation was highest showed that continued support from key staff and assessment of the impact on occupational roles helped to overcome these barriers.

Third, collective action varied between the sites and the organisational context emerged in the mental health site as the prime variable. A major organisational restructuring took place during the lifetime of the study, with key staff changes and new priorities having an impact, especially on the ability of staff in the community mental health team to operationalise the new verbal feedback process. Thus, the drive from 'product champions' disappeared and, coupled with uncertainty about the new structure and management, staff disengaged from the implementation process. In the acute site, participation declined as many staff did not make changes to their work as they felt that they should not be actively promoting digital feedback. Thus, they did not draw patients' attention to the kiosk or guide them if they did not know how to use it. In the primary care site where the lead GP and practice manager acted as the driving forces, regular assessment of the digital feedback took place in team meetings, followed by remedial action needed to maintain participation. Moreover, the PPG accepted that peer support had become an integral part of their role and thus collective action was enabled across staff and the PPG.

Finally, although it may be too early to formulate robust findings on reflexive monitoring, there were indications that staff considered the evidence to date to form views as to the usefulness of digital feedback and the chances of it becoming routinised. In the mental health site an understanding of the barriers to adoption emerged and some ideas of how to overcome these were being formulated. In the acute site staff described the benefits of receiving structured feedback but this was too general to do anything about and they expressed a preference for collecting more tailored feedback to inform their clinical practice.

In primary care site C2, plans for sustaining the kiosk were being advanced, coupled with clear ideas about the organisational structure and roles. The essential contribution of the PPG was recognised and included in forward planning. Across all sites, patients' preferences for a choice of feedback methods were recognised. Staff described benchmarking as a way forward to enable change at a systems level and digital methods as a way of capturing longitudinal feedback.

Adopting a theory-based framework such as the NPT has allowed the analysis of the implementation of digital feedback to be more context specific and to take account of the multiple actors involved in formulating and actioning the work required. Analysing the diversity of perspectives provides insights into how interventions should be targeted and tailored to specific needs. Continual adaptation to changing circumstances is integral to ensuring that an intervention remains relevant and thus embedded in daily practice.

Health economics

This section focuses on the resources and costs involved in developing and implementing the toolkit, addressing the following questions:

- What are the costs of the co-design activities carried out to develop the toolkit components?
- What are the costs of developing the text-mining and reporting elements of the toolkit?
- What are the costs of initial implementation of the toolkit/kiosks in each of the sites?
- What are the costs of analysing and reporting the data generated by the toolkit/kiosks?

Costs of developing the toolkit components

Table 12 summarises the NHS and research staff and service user/carer time spent participating in co-design meetings, interviews and focus groups to develop the toolkit and kiosk intervention. Appendix 8 (see Table 24) provides more detail about the service use/carer and staff time used for the co-design activities. There was some variation between the sites in the number of activities and quantity of staff time associated with the co-design activities, with site B having the highest number of focus groups and participants. Some of the activities were shared between sites. Accordingly, only the total time is reported for service user/carer trust lead meetings and NHS staff interviews, PPI meetings, focus groups and interviews with service users and carers and research staff time. The main resource used was research staff time (2869 hours) to facilitate and analyse the co-design activities and develop the components of the intervention. These activities also involved 216 hours of NHS staff time and 136 hours of service user and carer time.

Table 13 summarises the costs associated with the co-design activities and analysis of the qualitative data. These data are reported in more detail in Appendix 8 (see Table 25). Staff time was costed using the salary and on-costs of NHS and university staff, by staff category, and service users' and carers' time was costed according to the reimbursements paid to participants. The total cost of NHS staff and service user and carer time to participate in the co-design work was £9223 (£7513 for NHS staff and £1170 for service users and carers). The cost of research staff time to facilitate the co-design meetings, conduct interviews and focus groups and analyse the qualitative data was £84,831. The total cost of the activities carried out to co-design the toolkit/kiosk component of the intervention was £94,054.

TABLE 12 Staff and service user/carer time spent on co-design meetings, focus groups and interviews

Activity	Staff time (hours)				Total
	Site A	Site C1	Site C2	Site B	
NHS staff					
Focus group/meetings	35	41	53	68	195
Individual interviews (hours)					21
Total NHS staff time					216
Total service user and carer time					136
Research staff					2869

Time spent is rounded to the nearest hour. Due to this rounding, some totals may not correspond with the sum of separate figures.

TABLE 13 Costs of staff time and service user/carer time to develop the toolkit/kiosk (2016–17 prices)

Activity	Cost of staff time (£)				Total
	Site A	Site C1	Site C2	Site B	
Co-design meetings and focus groups					
NHS staff: focus groups	1677	1254	1974	1880	6785
NHS staff: individual interviews					728
Total cost of NHS staff time					7513
Service users and carers					1710
Research staff					84,831
Total cost of all staff time					94,054

Table 14 presents the staff time and costs to develop the text-mining tools and design the interface, reporting templates and information materials for the toolkit. These were costed using the salary and on-costs of university staff. The total cost was £59,417.

Costs of implementing the toolkit/kiosk intervention

The staff time and resources used to implement the toolkit/kiosk intervention are shown in *Appendix 8* (see *Table 26*). *Table 15* reports the costs of implementing the toolkit/kiosk intervention. The resource use and costs of staff induction across the sites and for video production in one site are included as implementation costs on the assumption that these will need to be regularly updated; however, they could be considered as development costs if this is not the case.

For the 9-month evaluation period, the main resource used was research staff time to analyse the data and provide reports to each of the four sites. This is reflected in the costs in *Table 15*. The total cost for the 9-month implementation period in the study was £23,589. If all costs are included, the annual cost was estimated at £31,452. If staff induction and video production are considered to be development activities rather than implementation activities, the estimated annual cost is £4833 lower at £26,619.

Total costs

Table 16 summarises the total costs of developing the toolkit/kiosk and implementation of the intervention in the four sites. This assumes that development and implementation costs are shared equally between the four sites. *Appendix 8* (see *Table 27*) explores how the costs vary if these costs

TABLE 14 Staff time and costs associated with development of the text-mining tools, interface design and reporting templates and information materials (2016–17 prices)

Staff activity	Time (hours)	Cost (£)
Developing text-mining tools	1607	43,745
Interface design	510	10,761
Reporting templates and information materials	161	4911
Total	2278	59,417

TABLE 15 Costs of implementing the toolkit/kiosks: 9-month evaluation period (2016–17 prices)

Activity	Costs (£)				
	Site A	Site C1	Site C2	Site B	Total
Kiosk rental (block contract)	NA				10,000
Induction: NHS staff	151	54	54	46	305
Induction: research staff	69	69	69	207	414
Kiosk support: research staff	NA				2419
Kiosk support: volunteers	NA				330
Information materials: research staff (hours)	NA				955
Video support: video production	0	0	0	1500	1500
Video support: NHS staff	0	0	0	247	247
Video support: research staff	0	0	0	204	204
Data analysis: software	NA				475
Data analysis: manual coding and report, research staff	NA				2756
Data analysis: auto coding and report, research staff	NA				929
Data analysis: populating report templates, research staff	NA				713
Data analysis: report to sites and discussion, research staff	586	586	586	586	2342
Total cost, 9 months	NA				23,589
Estimated total cost per year, all costs	NA				31,452
Estimated total cost per year, excluding staff induction and video production	NA				26,619

NA, site-by-site breakdown not available.
Cost is rounded to the nearest pound. Due to this rounding, some totals may not correspond with the sum of the separate figures.

TABLE 16 Total annual costs per site by length of annuitisation^a (2016–17 prices)

Assumption	Costs (£)				
	Site A	Site C1	Site C2	Site B	Total
Development costs excluded	7267	7137	7137	9912	31,452
Development costs annuitised over 1 year	61,297	60,634	61,594	63,551	247,076
Development costs annuitised over 2 years	33,853	33,468	33,956	35,975	137,253
Development costs annuitised over 3 years	24,709	24,416	24,747	26,786	100,658
Development costs annuitised over 4 years	20,140	19,894	20,146	22,194	82,373
Development costs annuitised over 5 years	17,400	17,182	17,387	19,442	71,409

^a Assumes equal allocation of shared costs.

are allocated pro rata according to the site-specific development costs. Table 16 also illustrates how the costs vary depending on assumptions about the lifespan of the toolkit/kiosk intervention. The development costs were converted to annual equivalent costs (annuitised) by discounting them over the assumed lifespan of the intervention, using a discount rate of 3.5%, applied at the beginning of each year.

If the development costs are excluded, then the estimated annual cost per site is the implementation cost only, which ranges between £7137 and £9912. The higher cost for site B is a result of the additional costs for video production and support. If the development costs are assumed to have a lifespan of 1 year, then the annual equivalent costs of development and implementation per site range from £60,634 to £63,551, with a total cost across all four sites of £247,076. If the toolkit/kiosk is expected to have a lifespan of 5 years, the annual equivalent cost of development and implementation across all four sites is £71,409, with a cost per site between £17,182 and £19,442.

Costs of the Friends and Family Test

The review of published research and Department of Health and Social Care policy and guidance indicated that very little is known about the local costs of implementing the FFT. Guidance about the implementation of the FFT locally indicates that there is no additional funding for the test and that it is expected to be resourced within overall budgets of individual providers.⁶¹ The guidance also makes it clear that local providers can be flexible in how the test is implemented and whether it is implemented in-house or outsourced.⁶¹ One example of outsourcing the provision of the test indicates a cost of around £40,000 per year.¹⁰ If the costs to a local provider are primarily for data entry of paper surveys and analysis of the data, the costs of the FFT may be substantially lower than this, depending on the size of the trust and number of patients who complete the test. For example, Manacorda *et al.*⁹ indicated that the FFT is perceived to add little extra work for staff in a primary care practice. The national costs of central support for the FFT are expected to be up to £1.5M per year.¹⁰

Overall, the estimated annual equivalent costs of implementing the toolkit/kiosk may be in the range of those incurred by trusts that tailor the FFT for their own patient population and information needs. For providers with smaller populations and response rates the annual costs of implementing the toolkit may be relatively high, depending on the lifespan of the toolkit and how the shared development costs are allocated.

Whether or not the toolkit/kiosk as developed is value for money will depend on the relative effectiveness of the intervention in achieving trust aims and/or quality improvement. The evaluation of the toolkit/kiosk intervention demonstrated the feasibility and acceptability of the approach. However, the evaluation was not designed to assess the relative effectiveness of the intervention compared with current methods of administering and using the results of the FFT at a local level. Nevertheless, the resource use and cost information collected as part of the evaluation provide useful information to inform future development and evaluation of methods to collect, analyse and use patient experience data.

Text mining versus qualitative analysis of free-text feedback received in general hospital and mental health service settings: a descriptive comparison of findings

Text mining: comparison of service types

Comments by mental health service users were framed much more by 'care quality' than general hospital comments, which were more focused on 'staff attitude and professionalism' (Table 17). Whereas roughly equal proportions of general hospital comments about 'care quality' were positive and negative, proportionately more mental health service comments about care quality were positive (57%) than negative (40%). Differences concerning the 'environment' were limited, although mental health service users were more likely to be critical in this regard. The proportions of negative comments around 'staff attitude and professionalism' were similar between both data sets, although general hospital users appeared more likely to express positive sentiments in this regard.

The main differences identified between the two data sets/health service settings in text-mining output were as follows. The high volume of words classified as 'not feedback' in the general hospital data set seems to reflect the use of mobile phones as data collection tools. Care quality was more of a focus in mental health service feedback; although it was equally of positive or negative

TABLE 17 Text mining: descriptive comparison of findings between mental health and general hospital settings

Theme	General hospital, % of the whole sample		Mental health, % of the whole sample	
	Positive	Negative	Positive	Negative
Care quality	11.950	11.006	56.757	39.623
Environment	2.516	1.572	2.703	7.547
Other	2.673	6.604	0.901	9.434
Staff attitude and professionalism	35.849	13.941	24.324	14.151
Waiting time	5.084	8.805	0.000	10.377

sentiment in general hospital settings, higher positive sentiment was seen compared with negative sentiment in mental health settings. 'Environment' was the least populated topic; the 'environment' was more of an issue in mental health services and there were more negative than positive sentiments. The ratio of positive to negative comments around staff attitude and professionalism was roughly similar between service settings, although the general hospital setting had proportionately more positive feedback. In respect of waiting times, whereas both positive and negative sentiments were provided in the general hospital feedback, the mental health service feedback was overwhelmingly negative.

Qualitative results

The qualitative analysis included three steps: (1) an early descriptive survey of around half of the data mid-way through the coding process, aimed at identifying potential differences between the two service settings (see *Report Supplementary Material 4*), (2) a detailed presentation of count data, organised by sentiment and service setting, to facilitate comparison with the text-mining results presented in the previous section (see *Report Supplementary Material 5*) and (3) a focused descriptive account, using exemplar patient comments, to explore the differences identified at step 2 (see *Report Supplementary Material 6*).

Comparison between count data obtained using the different analytical methods

Table 18 provides a comparison between count data obtained using text mining and count data obtained using AGT.

Having count data available from both methods allows a comparison of some aspects of the results in a more objective fashion. However, caution is required for two main reasons. First, both methods are based on different epistemological foundations and have conceptualised the feedback in subtly different ways. Both methods potentially involve multiple counting of comments (or parts of comments) especially in the AGT method. Some of this variation may also be because the qualitative analysts were able to incorporate some comments that the text-mining algorithms categorised as 'junk' or 'not feedback'. The analyses undertaken so far do not address this specific aspect. All of this would suggest that the AGT analysis has the potential to 'magnify' or exaggerate sentiment, but no discernible pattern is evident. However, there are some instances where this may have been an issue: in the general hospital care quality and staff attitude and professionalism. Reference to *Appendix 6* shows that many different descriptors were used for these aspects and they may have been magnified through a process of disaggregation followed by multiple counting.

TABLE 18 Count data: text mining versus adapted grounded theory

Categories (text mining)	General hospital (text-mining denominator = 1731; AGT denominator = 2114)		Mental health (text-mining denominator = 200; AGT denominator = 260)	
	Positive sentiment	Negative sentiment	Positive sentiment	Negative sentiment
Care quality	Text mining = 228 (13.17%); AGT = 595 (28.15%)	Text mining = 210 (12.13%); AGT = 150 (7.10%)	Text mining = 78 (39.0%); AGT = 98 (37.69%)	Text mining = 52 (26.0%); AGT = 74 (28.46%)
Environment	Text mining = 48 (2.77%); AGT = 94 (4.45%)	Text mining = 30 (1.73%); AGT = 56 (2.65%)	Text mining = 4; AGT = (1 + 1)3 = 2	Text mining = 8; AGT = 1
Staff attitude and professionalism	Text mining = 684 (39.51%); AGT = 1053 (49.81%)	Text mining = 266 (15.37%); AGT = 53 (2.50%)	Text mining = 31 (15.50%); AGT = 59 (22.69%)	Text mining = 15 (7.50%); AGT = 36 (13.85%)
Waiting time	Text mining = 97 (5.60%); AGT = 63 (2.98%) ⁷	Text mining = 168 (9.71%); AGT = 50 (2.37%) ⁸	Text mining = 0; AGT = 0	Text mining = 12 (6.0%); AGT = 0 ¹¹
AGT, adapted grounded theory.				

With all of this in mind, the principal differences or similarities seen between services as demonstrated by the two methods were as follows:

- Care quality. The results were the same for the mental health service setting, with both methods finding sentiments approximately 38% positive and 27% negative. In the general hospital setting, text mining found an equal proportion of negative and positive sentiments (12.13% vs. 13.17%, respectively), whereas the AGT analysis found sentiments 28% positive and 7% negative.
- Environment. Both methods produced similar results with regard to general hospital feedback (site A). The AGT results on environment for mental health services show that there were fewer comments than on the text-mining output, but all numbers involved in this category were small.
- Staff attitude and professionalism. Both methods identified this as the biggest component of feedback in the general hospital setting. However, and as for 'care quality', the AGT analysis appeared to accentuate positive sentiments such that more positive and fewer negative sentiments were identified than in the text-mining output. A similar result was found for the mental health service feedback, although the proportions involved were smaller.
- Waiting time. Differences here seemed to run counter to the differences outlined for the previous categories. In the general hospital setting, text mining found more negative (10%) than positive (6%) comments. In the AGT analysis, a small number of comments were more equally distributed between positive (3%) and negative (2%) sentiments. The AGT analysis found that mental health service users' comments under this theme tended to focus on the length of wait prior to commencing care within a service, which was also seen in the text-mining output.

Chapter 5 Discussion and conclusions

In this study we have generated new insights into the problems of existing methods of collecting and using patient feedback data in different health service settings for people with long-term physical and mental health conditions.

We have also created and tested new digital and non-digital tools to support the collection, analysis and presentation of patient feedback. These new tools require resources beyond the capacity of this study to sustain participation and to enable informed changes in service delivery. However, we have generated extensive learning towards the aims of the project, to enhance the credibility, usefulness and relevance of patient experience data using digital data capture and to enhance the analysis of narratives.

Our findings are discussed in the following sections in relation to the four WSs.

Improving the collection and usefulness of patient experience data: perspectives of patients, carers and staff

It has been stated by some that we already collect sufficient data and that we need to shift attention to action in response to data.¹¹ In our study, many patients and carers reported little experience of giving feedback, often because they had not been asked, and there was a lack of understanding regarding the purpose of feedback. Staff across all settings viewed low participation rates and the selective nature of feedback (e.g. reporting of extreme experiences by a selective proportion of patients) to be a problem. Our findings highlight that rates of participation (especially for particular patient groups and carers) remain a concern. There is a perceived need for greater and more routine participation among patients to generate insights that will be perceived as credible and useful to staff.

Previously, concerns have been raised regarding the timing of feedback and the need to enable greater capture and use of timely feedback as a means of ensuring patient safety.⁷ In our study, many patients, carers and staff thought that digital methods could help overcome some of the limitations of existing methods by increasing the range of ways in which people might give feedback when the timing is right for them. Some patients thought that they would be more likely to give rapid feedback at the point of receiving care if they were enabled to do this through the use of digital devices in waiting areas, whereas others thought that they might give feedback online at a later time (if clearly signposted) when not rushing following an appointment. However, there was a recognised need for support for many patients and carers to enable them to give feedback digitally, as well as a need for flexible solutions to enable participation by people who choose not to give feedback digitally. Respondents highlighted the importance of enabling feedback using traditional pen and paper methods, as well as enabling verbal feedback anonymously (e.g. by answerphone), or even openly and in the context of usual health-care interactions and consultations.

There has been much critical discussion of the limitations of generic survey questions and specific criticism of the continued policy of mandating the collection of feedback using the FFT when there has been wide acknowledgement that it generates few useful data.¹⁰ In our study, there was universal acknowledgement among patients, carers and staff that data should be more meaningful than those generated from the current brief surveys (such as the FFT) and that there needs to be greater awareness of the value of positive feedback, rather than an assumption that feedback is worth giving only if it is negative. All groups viewed narrative comments alongside structured questions as providing the detail to allow more meaningful analysis of generic survey questions. Such narrative comments were viewed as being important for explaining scores that were otherwise meaningless. For example, someone might select an average category (neither likely or unlikely) but might provide comments with details of specific good and bad experiences. There was acknowledgement that narrative comments are useful for informing service improvements only if the data are adequately accounted for in formal analyses.

Digital collection and automated analysis using text mining was considered to be particularly valuable as a means of managing narrative data. Managers were particularly concerned about time and resource constraints with regard to the analysis and reporting of patient feedback, including free-text comments. In the large trusts, a small number of senior nurse managers had responsibility for assessing narrative feedback and this was largely carried out by 'eyeballing' the data separately from the analysis of the quantitative data from structured questions. General practice sites (C1 and C2) dealt with a smaller number of data and reports were generally prepared by the practice managers. However, they also felt that they had very limited time and capacity to process and summarise the data. Consequently, having a reliable and automated method to analyse patient narratives was viewed as helpful.

Although senior nurse managers had a substantial role in generating and assessing patient feedback, staff working on the front line often reported feeling disconnected from current mechanisms of collecting and reporting on patient feedback. For many front-line staff, the collection and processing of patient feedback was something carried out by managerial staff as an organisational requirement to ensure quality assurance for reporting to regulators. The generation and use of data needs to be perceived as meaningful for staff, and ways of enabling this might vary in different settings.

The comparison between different service settings has drawn attention to the importance of context and the differences as well as commonalities across settings. Staff in the different settings talked about a number of ways in which they might feel a greater level of investment in the collection and use of data. For example:

- In the rheumatology outpatients service, consultants talked about the value of designing specific questions related to their own service, rather than being limited to general issues that they had little control over within a large organisation (such as food and parking).
- In the community mental health team, team members talked about the value of having a mechanism for generating feedback that fits easily with their way of working, which entails home visits and extensive one-to-one discussions about service users' experiences.
- In primary care, GPs talked about their experiences of collecting more detailed feedback on aspects of the service that they had had specific challenges with (e.g. reaching a good level of attendance for the flu clinic or other clinics).

The qualitative research drew attention to the perceived need for flexibility for all settings. Patients, carers and staff often felt that there should be more opportunities to collect and use verbal feedback, which was frequently delivered using more informal mechanisms. This was considered important across all settings, but was viewed as particularly important in the mental health service context. Here, service users and staff talked more often about the challenges of enabling feedback when unwell, because many people find it harder to write or type feedback. There was a view that, at some points, people may prefer to give verbal feedback in the context of a trusted and valued therapeutic relationship, but that they also needed to be aware of the alternative possible mechanisms for providing feedback, including anonymised surveys and independent discussions. The requirement for flexibility, and combining multiple options for giving feedback, were also considered likely to enhance participation rates and the utility of feedback data.

Improving the processing and analysis of narrative data alongside quantitative data

Narrative feedback is valuable for providing greater detail and an understanding of context, but processing narrative data is a challenging task. The key challenges identified in this study include:

- Multi-themed comments. Patient feedback comments typically refer to several themes, often with contrasting sentiments (e.g. 'nice food, tired waiting area'). This requires the identification of text segments that refer to specific themes, rather than the processing of entire comments. Although frequent and well-defined themes can be identified accurately, rare and vague expressions still need the involvement of users.

- The complexity of patient language. Narratives are typically informal, with unstructured grammar and syntax. They can be short (two to three words) or extremely long (up to 18 sentences). Therefore, standard linguistic-based rules are difficult to engineer, suggesting that machine-learning approaches are a preferred approach for processing narrative data.
- Training data. Although we generated high-quality training data sets, their size and theme distributions were not sufficient to capture all of the variability in the narratives. Additional training data and alternative methods for semi-supervised and incremental training need to be explored further.

Overall, this study has demonstrated that automated processing of free-text comments is promising and can be integrated in a semi-automated toolkit that gives an overview of comments on a large scale, presents typical examples of comments (for particular themes) and provides an opportunity to drill down to specific or rare comments for further manual analysis.

Co-design of tools to improve the collection, analysis and presentation of patient experience data for staff to maximise the potential for stimulating service improvement

Drawing on the findings of the two study components described in the previous sections, we worked with patients, carers and staff to co-design a toolkit comprising the following:

- A survey utilising the FFT with space for free-text comments to be completed using a digital kiosk in study sites, online or using a pen and paper version.
It was clear that most patients, carers and staff thought that there was value in enabling participation through the digital collection of feedback in the multiple service contexts. Although many recognised the limitations of the FFT, they were keen to keep the process as simple as possible and use of the FFT question was viewed to be a simple approach, with emphasis placed on the value of providing free-text comments to complement responses to the question. Staff in all sites were in favour of using digital data collection to help them in fulfilling their obligations to collect FFT data. The positioning of pen and paper versions and a link to the online questionnaire responded to concerns about ensuring that a range of options for giving feedback were available.
- Guidance and information for staff, patients and carers to support use of the new tools.
The development of guidance and information for staff, patients and carers was considered essential, given the finding that many patients and staff thought that there was insufficient awareness about the different ways to give feedback. Flexibility was also considered important to ensure quality and safety. The kiosk installations were also used as an opportunity to showcase previous feedback and its use in informing service delivery (using co-designed posters and flyers).
- New text-mining programs for analysing patient feedback data.
These programmes were developed during the initial phases of the study and used to analyse archived narrative data provided alongside the FFT question. However, they were not tested further during the testing period because of the small volumes of data collected from the kiosks in each of the four sites. Instead, the findings of analyses of archived data were used in presentations to trigger discussions about the value of this kind of analysis and as a means of informing development of appropriate reporting templates.
- New templates for reporting feedback from multiple sources.
Following initial discussions on possible formats for reporting narrative data alongside quantitative data, a number of reporting templates were developed. An automated template was created to enable outputs from the text-mining process to be summarised and as a means of prompting discussions about the most appropriate ways of displaying data, for example in the form of graphs, as well as boxes to display examples of free-text comments provided and how they had been categorised according to themes.

We also developed a number of alternative templates with different combinations of graphs and figures alongside text. Following initial feedback in co-design focus groups, these templates were adapted to form a single template to be used for monthly reports in each site during the testing period. This template included a headline summary of data in terms of the completion rates for the month, alongside summarised categories of responses to the FFT and examples of free-text comments summarised into key themes manually by members of the research team.

- A new process for eliciting and recording verbal feedback in community mental health services. Given the particular emphasis placed on the value of informal and verbal feedback by patients and staff in mental health services, a process for enabling and recording such feedback in routine practice was developed for testing during the evaluation period. This was perceived to allow meaningful data to be collected in the 'normal' way in which this was viewed to already take place. Although community mental health team staff reported having such interactions regularly and informally, the testing period was considered to be an opportunity to document this process more formally.

Implementation and process evaluation of new tools to enhance the collection, presentation and use of patient feedback data

This study found that digital tools can help to enhance the capture, analysis and perceived usefulness of feedback data. At the start of the study, no sites were routinely collecting patient experience data digitally. Once the sites began the digital capture of patient feedback using the kiosks, the increased rates of participation increased the volume of feedback captured at monthly intervals across three sites compared with the period before introduction of the new tools. However, one of the sites (site B, mental health trust) was slow to engage and the rates of participation declined throughout the study period.

We demonstrated proof of concept for routine digital data capture in a range of settings and showed some improvements in the cycle of data collection, analysis and use. However, such tools require additional investment of time and support and there were multiple barriers to adoption and little evidence of impact in the short period of data collection. The most successful adoption was in one of the primary care sites, where peer support proved a successful way of supporting participation. The findings also highlight the need to consider alternative ways of capturing feedback. The new tools were not found to influence changes in service delivery during the evaluation period; the data contained in monthly reports were limited and, although staff found the data informative, they did not find them sufficient to give clear indications of what (if any) changes were needed.

The comparison of the qualitative analysis of text with the text-mining approach demonstrated some of the strengths and limitations of each of these approaches. The qualitative analysis was able to extract more detailed themes and insights into issues that are particularly important in specific service contexts (e.g. access and timing of discharge for mental health services). However, this approach to analysis is feasible only for smaller data sets because it is time-consuming. In contrast, the text-mining method enables faster processing of large data sets but the transformation of data into a small number of generic categories may miss important information relating to more unusual experiences and potential 'patient safety' concerns.

Qualitative findings from the evaluation period were aligned to four key themes informed by core constructs of NPT:

1. Coherence: perceived value of digital tools for data collection and analysis. The tools made sense to staff based on perceived deficits in previous systems; however, there was a perceived need for additional resources to support use of the tools and a need for greater flexibility to enable relevant feedback to be collected and used.

2. Cognitive participation: information and support needs of patients and carers. Staff engagement with the new tools varied and observation in the centres indicated that patients were apprehensive about using the kiosks spontaneously, but would often participate with support. Peer support from a PPG in one site demonstrated the potential value of this. Staff turnover presented a challenge for supporting use of the new tools
3. Collective action: organisational and technical work for sustaining the new tools. The workload of staff was often highlighted as a barrier to implementation of the kiosks. Technical problems encountered included the machines becoming frozen and a lack of capacity and responsibility for maintenance of the kiosks.
4. Reflexive monitoring: embedding the new intervention. Although staff were positive about the feedback reports generated during the study and found these to be helpful for discussing any issues raised, changes to service delivery were not observed during the evaluation period.

Implications for health services

This study adds to a growing body of literature on the implementation of patient feedback and a number of key implications for health services were identified:

- There is a need for increased participation in the provision of patient feedback and raised awareness of why feedback is collected and what data can and have been used for.
- There is a need for support to enable digital participation in diverse NHS settings.
- There is a need to capture more meaningful data and emphasise the value of informal interactions and positive feedback.
- The value and current limitations of a text-mining approach should be emphasised, as well as the value of manual analysis of text for small bodies of data and in specific settings.
- It is important how data are presented.
- Context and flexibility are important.
- There is a need for greater buy-in from staff and senior leaders to foster enthusiasm for generating useful data and for acting on the findings.
- The costs and resources required to support improvements in data capture, analysis and use are potentially high. The extent of the costs depends on the costs of developing the methods and tools for data capture and analysis in any roll-out beyond the four sites in this study. The costs also depend on the extent to which there are economies of scale in the shared costs of both development and implementation. If single sites have to bear all of the shared development costs then the total cost of the toolkit per year would range between approximately £19,000, if the toolkit had a lifespan of 5 years, and £122,000, if the toolkit had a lifespan of 1 year only. The value of any additional costs associated with the toolkit depend on the extent to which it adds benefit to the organisation, staff and patients in terms of improved care processes, satisfaction, well-being and health.
- There are challenges in enabling feedback to inform service improvements.

Implications for future research

Further research is needed:

- on different ways of recording informal and positive feedback
- on the divisions and overlaps between feedback and complaints
- to develop and test approaches to peer support in service settings
- on the methodological aspects of text mining
- to develop and test more qualitative approaches initiated and tested by staff teams to gather and use meaningful feedback.

Acknowledgements

The co-authors would like to thank the following people and bodies.

The participating trusts and general practices for taking part in the study.

All of the patients and staff in the NHS participating sites who took part and supported the study.

All of the members of our PPI project advisory groups: Annmarie Lewis, Neal Sinclair, Dawn Allen, Kate Lurie, Helen Yeoman, Jane Reid-Peters, Annette Barber and Susan Moore.

Members of our study steering committee for their generous contributions in discussing and advising on many aspects of the project: Professor James Barlow (Chairperson), Professor Martin Knapp, Konstantina Poursanidou, Professor Nigel Collier, Simon Stones, Professor Carolyn Chew-Graham and Professor Tim Doran.

Sue Pargeter from the NIHR National Evaluation, Trials and Studies Coordinating Centre (NETSCC) for helpful advice throughout the project.

Cathy Lovatt, Head of Service User and Carer Involvement, Greater Manchester Mental Health NHS Foundation Trust, for advising on and supporting the project, from the co-design and evaluation phase to the dissemination of key findings.

Dedication

In memory of Neal Sinclair and Jane Reid Peters, who died during the study period. We are grateful to both Neal and Jane, who brought their energy, enthusiasm, expertise and experience to their advisory roles in the project. They leave a lasting influence on our work.

Contributions of authors

Caroline Sanders (<https://orcid.org/0000-0002-0539-928X>) (Professor in Medical Sociology; Chief Investigator) led the overall design of the study and data collection, contributed to the analysis and led the writing of the final report.

Papreen Nahar (<https://orcid.org/0000-0002-5817-8093>) (Research Fellow, Anthropology) contributed to the qualitative research, data collection and analysis, and writing and approval of the final report.

Nicola Small (<https://orcid.org/0000-0002-7879-7967>) (Research Associate, Health Services Research) contributed to the qualitative research, data collection and analysis, and writing and approval of the final report.

Damian Hodgson (<https://orcid.org/0000-0002-9292-5945>) (Professor of Organisational Analysis) contributed to the study design, was the WS4 lead and contributed to the qualitative research, data collection and analysis, and writing and approval of the final report.

Bie Nie Ong (<https://orcid.org/0000-0001-8138-8139>) (Emerita Professor of Health Services Research) contributed to the qualitative analysis and writing and approval of the final report.

Azad Dehghan (<https://orcid.org/0000-0001-7000-2835>) (Postdoctoral Research Associate, Computer Science) contributed to the computer science research, analysed the data to develop and run the text-mining programs and contributed to the analysis and writing and approval of the final report.

Charlotte A Sharp (<https://orcid.org/0000-0003-4051-2281>) (Clinical Academic Fellow) had a clinical advisory role for the rheumatology aspects of the study and contributed to the qualitative research, data collection and analysis, and writing and approval of the final report.

William G Dixon (<https://orcid.org/0000-0001-5881-4857>) (Professor of Digital Epidemiology, Clinician) contributed to the study design, was the clinical lead for the rheumatology aspects of the study, co-ordinated research activities for the acute trust and contributed to data collection and analysis and writing and approval of the final report.

Shôn Lewis (<https://orcid.org/0000-0003-1861-4652>) (Professor of Psychiatry) contributed to the study design, was the clinical lead for the psychiatry aspects of the study, provided mentorship in project management to the Chief Investigator and contributed to data collection and analysis and writing and approval of the final report.

Evangelos Kontopantelis (<https://orcid.org/0000-0001-6450-5815>) (Professor in Data Science and Health Services Research) was the lead for the statistical aspects of the study, created the code and automated reports for the text-analytics work and contributed to writing and approval of the final report.

Gavin Daker-White (<https://orcid.org/0000-0002-3538-8805>) (Research Fellow) led the comparison of the qualitative analysis with text mining and contributed to analysis of the primary qualitative data and writing and approval of the final report.

Peter Bower (<https://orcid.org/0000-0001-9558-3349>) (Professor of Health Services Research) contributed to the study design, was the primary care lead and contributed to the analysis and writing and approval of the final report.

Linda Davies (<https://orcid.org/0000-0001-8801-3559>) (Professor of Health Economics Research) contributed to the study design, was the lead for the costing analysis and contributed to writing and approval of the final report.

Humayun Kayesh (<https://orcid.org/0000-0002-9975-5862>) (Research Associate) contributed to the computer science research, analysed the data for the text-mining component of the study and contributed to writing and approval of the final report.

Rebecca Spencer (Programme Manager) co-ordinated and managed the project during year 1, provided project management oversight during year 2, led co-ordination of the components for the toolkit, led the writing of multiple progress reports with the Chief Investigator and contributed to writing and approval of the final report.

Aneela McAvoy (Project Manager) co-ordinated and managed the project during year 1, had a lead role in the PPI workshop, co-ordinated study activities and contributed to writing and approval of the final report.

Ruth Boaden (<https://orcid.org/0000-0003-1927-6405>) [Director, NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC) Greater Manchester, and Professor of Service Operations] contributed to the study design, the analysis and writing and approval of the final report.

Karina Lovell (<https://orcid.org/0000-0001-8821-895X>) (Director of Research and Professor of Mental Health) contributed to the study design, advised on research activities and interpretation of findings, advised on PPI aspects and contributed to writing and approval of the final report.

John Ainsworth (<https://orcid.org/0000-0002-2187-9195>) (Professor of Health Informatics) contributed to the study design, was the lead for the informatics aspects of the study, advised on the digital tools and contributed to interpretation of the findings and writing and approval of the final report.

Magdalena Nowakowska (<https://orcid.org/0000-0003-1386-2534>) (Research Assistant) was involved in the statistical aspects of the study, contributed to the analysis for autoreporting of the text-mining work, led the analysis of participation rates in WS4 and contributed to writing and approval of the final report.

Andrew Shepherd (<https://orcid.org/0000-0001-6589-746X>) (NIHR Clinical Lecturer) contributed to the comparison of the qualitative analysis with text mining, analysis of the primary qualitative data and writing and approval of the final report.

Patrick Cahoon (<https://orcid.org/0000-0002-3594-6811>) (Head of Patient Experience) contributed to the study design, played a lead role in enabling research activities in the mental health trust, contributed to discussions regarding the data collected and to interpretation of the findings, advised on clinical issues and quality aspects for mental health services and contributed to writing and approval of the final report.

Richard Hopkins (Honorary Senior Lecturer) contributed to the study design, advised on research activities and interpretation of the findings, advised on clinical issues and quality aspects for mental health services and contributed to writing and approval of the final report.

Dawn Allen (PPI Member) was a member of our PPI group and played a lead role in PPI activities, contributed to discussions regarding research findings and interpretation, carried out presentations on behalf of the study team, worked with us on the design of information materials and carer recruitment and contributed to writing and approval of the final report.

Annmarie Lewis (PPI Member, PPI Co-investigator) was a member of our PPI group and played a lead role in PPI activities, contributed to discussions regarding research findings and interpretation, carried out presentations on behalf of the study team, worked with us on the design of information materials and carer recruitment and contributed to writing and approval of the final report.

Goran Nenadic (<https://orcid.org/0000-0003-0795-5363>) (Professor of Computer Science) contributed to the research design, was the WS2 lead and contributed to the analysis, creation of the text-mining programs and writing and approval of the final report.

Publications

Ong BN, Sanders C. Exploring engagement with digital screens for collecting patient feedback in clinical waiting rooms: the role of touch and place [published online ahead of print December 9, 2019]. *Health* 2019. <https://doi.org/10.1177%2F1363459319889097>

Ong BN, Hodgson D, Small N, Nahar P, Sanders C. Implementing a digital patient feedback system: an analysis using Normalisation Process Theory. *BMC Health Serv Res* 2020;**20**:387.

Further dissemination information is provided in *Appendix 4*.

Data-sharing statement

All qualitative data generated that can be shared are contained within the report. All data queries and requests should be submitted to the corresponding author for consideration.

Patient data

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

References

1. Health Information and Quality Authority. *International Review on the Use of Patient Experience Surveys in the Acute Sector*. Cork: Health Information and Quality Authority; 2016.
URL: www.hiqa.ie/sites/default/files/2017-02/Intl-Review-Model-Methodology-to-implement-NPE-Survey.pdf (accessed 18 February 2020).
2. Anhang Price R, Elliott MN, Zaslavsky AM, Hays RD, Lehrman WG, Rybowski L, et al. Examining the role of patient experience surveys in measuring health care quality. *Med Care Res Rev* 2014;**71**:522–54. <https://doi.org/10.1177/1077558714541480>
3. Department of Health and Social Care (DHSC). *NHS Patient Experience Framework*. London: DHSC; 2012.
4. Merkley K, Bickmore AM. *The Top Five Recommendations for Improving the Patient Experience*. Salt Lake City, UT: Health Catalyst; 2017. URL: <https://downloads.healthcatalyst.com/wp-content/uploads/2016/09/Top-Five-Recommendations-for-Improving-Patient-Experience.pdf> (accessed 18 February 2020).
5. Care Quality Commission. *NHS Patient Surveys*. URL: <https://nhssurveys.org> (accessed 18 February 2020).
6. Francis R. *Report of the Mid-Staffordshire NHS Foundation Trust Public Inquiry*. London: The Stationery Office; 2013.
7. National Advisory Group on the Safety of Patients in England. *A Promise to Learn – A Commitment to Act*. London: Williams Lea; 2013.
8. NHS England. *NHS England Review of the Friends and Family Test*. 2014. URL: www.england.nhs.uk/wp-content/uploads/2014/07/fft-rev1.pdf (accessed 15 November 2019).
9. Manacorda T, Erens B, Black N, Mays N. *Implementation and Use of the Friends and Family Test as a Tool for Local Service Improvement in NHS General Practice in England*. London: PIRU Policy Innovation Research Unit; 2016.
10. Robert G, Cornwell J, Black N. Friends and family test should no longer be mandatory. *BMJ* 2018;**360**:k367. <https://doi.org/10.1136/bmj.k367>
11. Coulter A, Locock L, Ziebland S, Calabrese J. Collecting data on patient experience is not enough: they must be used to improve care. *BMJ* 2014;**348**:g2225. <https://doi.org/10.1136/bmj.g2225>
12. Gleeson H, Calderon A, Swami V, Deighton J, Wolpert M, Edbrooke-Childs J. Systematic review of approaches to using patient experience data for quality improvement in healthcare settings. *BMJ Open* 2016;**6**:e011907. <https://doi.org/10.1136/bmjopen-2016-011907>
13. Bourne T, Wynants L, Peters M, Van Audenhove C, Timmerman D, Van Calster B, Jalmbrant M. The impact of complaints procedures on the welfare, health and clinical practise of 7926 doctors in the UK: a cross-sectional survey. *BMJ Open* 2015;**5**:e006687. <https://doi.org/10.1136/bmjopen-2014-006687>
14. Smither JW, Walker AG. Are the characteristics of narrative comments related to improvement in multirater feedback ratings over time? *J Appl Psychol* 2004;**89**:575–81. <https://doi.org/10.1037/0021-9010.89.3.575>
15. Civica and InHealth Associates. *Making Sense and Making Use of Patient Experience Data*. London: Civica and InHealth Associates; 2015. URL: www.civica.com/globalassets/7.document-downloads/2.uk-docs/white-papers/engagement-solutions/making-sense-and-making-use-of-patient-experience-2015.pdf (accessed 18 February 2020).

16. Raleigh V, Thompson J, Jabbal J, Graham C, Sizmur S, Coulter A. *Patients' Experience of Using Hospital Services: Lessons From an Analysis of Trends in 2005–2013*. London: The King's Fund and Picker Institute Europe; 2016. URL: www.kingsfund.org.uk/sites/default/files/field/field_publication_file/Patients-experience-Kings-Fund-Dec-2015.pdf (accessed 18 February 2020).
17. Sheard L, Marsh C, O'Hara J, Armitage G, Wright J, Lawton R. The Patient Feedback Response Framework – understanding why UK hospital staff find it difficult to make improvements based on patient feedback: a qualitative study. *Soc Sci Med* 2017;**178**:19–27. <https://doi.org/10.1016/j.socscimed.2017.02.005>
18. Staniszewska S, Churchill N. Patients' experiences in the UK: future strategic directions. *Patient Exp J* 2014;**1**:140–3. <https://doi.org/10.35680/2372-0247.1017>
19. National Quality Board. *Improving Experiences of Care: Our Shared Understanding and Ambition*. National Quality Board; 2015. URL: <https://webarchive.nationalarchives.gov.uk/20161103234108/https://www.england.nhs.uk/wp-content/uploads/2015/01/improving-experiences-of-care.pdf> (accessed 18 February 2020).
20. Wolf JA, Niederhauser, V, Marshburn D, LaVela SL. Defining patient experience. *Patient Exp J* 2014;**1**:7–19.
21. Ziewitz M. Experience in action: moderating care in web-based patient feedback. *Soc Sci Med* 2017;**175**:99–108. <https://doi.org/10.1016/j.socscimed.2016.12.028>
22. Paterson BL. The shifting perspectives model of chronic illness. *J Nurs Scholarsh* 2001;**33**:21–6. <https://doi.org/10.1111/j.1547-5069.2001.00021.x>
23. Porter T, Sanders T, Richardson J, Grime J, Ong BN. Living with multimorbidity: clinical and patient perspectives. *Int J Clin Rheumatol* 2015;**10**:111–19. <https://doi.org/10.2217/ijr.15.6>
24. Mazanderani F, Locock L, Powell J. Biographical value: towards a conceptualisation of the commodification of illness narratives in contemporary healthcare. *Sociol Health Illn* 2013;**35**:891–905. <https://doi.org/10.1111/1467-9566.12001>
25. Lupton D. The commodification of patient opinion: the digital patient experience economy in the age of big data. *Sociol Health Illn* 2014;**36**:856–69. <https://doi.org/10.1111/1467-9566.12109>
26. Vogus TJ, McClelland LE. When the customer is the patient: lessons from healthcare research on patient satisfaction and service quality ratings. *Hum Res Manage Rev* 2016;**26**:37–49. <https://doi.org/10.1016/j.hrmr.2015.09.005>
27. NHS Institute for Innovation and Improvement. *Patient Feedback Survey 2012: National and Strategic Health Authority Summary Report*. London: NHS Institute for Innovation and Improvement and Ipsos Mori; 2012. URL: www.ipsos.com/sites/default/files/publication/1970-01/sri_Patient_Feedback_Survey_20122.pdf (accessed 18 February 2020).
28. Insight Team NHS England. *NHS England Review of the Friends and Family Test*. London: NHS England; 2014. URL: www.england.nhs.uk/wp-content/uploads/2014/07/fft-rev1.pdf (accessed 18 February 2020).
29. Picker Institute Europe. *NHS Friends and Family Test 'Unreliable' Comparison Tool Says Picker*. 2014. URL: www.picker.org/news/nhs-friends-family-test-unreliable-comparison-tool-says-picker-institute-europe/ (accessed 20 December 2018).
30. NHS Institute for Innovation and Improvement. *The Patient Experience Book*. London: NHS Institute for Innovation and Improvement; 2013.
31. Brookes G, Baker P. What does patient feedback reveal about the NHS? A mixed methods study of comments posted to the NHS Choices online service. *BMJ Open* 2017;**7**:e013821. <https://doi.org/10.1136/bmjopen-2016-013821>

32. Griffiths A, Leaver MP. Wisdom of patients: predicting the quality of care using aggregated patient feedback. *BMJ Qual Saf* 2018;**27**:110–18. <https://doi.org/10.1136/bmjqs-2017-006847>
33. King's College London and The King's Fund. *What Matters to Patients*. London: King's College London and The King's Fund; 2011.
34. Overeem K, Lombarts MJ, Arah OA, Klazinga NS, Grol RP, Wollersheim HC. Three methods of multi-source feedback compared: a plea for narrative comments and coworkers' perspectives. *Med Teach* 2010;**32**:141–7. <https://doi.org/10.3109/01421590903144128>
35. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. *J Med Internet Res* 2013;**15**:e239. <https://doi.org/10.2196/jmir.2721>
36. Cole-Lewis H, Varghese A, Sanders A, Schwarz M, Pugatch J, Augustson E. Assessing electronic cigarette-related tweets for sentiment and content using supervised machine learning. *J Med Internet Res* 2015;**17**:e208. <https://doi.org/10.2196/jmir.4392>
37. Gibbons C, Richards S, Valderas JM, Campbell J. Supervised machine learning algorithms can classify open-text feedback of doctor performance with human-level accuracy. *J Med Internet Res* 2017;**19**:e65. <https://doi.org/10.2196/jmir.6533>
38. Padmavathy P, Leema AA. Sentiment mining from online patient experience using latent Dirichlet allocation method. *Indian J Sci Technol* 2016;**9**(19). <https://doi.org/10.17485/ijst/2016/v9i19/93876>
39. Tapi Nzali MD, Bringay S, Lavergne C, Mollevi C, Opitz T. What patients can tell us: topic analysis for social media on breast cancer. *JMIR Med Inform* 2017;**5**:e23. <https://doi.org/10.2196/medinform.7779>
40. Wagland RR-S, Simon A, Bracher M, Hunt M, Foster K, Downing C, *et al*. Development and testing of a text-mining approach to analyse patients' comments on their experiences of colorectal cancer care. *BMJ Qual Saf* 2016;**25**:604–14. <https://doi.org/10.1136/bmjqs-2015-004063>
41. Brody S, Elhadad N. *An Unsupervised Aspect-sentiment Model for Online Reviews*. Human Language Technologies: the 2010 Annual Conference of the North American Chapter of the ACL, 2–4 June 2010, Los Angeles, CA, USA, pp. 804–12.
42. Choi Y, Cardie C. *Hierarchical Sequential Learning for Extracting Opinions and their Attributes*. Proceedings of the ACL 2010 Conference Short Papers, 11–16 July 2010, Uppsala, Sweden, pp. 269–74.
43. Yu D, Wang S, Deng L. Sequential labeling using deep-structured conditional random fields. *IEEE J Sel Top Signal Process* 2010;**4**:965–73. <https://doi.org/10.1109/JSTSP.2010.2075990>
44. Yu D, Wang S, Deng L. *Sequential Labeling Using Deep-structured Conditional Random Fields*. Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 3–8 June 2012, Montreal, QC, Canada.
45. Hai Z, Cong G, Chang K, Cheng P, Miao C. Analyzing sentiments in one go: a supervised joint topic modeling approach. *IEEE Trans Knowl Data Eng* 2017;**29**:1172–85. <https://doi.org/10.1109/TKDE.2017.2669027>
46. Care Quality Commission, NHS Patient Surveys. *NHS Surveys: Focused on Patients' Experience*. 2017. URL: <https://nhssurveys.org> (accessed 25 February 2020).
47. Asprey A, Campbell JL, Newbould J, Cohn S, Carter M, Davey A, Roland M. Challenges to the credibility of patient feedback in primary healthcare settings: a qualitative study. *Br J Gen Pract* 2013;**63**:e200–8. <https://doi.org/10.3399/bjgp13X664252>

48. Robert G, Cornwell J, Brearley S, Foot C, Goodrich J, Joule N, et al. *What Matters to Patients? Developing the Evidence Base for Measuring and Improving Patient Experience*. London: The King's Fund; 2011.
49. Spasić I, Livsey J, Keane JA, Nenadić G. Text mining of cancer-related information: review of current status and future directions. *Int J Med Inform* 2014;**83**:605–23. <https://doi.org/10.1016/j.ijmedinf.2014.06.009>
50. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;**350**:h1258. <https://doi.org/10.1136/bmj.h1258>
51. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociology* 2009;**43**:535–54. <https://doi.org/10.1177/0038038509103208>
52. Chang CK, Hayes RD, Broadbent M, Fernandes AC, Lee W, Hotopf M, Stewart R. All-cause mortality among people with serious mental illness (SMI), substance use disorders, and depressive disorders in southeast London: a cohort study. *BMC Psychiatry* 2010;**10**:77. <https://doi.org/10.1186/1471-244X-10-77>
53. NHS England. *Musculoskeletal Conditions: Why Is It Important?* URL: www.england.nhs.uk/ourwork/clinical-policy/ltc/our-work-on-long-term-conditions/musculoskeletal/ (accessed 20 December 2018).
54. Rhodes P, Sanders C, Campbell S. Relationship continuity: when and why do primary care patients think it is safer? *Br J Gen Pract* 2014;**64**:e758–64. <https://doi.org/10.3399/bjgp14X682825>
55. While D, Bickley H, Roscoe A, Windfuhr K, Rahman S, Shaw J, et al. Implementation of mental health service recommendations in England and Wales and suicide rates, 1997–2006: a cross-sectional and before-and-after observational study. *Lancet* 2012;**379**:1005–12. [https://doi.org/10.1016/S0140-6736\(11\)61712-1](https://doi.org/10.1016/S0140-6736(11)61712-1)
56. Bate P, Robert G. *Bringing User Experience to Healthcare Improvement: The Concepts, Methods and Practices of Experience-based Design*. Oxford: Radcliffe Publishing; 2007.
57. Strauss AL, Corbin JM. *Basics of Qualitative Research Techniques*. London: SAGE Publications Ltd; 1998.
58. Charmaz K. *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. London: SAGE Publications Ltd; 2006.
59. Corbin J, Strauss A. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. London: SAGE Publications Ltd; 2008. <https://doi.org/10.4135/9781452230153>
60. Ong BN, Hodgson D, Small N, Nahar P, Sanders C. Implementing a digital patient feedback system: an analysis using Normalisation Process Theory. *BMC Health Serv Res* 2020;**200**:387.
61. NHS England. *FAQs for the Friends and Family Test – Updated 24/02/2017*. 2017. URL: www.england.nhs.uk/wp-content/uploads/2015/10/fft-imp-guid-faqs-oct15.pdf (accessed 15 November 2019).
62. Daker-White G, Dehghan A, Shepherd A, Nowokowska M, Kontopantelis E, Nenadic G, Sanders C. *Humans Versus Machines: Text Mining Versus Adapted Grounded Theory in the Analysis of Free Text Data from Patient Feedback Surveys*. Health Services Research UK (HSRUK) Conference, Nottingham, UK, 4–5 July 2018.
63. Dehghan A, Kayesh H, Daker-White G, Sanders C, Nenadic G. *Mining Free-text Patient Feedback Comments*. Heal-TAC, Manchester, UK, 18–19 April 2018.
64. Nahar P, Ong BN, Small N, Hodgson D, Nenadic G, Bower P, et al. *Implementing and Evaluating Tools to Improve the Collection and Usefulness of Patient Experience Data in Multiple Service Contexts*. Health Services Research UK (HSRUK) Conference, Nottingham, UK, 4–5 July 2018.

65. Nahar P, Sanders C, Small N, Hodgson D, Daker-White G, Spencer R, Sanders C. *Creating Meaningful Patient Feedback Data for Health Service Improvement: Exploring the Formality and Informality of Feedback Mechanisms*. British Sociological Association (BSA) 49th Medical Sociology Annual Conference, York, UK, 13–15 September 2017.
66. Sanders C, Nahar P, Hodgson D, Small N, Daker-White G, Sharp C, Ong BN. *Constructing and De-constructing Patient Experience via Big Data and Small Data*. World Congress of Sociology, Toronto, ON, Canada, 15–21 July 2018.
67. Sanders C. *Story Specimens and Chemistry: a Creative Enquiry*. Society for Academic Primary Care (SAPC) 47th Annual Scientific Meeting, London, UK, 10–12 July 2018.
68. Sanders C. *The Ethics and Politics of Sharing Stories for Patient and Public Involvement*. British Sociological Association (BSA) 50th Medical Sociology Annual Conference, Glasgow, UK, 12–14 September 2018.
69. Small N, Nahar P, Daker-White G, Spencer R, Hodgson D, Bower P, Sanders C. *Can Digital Data Capture and Improved Analysis of Comments Make Patient Feedback More Meaningful and Useful for Primary Care?* Society for Academic Primary Care (SAPC) Annual Conference, Warwick, UK, 12–14 July 2017.
70. Robert G, Locock L, Sanders C, Sheard L. *Enhancing the Use of Patient Experience Data for Improving the Safety and Quality of Care*. ISQua 34th Annual Conference, London, UK, 1–4 October 2017.
71. Roberts G, Sheard L, Locock L, Sanders C. *Exploring and Enhancing the Use of Patient Experience Data for Improving the Quality of Care*. Health Services Research UK (HSRUK) Symposium, Nottingham, UK, 6–7 July 2017.
72. Sanders C. *The DEPEND Project Overview*. Paper presented in Manchester, 28 March 2018.
73. Allen D, Lewis AM, Small N. *What PPI has Worked Well in the DEPEND Project?* Paper presented in Manchester, 28 March 2018.
74. Nahar P. *PPI in the Bangladeshi Community*. Paper presented in Manchester, 28 March 2018.
75. NIHR Greater Manchester Patient Safety Translational Research Centre. *DEPEND: Use of Digital Methods for Collection and Use of Patient Experience Data*. 2018. URL: www.youtube.com/watch?v=BOYLxVJAzdl (accessed 23 December 2019).
76. Small N, Lewis AM, Allen D, Ong BN, Sanders C. *Co-designing New Tools for Collecting, Analysing and Presenting Patient Feedback in NHS Service: Working in Partnership with Patients and Carers*. International Perspectives on Evaluation of PPI in Research, Newcastle University, Newcastle upon Tyne, UK, 15–16 November 2018.

Appendix 1 Tables containing additional information for the text-mining methods

TABLE 19 Topics used for the initial manual coding

Category	Description
Waiting time	Comments on waiting times and service delays
Staff attitude (and professionalism)	Comments on staff attitude, behaviour and professionalism. Often describes how hospital staff are perceived by service users
Care quality	Comments on care quality as perceived by service users. Any mention of gratefulness/thankfulness vis-a-vis the service provider is included here
Food	Comments on catering services and food
Process	Comments on processes, e.g. discharge processes, hospital hand-off, referrals, admissions and similar processes. May describe communication failures between hospital staff (note the difference from communication)
Environment	Comments on cleanliness, noise level, temperature, directions
Parking	Comments on parking-related aspects
Communication	Comments on communication and inadequate dissemination of information between service providers and service users
Resource	Comments on resource availability or lack thereof
Not feedback	Comments that include any text that may have been included by error or that do not represent feedback (e.g. 'no', 'unlikely', 'extremely likely', 'likely', 'stop')
Other	Comments on any other topic that does not fall within the previously defined categories

TABLE 20 Merged themes

Merged theme	Includes
Staff attitude	Staff attitude, communication
Care quality	Care quality, resource
Waiting time	Waiting time
Environment	Environment, food, parking
Other	Not feedback, other

TABLE 21 Theme distribution in the 'gold standard' data set

Theme	Site A		Site B	
	Examples	Percentage	Examples	Percentage
Staff attitude	224	32.75	218	21.71
Care quality	156	22.81	520	51.79
Waiting time	103	15.06	98	9.76
Environment	45	6.58	60	5.98
Other	156	22.81	108	10.76
Total	684	100	1004	100.00

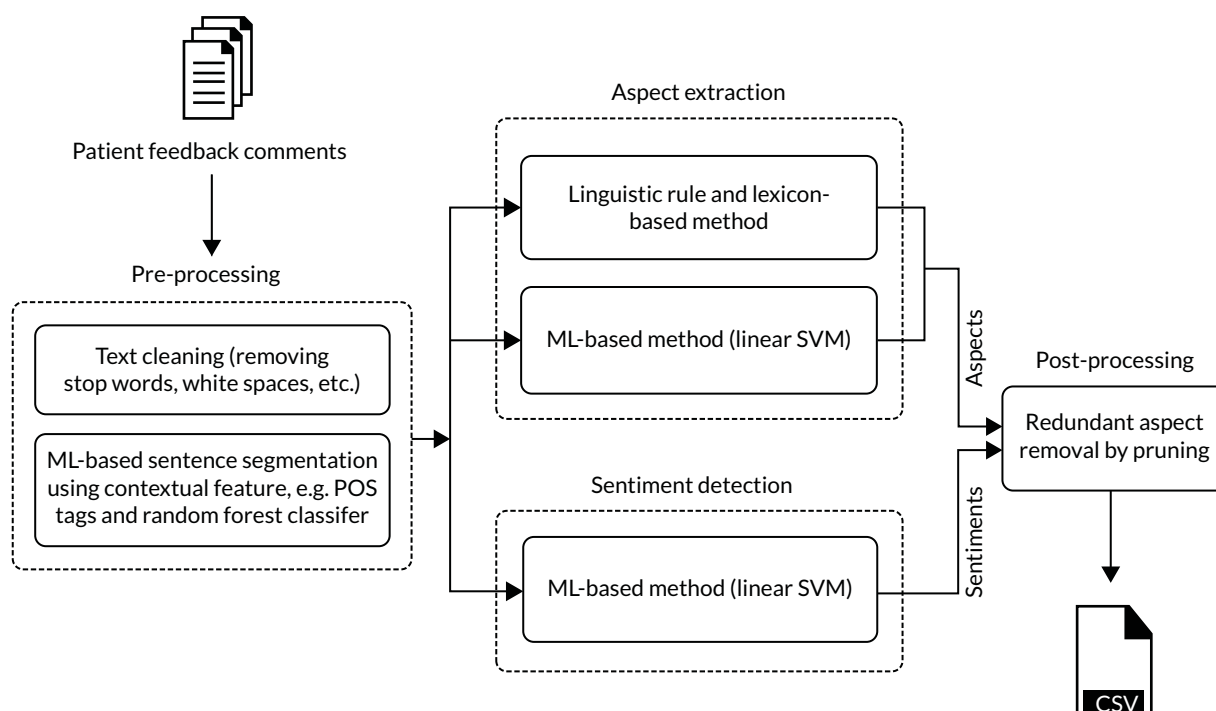


FIGURE 11 The SBM workflow. ML, machine learning; POS, part of speech; SVM, support vector machine.

BOX 2 Technical components and stages for the SBM

Segmentation: the aim was to predict whether or not a given candidate word indicates the beginning/end of a segment in the given context. A random forest classifier was used to identify segmentation points, with the feature set including lexical [e.g. part-of-speech (POS) tags for conjunctions] and contextual features that were extracted from the position of each segmentation marker (as provided in the training set by the annotators). The contextual features include lemmatised words, POS tags, distance of nearest conjunction (bidirectional) and dependency relations between each pair of the context words. A five-word window was used to extract context features around a candidate segmentation point.

Theme classifier: a theme classifier was trained using a multiclass support vector machine (SVM) with the linear kernel trained on uni- and bi-gram features. The features were weighted using term frequency over inverse document frequency (tf-idf). The data were pre-processed using text tokenisation, conversion to lower case, tokens/words lemmatisation, and white space and stop words removal.

Sentiment classifier was trained at the segment level using a binary SVM (positive and negative/neutral), with the same features and pre-processing as for the theme classifier.

Post-processing: in the final step, repetitive themes assigned to neighbouring segments were merged; the segments having the same themes and sentiments were treated as one pair of theme and sentiment.

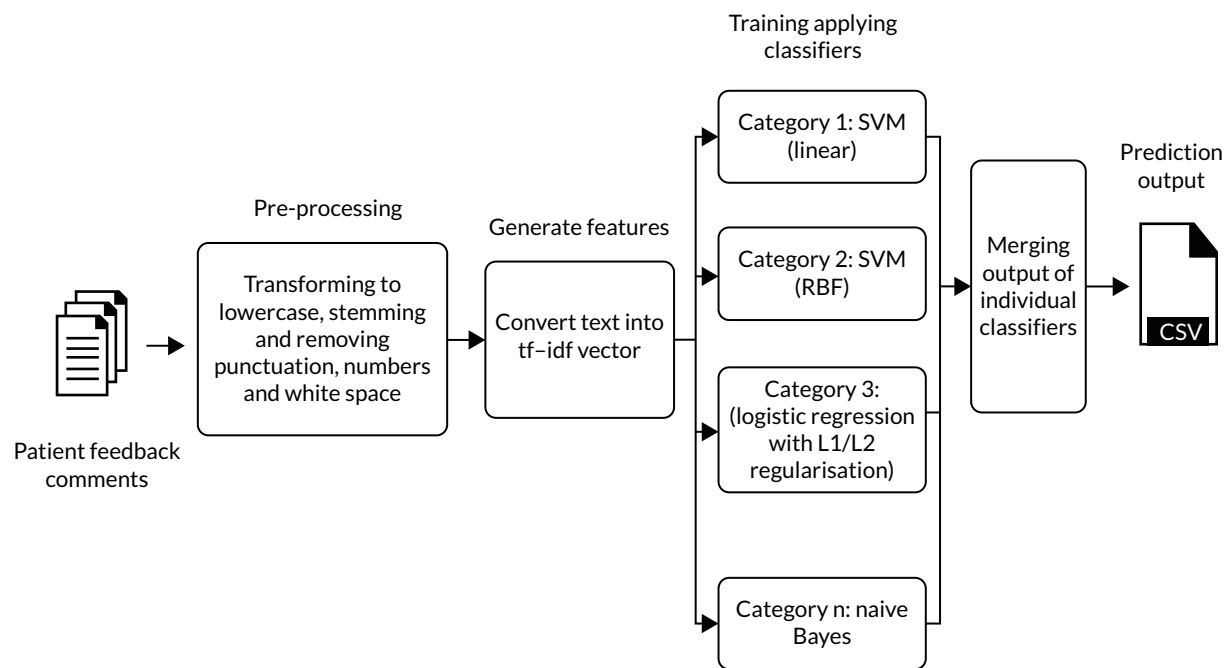


FIGURE 12 The CLM workflow. RBF, radial basis function; SVM, support vector machine; tf-idf, term frequency over inverse document frequency.

BOX 3 Technical components and stages for the CLM

Theme classifier: we experimented with several classifiers using uni-, bi- and tri-grams, with the features weighted using term frequency over inverse document frequency (tf-idf). We experimented with four different classifiers – (1) support vector machine (SVM) with linear kernel, (2) SVM with radial basis function kernel, (3) elastic net or logistic regression with L1 and L2 regularisation and (4) binary naive Bayes with Laplace smoothing – and aimed to identify those best performing for each theme (using cross-validation on the ‘gold standard’ data set).

Appendix 2 The DEPEND study staff participant information sheet and consent form



Developing and Enhancing the Usefulness of Patient Experience and Narrative Data

Chief Investigator: Dr Caroline Sanders, University of Manchester

Staff Participant Information Sheet

We would like to invite you to take part in a research study. Before you decide whether or not to take part, please read this information sheet which explains what this research is about, why it is being done and how it could involve you.

What is the study about?

Information about patients' experience of NHS services is routinely collected with the aim of informing ways of improving the quality, safety and effectiveness of services, and ensuring that they are sensitive to population needs. However, there are problems with patient experience data, and the way in which they are processed and summarised. Data often come from postal surveys that ask patients questions about events that took place sometime ago. Not only do these surveys receive low response rates, but healthcare professionals are often sceptical about the relevance of information due to the generalised nature of the questions that are asked, and concern that the views of vulnerable patients and carers may be excluded.

This study asks: can we make data about patient experience of NHS services easier to collect, and more useful, by improving the way we analyse and present this data?

By talking to and working alongside patients, carers and staff this study aims to collect information in 'real-time' – that is, as soon as possible after services have been experienced. We are interested in finding out what information needs to be collected and how best to do this.

We also want to look at more efficient ways of analysing the information that patients and carers provide and how to present it in a meaningful way that will enable staff to use in ways that will make a difference to NHS services.



Drawing on these conversations, we will formulate, introduce and train staff to use a toolkit that will provide:

- i. guidance about collecting patient feedback
- ii. new methods for analysing patient feedback data
- iii. new methods and guidance for presenting patient feedback data

This project will focus on the experience of patients with either a musculoskeletal condition or who have experienced serious mental illness, and the staff who support them in one of three settings:

-
- i. the rheumatology outpatients (OPD) within xxxxxxxx
 - ii. a community mental health team within xxxxxxxx Trust
 - iii. one of two general practices



Who is doing the research?

This project is being undertaken by the University of Manchester in collaboration with xxxxxx, and is funded by the Health Services and Delivery Research (HS&DR) Programme of the National Institute for Health Research (NIHR).

Why have I been invited?

You have been invited to participate in this study because you are either:

- i. a healthcare professional or other staff member working in xxxxxxxxxx in support of service users with a musculoskeletal condition, or
- iv. a mental healthcare professional or other staff member working in xxxxxxxxxx
- ii. in support of service users a mental health problem, or
- iii. a healthcare professional or other member of staff working in a GP practice with experience of supporting patients with either musculoskeletal condition, or a mental health problem.
- iv. A member of Information Management & Technology (IM&T) staff in NHS settings included in this study

Do I have to take part?

Participation in this study is entirely voluntary and you are free to withdraw at any point without giving a reason. Should you agree to participate, you will be asked to sign a consent form. Taking part or not taking part will have no effect on your current job.

What will I have to do if I participate?

There are 4 phases (known as 'work streams') to the study. During work streams 1, 3 and 4, members of staff will be invited to take part in either a group discussion or will be interviewed individually by a member of the research team. You will be invited to take part in one interview or focus group discussion when you join the study. You will be invited to take part in either a face-to-face interview, or a telephone interview. Some participants will also be invited to take part in another interview or focus group at later points in the study. You will be invited to take part in a maximum of 3 interviews or focus groups, but you do not need to take part in all of these to be a participant.

In work stream 1, staff will be encouraged to discuss their views and experiences relating to how patient experience data is currently collected. Staff will be asked about the strengths and weaknesses associated with the data that is gathered, the processes used to analyse it and the extent to which the information that emerges is effectively used.

In work stream 2 computer scientists who are members of the research team will design ways to analyse free text comments that are already being collected in the NHS. We are not asking staff to take part in this stage.

In work stream 3, a proportion of staff who took part in initial focus groups and interviews in work stream 1 will be asked to take part in a group discussion to help co-design the toolkit for the NHS that will offer innovative solutions for collecting, analysing and presenting patient feedback data.

In work stream 4, we will introduce the toolkit into the work settings of the teams who participated in the earlier discussions. Staff in each setting will be trained in how to use it and after a period of time, will be asked if they would be willing to take part in an exercise during which they will be observed using the new toolkit. Staff will then be invited to participate in a third group discussion or interview to assess the impact of the toolkit on enhancing the usefulness and relevance of the data collected, the functionality of the toolkit, and its influence on service change.

As part of an economic evaluation of the toolkit, staff will also be invited to take part in a short survey aimed at understanding the time and resources required to use it.

Focus groups will take place at central locations, and individual interviews will take place at a place and time that is convenient for you. We will audio-record these discussions so that we can fully consider and review all that is said. A University of Manchester approved supplier, '1st class secretarial', will transcribe them before we analyse them in our research. 1st class secretarial trades under the name Lawson Hardwick Ltd and has signed a confidentiality agreement with the University of Manchester Website: <https://www.1stclass.uk.com/>

The focus group discussions will last approximately 1½ - 2 hours. The individual interviews will last approximately 30 minutes.

Will I be paid?

You will not be paid directly for your participation in the study but, as is usual practice, time spent participating in the study will be reimbursed to your NHS trust via support costs agreed by the Comprehensive Local Research Network (CLRN). Your involvement in the research will be during normal working hours and should not change the general requirements of your work or the size of your caseload.



What are the possible disadvantages and risks of taking part?

Because of work commitments you may not be able to take part in all activities. However, you can choose to take part in some or all activities. You can also take a break at any time.

Can I withdraw from the study?

You may withdraw from the study at any point without giving a reason.

What are the possible benefits of taking part?

There are no direct personal benefits to taking part, although you will have the opportunity to participate in research that aims to contribute to the improvement of NHS services by increasing the credibility, usefulness and relevance of patient experience data in services for people with long-term conditions.

What happens when the research study stops?

We will send you a summary of our findings. Please be aware that consent forms, audio recordings and written notes based on interviews and focus groups, will be securely stored at the University of Manchester for a period of ten years.

Will my taking part in the study be kept confidential?

Yes. Whatever you say in interviews or focus groups will be kept confidential and will not be discussed with anyone else. Any information that identifies you will be anonymised or changed to your research ID at the earliest opportunity. Your real name and job title will not be used in anything that we write or publish, although we may provide your generic role e.g. member of clinical staff, nurse, a member of hospital management.

Individuals from the University of Manchester, NHS Trust or regulatory authorities may need to access data collected during this study to ensure the research is being conducted appropriately. With your permission, this will include data from which you could be identified. All individuals have a duty of confidentiality to you as a participant in this research.

What will happen to the results of the research study?

We will make the findings and toolkit available to NHS management nationally and in other hospitals so they can use the results to consider wider implementation and improvement to services in the NHS. We will also publish the results in research publications and inform patient groups about the results.

Your rights:

You should decide on your own whether or not you want to take part in this study. If there are any questions that you do not wish to answer, then you can leave these out. You may stop participating in the study or withdraw some of your data at any time, and this will have no effect on your job.

Ethics Approval:

NHS Research Ethics Committee.

What if there is a problem?

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. If they are unable to resolve your concern or you wish to make a complaint regarding the study, please contact a University Research Practice and Governance Co-ordinator on 0161 275 2674 or 275 2046 or by email to research.complaints@manchester.ac.uk.

Contact Details:

If you wish to discuss participation in the study, any aspects of this information sheet or have any questions or concerns about the study please contact:

Dr Caroline Sanders: Tel: 0161 275 7619
Email: Caroline.Sanders@manchester.ac.uk

Dr Nicola Small: Tel: 0161 275 4844
Email: Nicola.Small@manchester.ac.uk

Dr Papreen Nahar: (Tuesday to Thursday) 0161 275 7640
Email: Papreen.nahar@manchester.ac.uk

Centre for Primary Care
University of Manchester
Williamson Building
Oxford Road
Manchester
M13 9PL

What do I need to do next?

You will be requested to take part in an interview/focus group related to your experience of collecting and using patient feedback within your service area. The interview will be conducted in your convenient place. If any question is unclear to you or if you want to know why are we asking you these questions, you can ask us. It would take about 30 minutes for the interview, and 90 minutes for a focus group. We will take all the care to make you comfortable but if you do not want to continue to talk or want to arrange it another time, we will do that accordingly.

Staff written consent form

Institute of Population Health, The University of Manchester
Williamson Building, Oxford Road, Manchester, M13 9PL
Tel: 0161 275 7619 and/or 0161 275 7640

Study Title: Developing and Enhancing the Usefulness of Patient Experience and Narrative Data (DEPEND)
Chief Investigator: Dr Caroline Sanders

➤ Please **initial** the box if you agree with the statement:

1. I confirm that I have read and understand the information sheet (dated..... v.....) for the above study and have had the opportunity to ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason. ☐
3. I understand that the DEPEND team will hold my contact details so that I can be contacted during the project to arrange my participation in study activities. ☐
4. I understand that I may be invited to consent to participate in up to three interviews or group discussions, and I am happy to be contacted for this purpose. ☐
5. I give consent for any interviews or group discussions that I participate in to be audio recorded and transcribed. ☐
6. I agree to the use of anonymised quotations from interviews being reported in research reports, journal articles and presentations. ☐
7. I understand that later in the study (during Work Stream 4) I will be asked to attend a training session relating to the use of a toolkit designed to enhance the collection, analysis and presentation of patient satisfaction data and I am happy to be contacted for this purpose. ☐
8. Later in the study (during Work Stream 4) I am willing to allow a researcher observe me in the work place. ☐
9. I understand that I will be invited later in the study (during Work Stream 4) to complete a short survey relating to the economic evaluation of the toolkit that will be introduced as part of the study, and I am happy to be contacted for this purpose. ☐
10. I understand that data collected during the study may be looked at by individuals from the University of Manchester, from regulatory authorities or from the relevant NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my data. ☐
11. I understand that when this research is completed the audio file will be retained and securely archived for a period of 10 years. This archive can only be accessed by request from the research team and all files will be destroyed at the end of that period. ☐
12. Would you like to be contacted again with either (a) participation in a further stages of the project ☐ (b) the final results of the study ☐ (c) other similar studies. ☐

Name of Participant

Date

Signature

Researcher

Date

Signature

Appendix 3 The DEPEND study toolkit

- 3.1 A survey utilising the Friends and Family Test (FFT) with space for free text comments to be completed via digital kiosk within study sites, a website online, pen and paper version
- 3.2 Guidance and information to support use of the new tools for:
 - 3.2.1 Staff
 - 3.2.2 patients and carers;
- 3.3 New text mining programmes for analysing patient feedback data
 - 3.3.1 The open-source code for the text mining analysis is provided via the following website <http://gnteam.cs.manchester.ac.uk/depend/>
- 3.4 New templates for reporting feedback
- 3.5 A new process for eliciting and recording verbal feedback within community Site B services.

DEPEND online survey

Site C1 Online Survey Draft

Survey Created:
Survey Changed:

Client: DEPEND

Survey: Site C1 Medical Practice Survey ONLINE Version copy

ID: 14

Link to DEMO version (use Internet Explorer)

<https://secure.crtviewpoint.com/Online/Survey/Preview/624A14>

Start Message: Please tell us what you think about our general practice service at Site C1 Medical Practice

End Message: Thank you for helping us to improve our general practice service at Site C1 Medical Practice.

1. This is a short survey with two main questions. To keep your answers private, please do not enter any personal information such as names or contact details.

(Multiple Choice, select one only)

- Press here to continue

2. Are you a patient at Site C1 Medical Practice?

(Multiple Choice, select one only)

- Yes
- No

Route to: Q4

3. Do you provide care for someone who uses the general practice service at Site C1 Medical Practice?

(Multiple Choice, select one only)

- Yes
- No

Route to: Q9

4. 1. How likely are you to recommend this general practice service to your friends and family if they needed similar care or treatment?

(Multiple Choice, select one only)

- Extremely unlikely
- Unlikely
- Neither likely nor unlikely
- Likely
- Extremely likely

5. 2. Please tell us about your experience: (including good, bad or just ok)

(Open Ended)

- Free Format Text

6. Are you?**(optional question)***(Multiple Choice, select one only)*

- | | |
|-----------------------|--------------|
| • Male | Route to: Q8 |
| • Female | Route to: Q8 |
| • Other | Route to: Q7 |
| • I prefer not to say | Route to: Q8 |

7. Please specify*(Open Ended)*

- Free Format Text

8. What age are you?**(optional question)***(Multiple Choice, select one only)*

- | | |
|-----------------------|---------------|
| • 0-15 years | Route to: End |
| • 16 -24 years | Route to: End |
| • 25-34 years | Route to: End |
| • 35-44 years | Route to: End |
| • 45-54 years | Route to: End |
| • 55-64 years | Route to: End |
| • 65-74 years | Route to: End |
| • 75-84 years | Route to: End |
| • 85 years | Route to: End |
| • I prefer not to say | Route to: End |


9. Thank you for your interest in giving feedback. At the moment this survey is only for people who use the general practice service here at Site C1 Medical Practice.*(Multiple Choice, select one only)*

- | | |
|--------------------------------|---------------|
| • Press here to end the survey | Route to: End |
|--------------------------------|---------------|

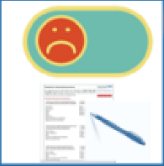
DEPEND patient and carer guidance and information to support use of the new tools

NHS
National Institute for
Health Research

Let us know about what you think about our service in rheumatology



**Describe what
we did well**



**Describe what
we could do to
improve**

"I was extremely impressed with the care I received from all the medical and support staff. They all treated me with great professionalism and humour"


"Good service. Staff hard pressed. Inter-departmental communication is a weakness"

"The department was extremely busy, and I was amongst a lot of other people, but I was still very well looked after"


"Everyone in rheumatology were professional and friendly, and made me feel more comfortable even when I was scared, or worried. The team are a credit to the NHS."

Please complete our survey using one of these methods



Touch screen survey at our
new kiosk



Pen and paper survey
available



Online survey available at

MANCHESTER
The University of Manchester

Staff Guidance for Feedback

A new approach to collecting and documenting feedback from service users for quality improvement

What is the DEPEND project and why is a new approach needed?

The [DEPEND research project](#) is investigating the best ways to collect and use feedback from service users and carers to make it useful for quality improvement within specific staff teams. The Site B (Site B) is working in collaboration with the University of Manchester for this project. Based on our initial stage of the research, we found that staff, service users and carers thought that survey questions are not always helpful. This is because the questions are very general, and many people will not respond, especially when unwell. Staff and service users talked about the value of face-to-face discussion about experiences of Site B services. This was viewed to be something that happens anyway during visits and conversations but is not always documented in detail or systematically.

What is the new approach?

The new approach suggested by the team is to collect and document feedback based on their visits systematically and formally within the 'plan, outcome, action' field of the electronic care record.

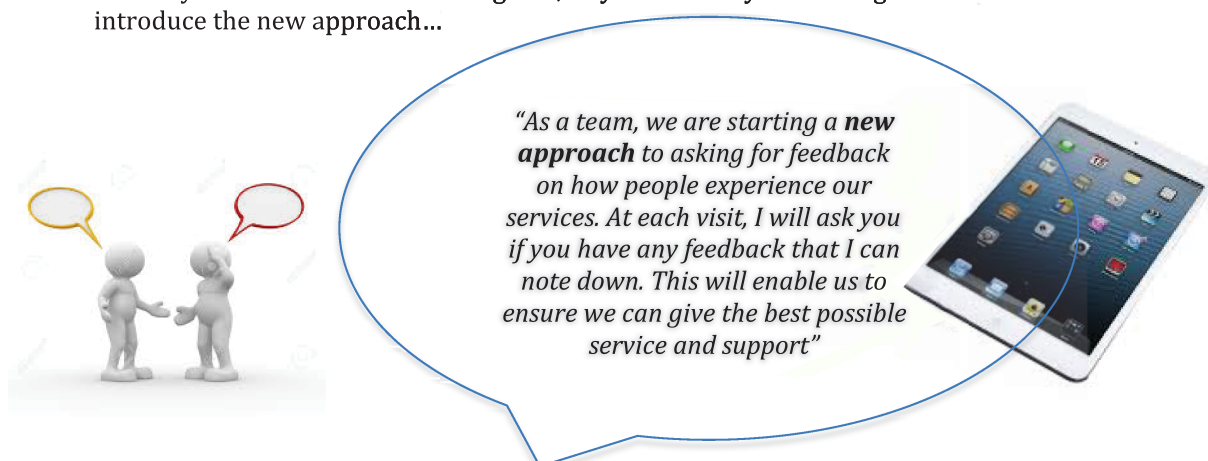
The team have also drafted some questions that might help to trigger this discussion routinely at the end of a visit. However, use of questions needs to be flexible and sensitive to individual circumstances and context of each visit.

Capturing the discussion formally as feedback will provide a more sensitive and inclusive process. This should also be more consistent with a recovery model of care. It might help to make feedback more meaningful and useful to individual staff, the team locally, and the wider Trust.

Having discussions about experience of service formalised as feedback and analysed in an aggregated way will also allow common issues to be identified and reflected upon at team meetings, and enable these to inform best practice.

How should staff tell service users about the new approach?

At home visits, staff should let service users and carers know that as a team you are testing out a new approach to the way you currently collect feedback and explain this. You may like to use this sheet as a guide, or you could say something like this to introduce the new approach...



Staff need to ensure service users know that they are not under any pressure to give feedback in this way, and they can also give feedback using the usual anonymous surveys done within the Trust.

Initiating a discussion about experience of service

The team recognise that sometimes feedback is given naturally in discussions with service users, and sometimes people ask specifically about this at the end of a visit, but this varies according to circumstances. Sometimes specific questions can trigger a discussion to enable clients to give feedback. The team have drafted a few questions (see Table 1) as a guide to prompt such discussions.

Table 1: Possible questions

Example trigger questions
How are you feeling after today's visit?
Can you say what has been helpful, or unhelpful?
Is there anything that you have found difficult?
What are your views about what we've talked about today?
Have we addressed all the issues needed today?
Have we covered all the questions that you'd like to ask?
Have your expectations been met today? [Possible prompts: timekeeping, communication]
Was everything useful [or 'of value'] that we discussed today?
Is there anything you'd like to change [or, 'look at in more detail'] as a result of today's visit?
What shall we plan in terms of the next visit?

Documenting the feedback?

The care coordinator will record the discussion about their experiences of care at the end of the visit within the 'plan, outcome and action' field in the electronic care record. Please note down any discussions or comments that are relevant, and if you used a specific question. It is also helpful to record reasons why feedback is not documented; for example, if the circumstances made this difficult, or if the client did not want to give any feedback. This will help us to understand how useful this new process is.



When will this new approach start?

We will start the new approach of collecting and documenting feedback at the beginning of September 2017 and it will be tested over a period of approximately 6 months until the end of February 2018.

How will the team and the Trust leads receive the analysis and presentation of feedback?

The DEPEND team will summarise and produce a report of what feedback is given in this way each month. We will liaise with the team to consider how best to summarise information to make it most useful for the team. We will also summarise issues raised by the team regarding the process itself, including any benefits and problems.

What other methods are available for service users and carers to give their feedback?

The Trust is also introducing a new programme of collecting feedback digitally using iPad devices in specific areas around the Trust. Service users can also give their feedback using an online version of the questionnaire, available here: [] or using the QR code here:

However, we know it is unlikely that one method of collecting feedback will suit all clients. Service users and carers will still be able to complete the existing pen and paper questionnaires to give their feedback. The idea is that this new process of collecting verbal feedback being tested out by the community team will provide an additional way of enabling people to give feedback. Importantly, the new process should provide more useful information for the team to reflect on their specific practice, enabling ongoing high quality service provision.

What support is available for the team?

There is a DEPEND researcher, [REDACTED] who will support your team throughout this 6 months to encourage the new process of collecting and documenting feedback to evolve and improve and embed into everyday practice. [REDACTED] will be observing how this works in practice. Please contact [REDACTED] if you have any questions:



Appendix 4 Dissemination

Conference presentations and publications

During the life of the study we presented data at multiple national and international conferences, as well as at academic seminars:

- Daker-White G, Dehghan A, Shepherd A, Nowokowska M, Kontopantelis E, Nenadic G, Sanders C. *Humans Versus Machines: Text Mining Versus Adapted Grounded Theory in the Analysis of Free Text Data from Patient Feedback Surveys*. Health Services Research UK (HSRUK) Conference, Nottingham, UK, 4–5 July 2018.⁶²
- Dehghan A, Kayesh H, Daker-White G, Sanders C, Nenadic G. *Mining Free-text Patient Feedback Comments*. Heal-TAC, Manchester, UK, 18–19 April 2018.⁶³
- Nahar P, Ong BN, Small N, Hodgson D, Nenadic G, Bower P, et al. *Implementing and Evaluating Tools to Improve the Collection and Usefulness of Patient Experience Data in Multiple Service Contexts*. Health Services Research UK (HSRUK) Conference, Nottingham, UK, 4–5 July 2018.⁶⁴
- Nahar P, Sanders C, Small N, Hodgson D, Daker-White G, Spencer R, Sanders C. *Creating Meaningful Patient Feedback Data for Health Service Improvement: Exploring the Formality and Informality of Feedback Mechanisms*. British Sociological Association (BSA) 49th Medical Sociology Annual Conference, York, UK, 13–15 September 2017.⁶⁵
- Sanders C, Nahar P, Hodgson D, Small N, Daker-White G, Sharp C, Ong BN. *Constructing and De-constructing Patient Experience via Big Data and Small Data*. World Congress of Sociology, Toronto, ON, Canada, 15–21 July 2018.⁶⁶
- Sanders C. *Story Specimens and Chemistry: a Creative Enquiry*. Society for Academic Primary Care (SAPC) 47th Annual Scientific Meeting, London, UK, 10–12 July 2018.⁶⁷
- Sanders C. *The Ethics and Politics of Sharing Stories for Patient and Public Involvement*. British Sociological Association (BSA) 50th Medical Sociology Annual Conference, Glasgow, UK, 12–14 September 2018.⁶⁸
- Small N, Nahar P, Daker-White G, Spencer R, Hodgson D, Bower P, Sanders C. *Can Digital Data Capture and Improved Analysis of Comments Make Patient Feedback More Meaningful and Useful for Primary Care?* Society for Academic Primary Care (SAPC) Annual Conference, Warwick, UK, 12–14 July 2017.⁶⁹
- Ong BN, Sanders C. Exploring engagement with digital screens for collecting patient feedback in clinical waiting rooms: the role of touch and place. *Health* 2019.
- Ong BN, Hodgson D, Small N, Nahar P, Sanders C. Implementing a digital patient feedback system: an analysis using Normalisation Process Theory. *BMC Health Serv Res* (accepted for publication).

Learning set and related dissemination workshop

We have contributed to a learning set with members from other projects funded under the same commissioned call. This has led to two joint conference presentation panels where Caroline Sanders presented with chief investigators from other studies on the use and usefulness of patient experience data:

1. Robert G, Locock L, Sanders C, Sheard L. *Enhancing the Use of Patient Experience Data for Improving the Safety and Quality of Care*. ISQua 34th Annual Conference, London, UK, 1–4 October 2017.⁷⁰
2. Roberts, G, Sheard L, Loockock L, Sanders C. *Exploring and Enhancing the Use of Patient Experience Data for Improving the Quality of Care*. Health Services Research UK (HSRUK) Symposium, Nottingham, UK, 6–7 July 2017.⁷¹

In addition, we presented at a dissemination workshop^{72–74} at The King's Fund in June 2018 for studies commissioned under the themed call on the 'use and usefulness of patient experience data'. Multiple

stakeholders attended including policy-makers and members of the NIHR. All studies formulated summary posters to outline core research findings and these were used as a basis for presentations and discussion and to distil the overarching key findings and implications for policy and practice. A plan was made to produce a joint output reporting on these key findings and implications for dissemination by the NIHR dissemination centre.

Animation

In March 2018 an animated video outlining the study findings was developed and produced to disseminate key learning from the study to a diverse audience, including members of the public, patients and carers, clinical staff and managers in NHS trusts and other health-care organisations:

NIHR Greater Manchester Patient Safety Translational Research Centre. *DEPEND: Use of Digital Methods for Collection and Use of Patient Experience Data*. 2018. URL: www.youtube.com/watch?v=BOYLxVJAzdl (accessed 23 December 2019).⁷⁵

Toolkit

The text-mining programs and a user manual for carrying out analysis using the software is available: <http://gnTEAM.cs.manchester.ac.uk/depend/> (accessed 18 October 2019).

Patient and public involvement

A PPI workshop was delivered in February 2018, with presentations on the DEPEND study (including PPI) research findings. Members of our PPI team co-presented the workshop and facilitated discussions. We co-designed an animated video reporting the study findings and this video was shown for the first time at the workshop. Nineteen PPI members attended the workshop, with representation from study sites and multiple PPI groups and networks across Greater Manchester. We had much positive feedback and were able to use the comments and contributions to make final tweaks to improve the animation further prior to general release.

Two members of our PPI group presented a paper with Nicola Small at an international conference:

Small N, Lewis AM, Allen D, Ong BN, Sanders C. *Co-designing New Tools for Collecting, Analysing and Presenting Patient Feedback in NHS Service: Working in Partnership with Patients and Carers*. International Perspectives on Evaluation of PPI in Research, Newcastle University, New Castle upon Tyne, UK, 15–16 November 2018.⁷⁶

The following paper is currently being finalised for submission to a leading PPI journal:

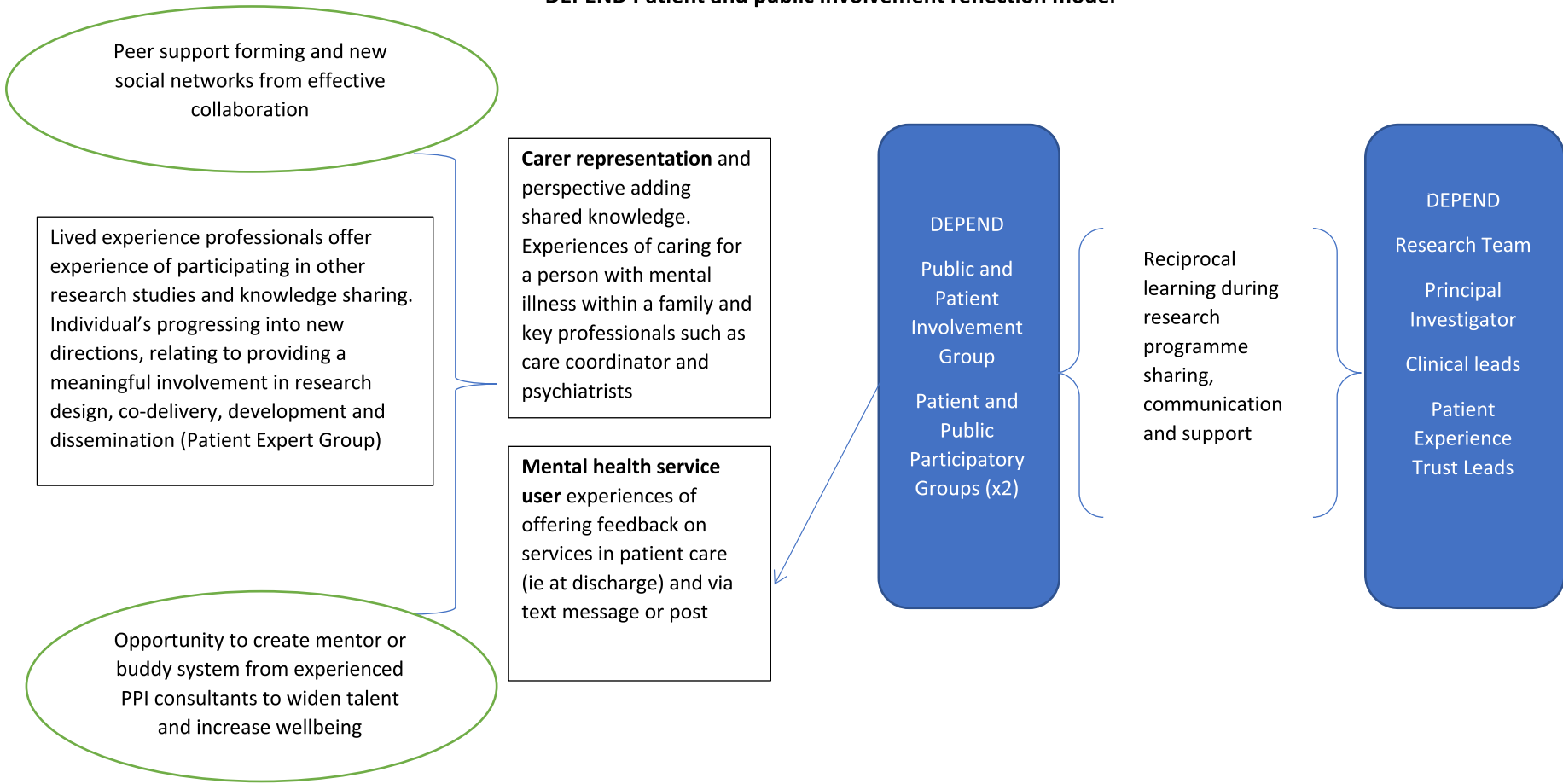
Small N, Ong BN, Lewis A, Allen D, Bagshaw N, Sanders C. Co-designing new tools for collecting, analysing and presenting patient feedback in NHS services: working in partnership with patients and carers *BMC Research Involvement & Engagement*.

Health-care providers

In addition to conference presentations and workshops, we have held specific dissemination seminars in each of the participating sites on completion of the study.

Appendix 5 The DEPEND study patient and public involvement reflection model

DEPEND Patient and public involvement reflection model



Appendix 6 Tables and supplementary material for the text-mining results

TABLE 22 Distribution of topics in the 'gold standard' data set

Aspect classes	Site A data set		Site B data set	
	Examples	Percentage	Examples	Percentage
Care quality	86	12.57	367	36.55
Staff attitude	170	24.85	176	17.53
Waiting time	103	15.06	98	9.76
Process	51	7.46	91	9.06
Resource	19	2.78	62	6.18
Communication	54	7.89	42	4.18
Environment	29	4.24	39	3.88
Food	10	1.46	18	1.79
Parking	6	0.88	3	0.30
Other	74	10.82	106	10.56
Not feedback	82	11.99	2	0.20
Total	684	100	1004	100.00

TABLE 23 Evaluation results (precision, recall, F₁-score) cross-validated on the 'gold standard' data sets for the integrated model

Theme	P-optimised			R-optimised			F ₁ -optimised		
	P (%)	R (%)	F ₁ (%)	P (%)	R (%)	F ₁ (%)	P (%)	R (%)	F ₁ (%)
Site A									
Care quality	54.56	12.21	19.73	47.09	66.28	55.01	47.09	66.28	55.01
Staff attitude	90.95	25.19	39.00	64.69	81.16	71.88	77.95	69.98	73.57
Waiting time	78.48	31.94	44.64	58.66	88.45	70.46	69.99	76.60	72.99
Environment	69.22	09.50	16.44	57.05	38.62	45.56	57.90	36.62	44.08
Other	48.53	11.26	18.22	43.80	55.55	48.92	43.80	55.55	48.92
Micro average (segment) ^a	71.17	19.02	29.86	54.22	69.92	61.08	58.12	64.36	61.06
Micro average (comment) ^a	56.13	16.19	24.98	32.10	75.89	45.10	35.95	66.58	46.66
Site B									
Environment	91.21	28.11	42.91	41.98	51.63	45.98	64.33	44.39	51.96
Waiting time	94.38	47.13	62.71	76.78	85.83	81.01	92.74	81.62	86.75
Staff attitude	75.92	18.37	29.53	61.52	51.50	55.98	66.56	48.17	55.88
Care quality	86.06	19.43	31.52	71.22	96.14	81.78	71.41	95.83	81.80
Other	57.84	05.25	09.45	34.58	11.83	17.48	34.58	11.65	17.27
Micro average (segment) ^a	85.25	20.97	33.62	67.02	73.08	69.91	70.89	71.35	71.11
Micro average (comment) ^a	77.67	16.80	27.58	55.22	83.63	66.49	60.11	80.89	68.94

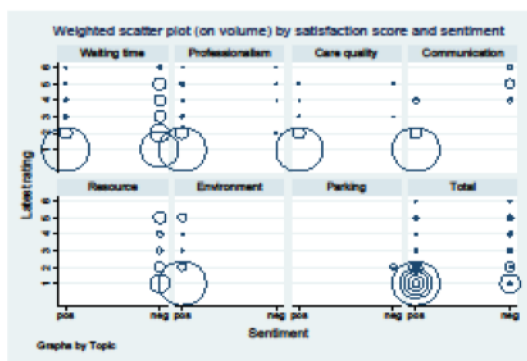
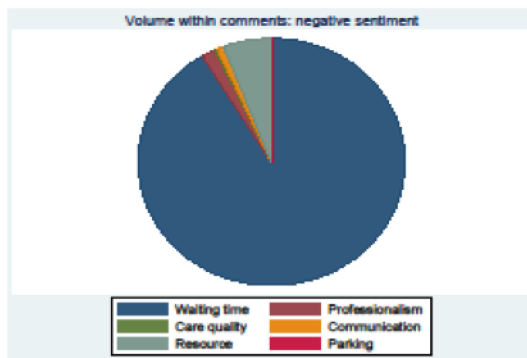
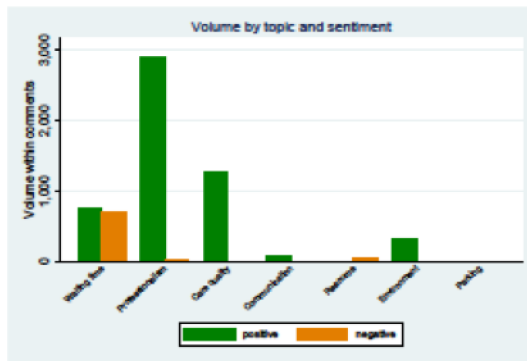
P, precision; R, recall.

^a Micro average (segment) and micro average (comment) refer to evaluation on individual segments and the whole comment, respectively.

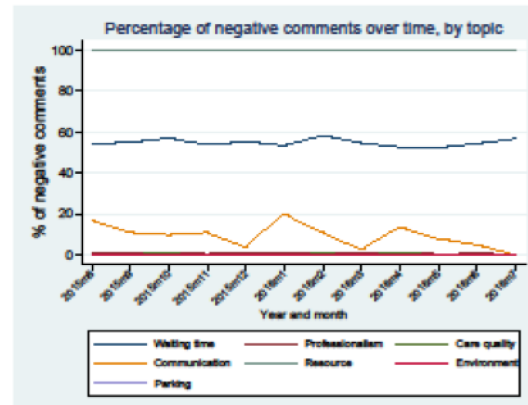
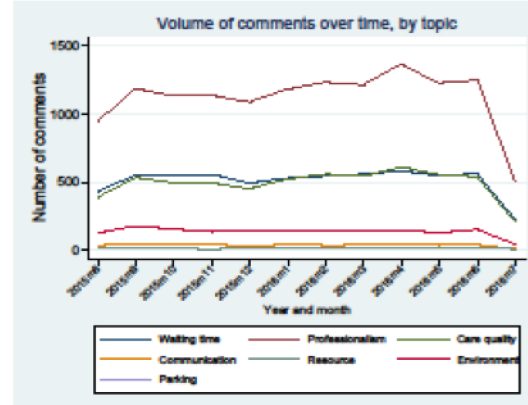
DEPEND 3 month report June 19, 2017

Evangelos Kontopantelis
e.kontopantelis@manchester.ac.uk

1 LAST 3 MONTHS, AGGREGATED



2 LAST 12 MONTHS, TIME TRENDS



3 LAST 3 MONTHS: MOST NEGATIVE COMMENTS (5 MAX FROM EACH TOPIC)

Professionalism

1: Rude, inconsiderate staff wrongly diagnosing my child. Taking one glance to say he was fine and asked every other child in first. The main issue was the know it all answers I got when asking why children coming in after were being seen first. I walked out, took him to another hospital where he was diagnosed with Scarlet fever and given a 2 week course of penicillin. Never will I be back again or recommend to anyone. In fact I strongly recommend not to go when it is brought up in conversation.

2: ** some text is missing ** ur hospital. I plan to

© 2017 Evangelos Kontopantelis e.kontopantelis@manchester.ac.uk

FIGURE 13 Automated report to summarise and display the results of the text-mining analysis (example).

Text mining

The open-source code for the text-mining analysis is available at <http://gnTEAM.cs.manchester.ac.uk/depend/> (accessed 11 October 2019).

Instruction manual

DEPEND SYSTEM

Introduction

This document includes the instructions to run the DEPEND aspect extraction system and classify new patient comments. The system consists three sub-systems: System A, System B and a script to combine two systems. The document also discuss the processes to setup necessary environment run the system.

Environment setup

This system was developed and tested on a Unix system (Ubuntu 16.04, 64-bit), so it is recommended to use a unix system for the best performance.

Install Python

Install python 2.7 to run the system. After installing python, nevigat to the project root directory and run the following commands to install the required python modules.

```
pip install -r requirements.txt
```

NLTK data

System A requires NLTK stopwords corpus to run. If you do not have have nltk python module installed, run the following command to install:

```
$ pip install nltk
```

To download nltk corpus, run the commands below after starting the python command prompt:

```
> import NLTK
> nltk.download('stopwords')
```

- In the Python command line type "import NLTK" and press Enter
- After that type "nltk.download('stopwords')"

Install R

System B was developed in R. It requires R version 3.2.3. Apart from the built in packages, the system requires the following R packages:

- plyr 1.8.4
- dplyr 0.7.4
- NLP 0.1-1d
- tm 0.7-1

- RWeka 0.4-34
- e1071 1.6-8
- caret 6.0-77
- RTextTools 1.4.2
- Glmnet 2.0-13
- kernlab 0.9-25
- Mlr 2.11

To install a R package, start the R command prompt and run:

```
> install.packages("plyr")
```

Run System

Run the following command to make prediction:

```
./run_prediction.sh <path/to/data/file.csv> <dataset type>
```

<path/to/data/file.csv> should be replaced by the csv file path that contains comments patient comments. An example datafile is given at 'sample_data_file.csv'. Moreover, <dataset type> should be replaced by either `IMBCT` or `SFFT` based on the type of the data source. By default, the outputs will be saved in the root directory of the system. There are three types of outputs:

1. `output.csv`: this file contains prediction per comment.
2. `top_comments_system_a.csv`: Top five comments predicted by the System A
3. `top_comments_system_b.csv`: top five comments predicted by the System B

Appendix 7 Staff Information Sheet

Patient and Carer Feedback for the Site A Service

The DEPEND research project is being conducted in collaboration with the University of Manchester and is taking place in our Rheumatology Department. The project aims to improve the way in which anonymous patient feedback is collected, analysed and used. One of the methods to improve the collection of digital feedback is to install a patient feedback kiosk which will look similar to this:



What questions will be asked on the kiosk?

The kiosk will have two main friends and family test questions on

1. How likely are you to recommend this service to your friends and family if they needed similar care or treatment? 2. Please tell us about your experience: For example, what went well? what could be improved?

There will also be screening questions to make sure that only patients and carers who use Site A services answer the survey and optional demographic questions at the end (sex and age). To see a demo version of the survey on your PC of this survey please [click here](#) (please use internet explorer).

Won't patients already receive a text with the Friends and Family test on?

While the kiosk is being tested out, patients will not get the text message asking for the friends the family feedback. This decision was made to ensure that patient's aren't asked twice (to lower the risk of survey fatigue).

When is the kiosk arriving?

The kiosk will arrive on Wednesday 28th June and will be tested over a period of approximately 9 months until March 2018.

Where will the kiosk be positioned?

The kiosk will be next to the desk (which is sometimes used by the eye hospital) in Area 4 Site A Outpatients. This position isn't ideal as it's quite tucked away but due to the size of the kiosk it was deemed the best place to start with so there were no health and safety concerns. Over the 9 month period we could try out moving it to other positions to see if it helps improve the number of people who use it. Staff name and Staff name will approve any position changes.

How should I encourage patients and carers to use the kiosk?

It would be great if you could mention to patients that we have a new kiosk to collect digital feedback and explain where it is. We are particularly trying to capture positive feedback as well as areas for improvement.

What other options are there for patients and carers to give their feedback?

There will still be the option for patients and carers to complete the pen and paper Site Patient Experience survey in the waiting area. Over the coming weeks we will also develop an online version of the kiosk survey and publicise the link in a poster and potentially the Site A website. We have a poster we can put up to show patients and carers the different ways they can provide their feedback.

Which company supplies the kiosk and software?

The company is called CRT Viewpoint and they are based in Coventry. In case of any faults with this kiosk their customer service number is...

Who should I contact if there is a problem?

Please contact Papreen for information about supporting patients to use the kiosk.

Researcher: Staff name

Staff email

Tel:

Project Manager: Staff name

Staff email

Tel:

Project lead at Site A: Staff name

Staff email

Project Lead: Staff name

Staff email

Tel:

What support is available for patients to use the kiosk?

There will be DEPEND researcher (Staff name) who will attend at various times throughout this 9 months to observe how it is being used. Papreen will also organise with the Site A volunteers extra support and explanations for patients and carers about how to use the kiosk while they are in the outpatient area. This will probably be only for the first few weeks after it has been installed. We will also put the Friends and Family test leaflet (from the NHS England website) in the outpatient's area to explain a bit more about the test.

How will the team receive feedback?

Survey responses will be collected by the DEPEND team and reports will be produced each month for the service which will be sent to Staff name and Staff name for discussion within the team. See the few examples attached for what this might look like. At a corporate level the reports will be sent to Staff name, Staff name and Staff name who have all be involved with the design of the tools for DEPEND.

How should the kiosk be cleaned?

The kiosk can be cleaned with a detergent wipe on a regular basis, however if it was to become contaminated with bodily fluids it will be cleaned with a bleach solution. The domestics manager Staff name has been advised of this.

Who has been involved in the approvals of the Kiosk?

The DEPEND project and tools to help improve feedback have been discussed in detail with Staff name, Staff name, Staff name, Staff name and Staff name.

From an information governance perspective, the survey has been approved by Staff name.

From an infection prevention and control perspective, the kiosk has been approved.

From an electrical safety perspective the kiosk will be PAT tested by an electrician from estates before first use.

In terms of the placement of the kiosk Staff name – Outpatient's Manager and Staff name have been involved in this decision.

Site Monthly Report

Developing and Enhancing the Usefulness of Patient Experience and Narrative Data:
DEPEND study

**Summary report of the Friends and Family Test data
Captured 1st Aug to 31st 2017 on the new digital kiosk**

Results

- 43 surveys were completed and 1 of these was invalid because that was not a patient or carer within the Site A service, it was a test survey.
- Of the 42 respondents, 71% were patients and 29% were carers.
- 47% of respondents were 'extremely likely' to recommend the Site A service to their friends and family if they needed similar care or treatment.
- 15% of respondents were 'extremely unlikely' to recommend the Site A service to their Friends & Family.
- The majority of the respondents were female (56%).
- The highest number of respondents were between the ages of 45-54 years of age (30%).

Free text comments

Total 17 respondents typed free text comments on their experiences of service at the practice.

Positive feedback	Negative feedback
Of the 17 respondents, 35% (15) typed positive comments about the service in the free text box, such as <i>'everything is right'</i> . They have also expanded on why they think so. As one respondent said, <i>'The service which I received was amazing, the staff and doctor was really polite and made me feel very welcome'. Thank you'</i> . Service users are so satisfied that one of them said, <i>'very thorough, no need to improve, well lead'</i> .	In comparison, only 2 people have expanded on why they had a negative experience. Just one main issue has been raised was waiting time. Examples are listed below:
<u>Staff attitude and professionalism</u> Many positive comments were expressed on staff professionalism such as, <i>'lovely staff, all amazing, and so helpful. Kind and loving people'</i> , and <i>'amazing'</i> . 'Excellent consultant (Dr. xxxx). Positive steps has been taken to prevent further joint damage (hopefully)'. 'Excellent staff'	
<u>Care quality</u> 'Helpful and understanding'. 'Nice staff , very friendly'.	

<p>'Lovely staff, all amazing and so helpful, kind, loving people'</p> <p>Below are the words used to describe the characteristics of the staff:</p> <p>'Very well', 'very good', 'nice', 'friendly', 'understanding', 'helpful', 'cheerful', 'polite', 'excellent', 'made me feel very welcome'.</p> <p>Also, one of them said, 'Injection for pain great improvement'.</p> <p>'Injection for pain, great improvement'</p>	
<p><u>Access</u></p> <p>'Booking and check in is good...'</p>	<p><u>Access</u></p> <p>'I am out of area and the blood tests taken by my GP are not accessible by Site A so this has affected my treatment at times'</p>
<p><u>Waiting time</u></p> <p>'Short wait time'</p> <p>'Usually seen on time'</p> <p>'Quick service'</p>	<p><u>Waiting time</u></p> <p>'Long wait'</p> <p>'Very late'</p>

We need your help

- If you notice the screen has frozen on the kiosk, please switch it off and on again at the socket and it should re-boot and work fine again.

We are looking for examples of where the feedback has led to positive improvements to the service which we can put up on a 'you said, we did' poster. It would be great if you indicate your preferred examples of this kind from the report.

- We are interested to hear from staff who may want to ask new questions on the survey that would be useful for the service and staff with any ideas for further advertisements of different feedback options. Please get in touch with Papreen Nahar to discuss this:
Email: Papreen.nahar@Manchester.ac.uk Phone: 01612757640

Next steps

- We are aware that the kiosk screen keeps 'freezing' so have placed a note on the kiosk to say if this happens it will need to be switched off and on again.
- We will be coming to interview staff and conduct a focus group to get staff views about the Kiosk use for feedback. We are hoping to discuss how the monthly report is useful to them and whether it is possible to improve further. Also we will be coming along to observe the staff team meetings where the feedback is being discussed to see how this is impacting on care.
- We have changed the wording of the free text question slightly to encourage more detailed responses.
- We have increased the font size of the electronic survey to make it more eye-catching and easy to read.

- We have developed a similar online version of this survey which has been advertised on a large poster in the Site A waiting area.
- To raise awareness among patients about providing feedback, we have made a handout to distribute to the patients via support staff from the blood pressure room.
- We will continue to have volunteers presented at least once per week to help patients/carers use the kiosk if needed.
- There are lots of typographical errors within the free text comments (see corresponding open ended results section of the report). We have contacted the company to ask that the free text box is expanded to more than one line to a) encourage more free text responses b) enable respondents to see what they have written and revise if needed.
- We will continue to provide a monthly summary to accompany the report at the end of each month until January 2018.
- We are monitoring the response rate of the electronic survey, paper survey, online and survey throughout this study to see what impact it has. The text survey is currently switched off during this study.

Appendix 8 Staff time and costs

TABLE 24 Staff and service user/carers time spent developing the toolkit/kiosk/data-mining tools

Activity	Time and resources used				
	Site A	Site C1	Site C2	Site B	Total
Co-design meetings and focus groups					
NHS staff: focus groups (number)	3	3	3	4	13
NHS staff: focus group participants (number)	23	27	35	43	128
NHS staff: focus group participants (total hours)	35	41	53	65	192
NHS staff: service user/carers trust lead meetings	0	0	0	3	3
NHS staff: individual interviews (number)					31
NHS staff: individual interviews (hours)					21
Service users and carers: PPI meetings (number)					7
Service users and carers: PPI participants (number)					35
Service users and carers: PPI participants (hours)					35
Service users and carers: PPG meetings (number)					10
Service users and carers: PPG participants (number)					50
Service users and carers: PPG participants (hours)					25
Service users and carers: focus groups (number)					4
Service users and carers: focus group participants (number)					20
Service users and carers: focus group participants (hours)					20
Service users and carers: interviews (number)					56
Service users and carers: interviews (hours)					56
Research staff (hours)					2869
Developing text-mining tools: research staff (hours)					1607
Interface design: research staff (hours)					116
Interface design: IT technician (hours)					394
Reporting templates: research staff (hours)					101
Information materials: research staff (hours)					60
Time spent is rounded to the nearest hour. Due to this rounding some totals may not correspond with the sum of the separate figures.					

TABLE 25 Costs of developing the toolkit/kiosk (2016–17 prices)

Activity	Cost of staff time (£)				
	Site A	Site C1	Site C2	Site B	Total
Co-design meetings and focus groups					
NHS staff: focus groups	1677	1254	1974	1820	6725
NHS staff: service user/carer trust lead meetings	0	0	0	60	60
NHS staff: individual interviews					728
Service users and carers: PPI meetings					700
Service users and carers: PPG meetings					250
Service users and carers: focus groups					200
Service users and carers: interviews					560
Research staff					84,831
Developing text-mining tools					43,745
Interface design					10,761
Reporting templates					3383
Information materials					1528
Total cost					167,615

TABLE 26 Staff time and resources associated with implementing the toolkit/kiosks

Activity	Time and resources used				
	Site A	Site C1	Site C2	Site B	Total
Kiosk rental	1	1	1	1	4
Induction: NHS staff (number)	7	3	3	10	23
Induction: NHS staff (hours)	4	2	2	15	22
Induction: research staff (number)	2	2	2	2	2
Induction: research staff (hours)	1	1	1	3	6
Induction: kiosk rental	1	1	1	1	4
Kiosk support: research staff (number)	NA				2
Kiosk support: research staff (hours)	NA				95
Kiosk support: volunteers (hours)	NA				33
Information materials: research staff (number)	NA				1
Information materials: research staff (hours)	NA				38
Video support: video production	0	0	0	1	1
Video support: NHS staff (number)	0	0	0	1	1
Video support: NHS staff (hours)	0	0	0	8	8
Video support: research staff (number)	0	0	0	1	1
Video support: research staff (hours)	0	0	0	8	8
Data analysis: software	NA				1
Data analysis: manual coding and report, research staff (hours)	NA				80
Data analysis: auto coding and report, research staff (hours)	NA				15
Data analysis: populating report templates, research staff (hours)	NA				28
Data analysis: report to sites and discussion, research staff (hours)	23	23	23	23	92

NA, site-by-site breakdown not available.

Time spent is rounded to the nearest hour. Due to this rounding some totals may not correspond with the sum of the separate figures.

TABLE 27 Total development and implementation costs per site by allocation of shared costs and length of annuitisation

Assumption	Costs (£)				
	Site A	Site C1	Site C2	Site B	Total cost
Implementation costs					
No shared implementation costs borne by local sites	806	708	708	839	5012
Shared implementation cost allocated equally to sites	5450	5352	5352	5483	23,589
Shared implementation cost allocated pro rata to sites	5440	4173	6160	5987	23,589
Development costs annuitised over 1 year					
No shared development costs borne by local sites	1677	1254	1974	1880	6786
Shared development cost allocated equally to sites	41,885	41,461	42,181	42,088	167,615
Shared development cost allocated pro rata to sites	41,793	31,247	49,171	46,449	167,615
Development costs annuitised over 2 years					
No shared development costs borne by local sites	853	638	1004	956	3451
Shared development cost allocated equally to sites	21,302	21,087	21,453	21,406	85,248
Shared development cost allocated pro rata to sites	21,256	15,892	25,008	23,624	85,248
Development costs annuitised over 3 years					
No shared development costs borne by local sites	578	432	681	648	2340
Shared development cost allocated equally to sites	14,444	14,298	14,546	14,514	57,802
Shared development cost allocated pro rata to sites	14,412	10,775	16,957	16,018	57,802
Development costs annuitised over 4 years					
No shared development costs borne by local sites	441	330	519	495	1785
Shared development cost allocated equally to sites	11,017	10,906	11,095	11,070	44,088
Shared development cost allocated pro rata to sites	10,993	8219	12,934	12,218	44,088
Development costs annuitised over 5 years					
No shared development costs borne by local sites	359	268	422	402	1452
Shared development cost allocated equally to sites	8962	8872	9026	9006	35,865
Shared development cost allocated pro rata to sites	8943	6686	10,521	9939	35,865

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