Automated analysis of free-text comments and dashboard representations in patient experience surveys: a multimethod co-design study

Carol Rivas,[1,2]* Daria Tkacz,[1] Laurence Antao,[1]
Emmanouil Mentzakis,[3] Margaret Gordon,[4]
Sydney Anstee[1] and Richard Giordano[1]

[1]Faculty of Health Sciences, University of Southampton, Southampton, UK
[2]Department of Social Science Research Unit, University College London, London, UK
[3]Economics within Social Sciences, University of Southampton, Southampton, UK
[4]PicoMeg Ltd, London, UK

*Corresponding author c.rivas@ucl.ac.uk

Declared competing interests of authors: none

Published July 2019
DOI: 10.3310/hsdr07230

Scientific summary

Free-text comments and dashboard representations in surveys
Health Services and Delivery Research 2019; Vol. 7: No. 23
DOI: 10.3310/hsdr07230

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Policy in the UK recognises the importance of the patient's perspective on the quality of care, both generally and for specific conditions. An example of a successful survey that aims to capture this perspective is the National Cancer Patient Experience Survey (CPES). Since 2010, this has been sent over a 3-month period each year to all patients treated for cancer, as inpatients or day cases, in NHS trusts in England. Currently, the psychometrically validated closed questions from such surveys provide both national and trust-/site-level summary statistics relating to performance. CPES closed-question data analyses have been linked to service improvement. However, currently there is no way of systematically, efficiently and usefully analysing and reporting the free-text responses in these surveys, despite this being a recognised need. The CPES generates > 70,000 comments each year and, accordingly, the conventional approach to this task, manual thematic analysis, can take months. Hence, the potential of free text to improve the patient experience depends on the willingness and capacity of staff. Template-based machine learning, otherwise known as memory-based or instance-based information retrieval (IR), has been tried as an alternative, but may be considered a semiautomated approach, in so far as a large number of the data still have to be analysed manually for themes that act as templates for the software to use. It is also a domain-non-specific, black-box approach, which has the potential to reduce accuracy. More consistent use of survey free-text data is asked for by patients and would provide insights into the closed-question responses, illustrating the processes and experiences that underpin them.

Aims and objectives

Rule-based IR, rather than template-based IR, was used to structure CPES free-text comments, writing rules based on natural language processing, such as syntactic structures, with the rules drawing on health-care domain-specific gazetteers of terms. This was explored for cancer as a specific condition, with the transferability to other health-care free-text feedback data and are conditions also explored. Rules and gazetteers are much more flexible and are easier to modify than templates for transferable use. The main aim was to optimise the use and usefulness of patient experience survey (PES) free-text comments in driving health service improvements that benefit the patient experience.

As well as aiming to produce rapid automated thematic analysis of large-volume survey free text, it was wished to display the results in a summary visual format in a digital dashboard display that could be drilled down to the original free text. The dashboard design was intended to drive service improvements, and so an interdisciplinary approach with co-design was used with all relevant stakeholders. This was important to maintain the patient voice in the data, while producing themes and displays that were meaningful to staff in terms of health-care improvement possibilities. The dashboard (developed into a toolkit), and the use of co-design with stakeholders, gives this approach added value over a simple thematic analysis, and the work was underpinned by Normalisation Process Theory (NPT), so that implementation and sustainability issues could begin to be explored.

Methods

The study was divided into three stages. The first stage of the study was a preliminary (or scoping) phase; this included a scoping review of clinical digital toolkit design, a stakeholder dashboard-scoping survey, a text-analysis term and theme mindmapping survey and a rapid review of themes. The surveys were disseminated through collaborator networks to determine what the different stakeholders in health care

(patients, carers, health-care providers, commissioners and policy-makers, and charities) thought should be incorporated into the work, from colloquialisms for the gazetteers to design requirements for the dashboard. The study aimed for 100 respondents, with a minimum of 15 required. Combining the results with those from a scoping review of clinical toolkit design and a rapid review of patient experience themes enabled the development of a working list of theme names and prototype dashboards. The rapid review included themes from previous manual thematic analysis of the 2013 Wales CPES (WCPES) data, which was also used to develop the approach.

The second stage of the study was the main development phase. The General Architecture for Text Engineering (GATE) (version 5.2.1; University of Sheffield, Sheffield, UK) Developer and GATE Embedded is open-source, 'text-engineering' (rule-based IR) software (originally developed by the University of Sheffield) and was modified using the working list of themes and half of the WCPES raw free-text data, refining the rules by testing and retesting on these data. This is different from the training used in template approaches, and was not a 'learning' phase, but a programming code refinement phase. Unlike the 'black-box' template approach, this enabled the exploration and refining of approaches to the analysis of fragmented sentences.

Stakeholders' views were explored on the themes and dashboard prototypes in group concept-mapping workshops, incorporating co-design work (aiming for 50–76 participants) and associated interviews (aiming for 15 participants).

The final stage of the study involved validation and use considerations for the approach. In addition to IR validation, costs and value-based preferences [using discrete choice experiment (DCE)], there was consideration, through walk-through techniques, of the usability and implementation of the dashboard. The rule-based IR work was validated by running the 50% of Welsh CPES data not used in the development work through the system and comparing the results with a manual analysis of the same data. Statistics on sensitivity, precision, accuracy and F-scores were considered. In the same way, to explore transferability without modification, two further data sets were validated against: one from patientopinion.org and one from a patient-reported outcome measure prostate cancer survey. The DCE aimed for 50 participants for a *D*-efficiency of around 85%, to determine the preferences for the different features in a health-care, toolkit. The final DCE design had 10 choice sets of three alternatives each and an opt-out. Within each choice set, only four attributes were allowed to vary across the three alternatives. Attributes were chosen on the basis of findings from the earlier stages of the study. Three versions of the design were created, with each respondent assigned to one of the three. In the walk-through, different types of NHS manager (aiming for 15), from three national NHS trusts (in Wessex, London and Leeds), were observed and recorded as they engaged in cognitive walk-through/heuristic evaluation of the toolkit to ensure that the toolkit was usable and could be put into practice. Questions targeting NPT constructs and the diffusion of innovations theory were included. Refinements of the toolkit followed this stage.

The key outcomes were rule-based IR accuracy, sensitivity, precision and F-score checks, documentation of transferability, a taxonomy of consensus-formed themes ranked by priority, a checklist of key features for digital clinical toolkits and ascertainment of the usability and goal-oriented support and marketability of the toolkit, with a view to exploring the further development and transferability of the approach.

Results

Preliminary phase (stage 1)

The scoping review of clinical digital toolkit design showed the importance of a clear definition of purpose and benefits on the homepage and a simple, short registration process. Continued engagement requires ease of access and integration with users' workflows.

In the dashboard-scoping survey, health-care professionals made it clear that they have limited time to spend on dashboards, despite expressing the need for the study's dashboard. They also desired access to the raw data; a combination on one computer screen of high-level overviews to highlight problem areas and benchmarks of regional and national performance; the potential to upload and incorporate their own quantitative data; a simple predictive tool; and good data reliability.

From the text-analysis term and theme mindmapping survey, participants' use of figures of speech suggested that patients were more focused on their daily experience and staff were more focused on cure. From this survey and the rapid review (including the previous WCPES work), 36 themes were identified. The most commonly described themes were transport-/travel-related issues; hospital environment; accessing the care system; clear information/communication between patients and staff; waiting for appointments/waiting on the day; co-ordinated versus fragmented care/communication between staff and staff, staff and institutions and institutions and institutions; and follow-up and aftercare.

Development phase (stage 2)

The study tried to accommodate, in the prototype dashboard design, all the features suggested in the preliminary phase, including individualisation features, information provided quickly and in a clear format and a range of filters to enable quick analytics. Theme icons could be clicked on to drill down to more detail, including original comments, enabling the rationale for higher-level directives to be understood at the clinical 'coalface' through detailed example, and with alert flags indicating data reliability. The prototype was discussed with the 34 participants in concept-mapping workshops (4–9 per group; modally, these were female participants, in their 50s, and they self-reported as being from a cross-section of ethnic groups); different stakeholders had different understandings of salient themes and dashboard usage. Health-care professionals' clearest conceptualisations concerned staff contact and hygiene, whereas for academics and other professionals these were staff attitude, communications (between patients and staff and between staff and staff), transport issues and hospital facilities, and for patients and carers these were waiting times (for appointments) and NHS organisational issues in general. Administration issues were blamed for most of the poor patient experience by the patient and carer group, so that, for them, these themes merged with other themes. One group comprised participants with multiple sclerosis (MS), who had different foci. Negotiation among participants resulted in 19 defined final themes used on the dashboard. Twenty-six participants rated themes by feasibility and 25 rated themes by importance; scores were similar for both, suggesting that participants conflated them, with a Pearson's correlation coefficient of 0.95. For the cancer-focused workshops, the top themes by importance for patients and carers were (1) legal and safety issues, (2) staff attitudes, (3) staff teamwork, (4) diagnosis and primary care issues, (5) funding and resources for the NHS and (6) facilities and environment of the hospital. For health-care professionals, the top themes by importance were (1) legal and safety issues, (2) staff training and skills, (3) facilities and environment of the hospital, (4) staff attitudes, (5) teamwork/communication and (6) including patients' family members in treatment and decision-making. For the MS-focused workshops, patient and carer ratings for the top themes by importance were (1) hospital cleanliness, (2) staff expertise and attitudes, (3) patients' consent and decision-making, (4) family support, (5) treatment choices and (6) hospital resources. For health-care professionals in the MS group, the top themes by importance were (1) staff expertise and attitudes, (2) treatment choices, (3) diagnosis, (4) availability of support, (5) human contact and empathy and (6) hospital cleanliness. This heterogeneity shows the importance of including theme choices in the dashboard.

Interviews with 12 workshop participants reached theme saturation. The workshops and interviews determined data-sharing issues, expectations around IR accuracy and survey sampling biases as critical topics for further debate and consideration before patient feedback data can be fully and optimally used.

Validation (stage 3)

In the DCE, the study sought to objectively validate the features of the dashboard/toolkit. The mean age of the 32 completer responders was 49 years, with 81% being female and 38% working for the NHS. Among these 38%, the most common band was band 7 (50%), and management (58%) was the most common professional area. All three models used pointed to similar ranking across attribute levels, with the search feature of a drop-down hospital list increasing the chances of an alternative being selected by

32 percentage points in the forced-choice model and by 28 percentage points in the nested logit (NL) model. The filter (by age, gender and condition) feature increased the chances by 32 percentage points and 25 percentage points, respectively. An increase of the annual fee from £250 to £1500 reduced the probability of purchase by about 18 percentage points in the forced-choice model and 10 percentage points in the NL model. A toolkit with minimal features would not be purchased. The full-featured option with the low-pricing strategy becomes profitable over 5 years only for 1000 potential clients (i.e. 890 purchases per year). For the full-featured option in the high-pricing strategy, the product becomes highly profitable for 500 potential clients (i.e. 403 purchases per year) and 1000 potential clients (i.e. 805 purchases per year).

Analysis accuracy was considered to be important to participants in the different stages. With the current approach, the following statistics were calculated for WCPES data comparing the rule-based IR with manual analysis of the same date: accuracy = 86%, precision = 88%, sensitivity = 96% and F-score = 92%. This is therefore close to human coding levels, taking into account human error, but still needs improvement. Automated performance with patientopinion.org and a patient-reported outcome measure prostate cancer survey was poor; this was a test to see if the system could give reasonable results with data sources other than the PES without any modification, which it cannot. However, the system has been designed to be easily modified for other data sources. Accuracy testing with modifications is continuing in collaborative work with the patient-reported outcome measure study and may be augmented with other types of IR.

The heuristic evaluation determined the need to refine the design in ways that were largely undertaken and are not reported here. All participants considered that the toolkit had the potential to help them to achieve goals in service improvement, justify requests for funding/designing new initiatives, gather data to show successes (and boost morale) and support appraisals, and to help them with reporting (although an export/report output function is needed). The majority of participants currently had no viable alternative, and they and their teams/organisations would probably support the use of such a system, and would see the purpose and value of it. More detailed implementation considerations were not possible to augment these generalisations, as governance issues within the NHS delayed the consideration of specific implementation models.

Patient and public involvement (PPI) was instrumental in the design and undertaking of the study. PPI representatives supported the study in the preparation of the public-facing materials and the group concept-mapping workshops. The group also played a key role in the recruitment of both patients and professionals for the workshops and interviews, and in designing the toolkit in such a way that the patient voice remained. The core values and principles for good PPI work were followed, as recognised by the PPI members in feedback.

Conclusions

The systematisation of patient experience free-text comments has been achieved in a way that should drive health-care improvements, with process transferability inbuilt. The evidence, in terms of the changes that were made to the taxonomy of themes, and the toolkit development, through the 'small study' co-design work and evaluation and validation stages, makes it clear that such novel interdisciplinary research is needed and has considerable benefits. The importance of this cannot be emphasised enough, as feedback about care provision does not lead to health-care improvements if suitable understandings, coherence and meaning-making are not achieved. Significant contributions have also been made to data sharing, IR accuracy and survey sampling debates.

Further research is needed to move from this proof-of-concept study to implementation. The first research priority is to explore the implementation and potential sustainability of the approach in practice for CPES data. This might be developed centrally with Insight NHS England, or locally within Wessex and then spread to other areas. Despite good process transferability, the rules for the rule-based IR need some further development for use in practice. Local adoption would enable careful evaluation using, for example, ethnography and implementation science methods to facilitate a national roll-out. One route forward could be to involve strategic clinical networks. More detailed cost–benefit assessments could be made at this stage using outcomes from this further research. Secondary priorities are to further explore the use of the approach with other health-care data sets to drive care improvements, and for the different settings of health care, research and teaching. More work is needed to develop the text-analysis approach, so that it performs adequately on other data.

The co-design processes used could also be further developed and refined to improve their usefulness. IR is a rapidly developing area, but for thematic analyses more sophisticated processes are not needed. However, a scoping review of what is being undertaken globally and its underlying theoretical frameworks might be a useful way forward, so that guidelines on good practice can be developed. This might also inform the refinement of the analytical approach for better transferability.

Funding

Funding for this study was provided by the Health Services and Delivery Research programme of the National Institute for Health Research.

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 or programmes which preceded the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HS&DR programme

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

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

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

This report

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 14/156/15. The contractual start date was in November 2015. The final report began editorial review in July 2017 and was accepted for publication in December 2017. 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.

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

NIHR Journals Library Editor-in-Chief

Professor Ken Stein 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 Honorary Professor, University of Manchester, and Senior Clinical Researcher and Associate Professor, 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