National Institute for Health Research

University of Southampton

Protocol

Study Title: PRESENT: Patient Reported Experience Survey Engineering of Natural Text: developing practical automated analysis and dashboard representations of cancer survey freetext answers

Principal Investigator: Dr Carol Rivas, Faculty of Health Sciences

Funder: NIHR HS&DR stream

Sponsor: University of Southampton

ABBREVIATIONS

CPES Cancer Patient Experience Survey

CQC Care Quality Commission

GCP Good Clinical Practice

NLP natural language processing

NPT Normalisation Process Theory

PPI Patient and public involvement

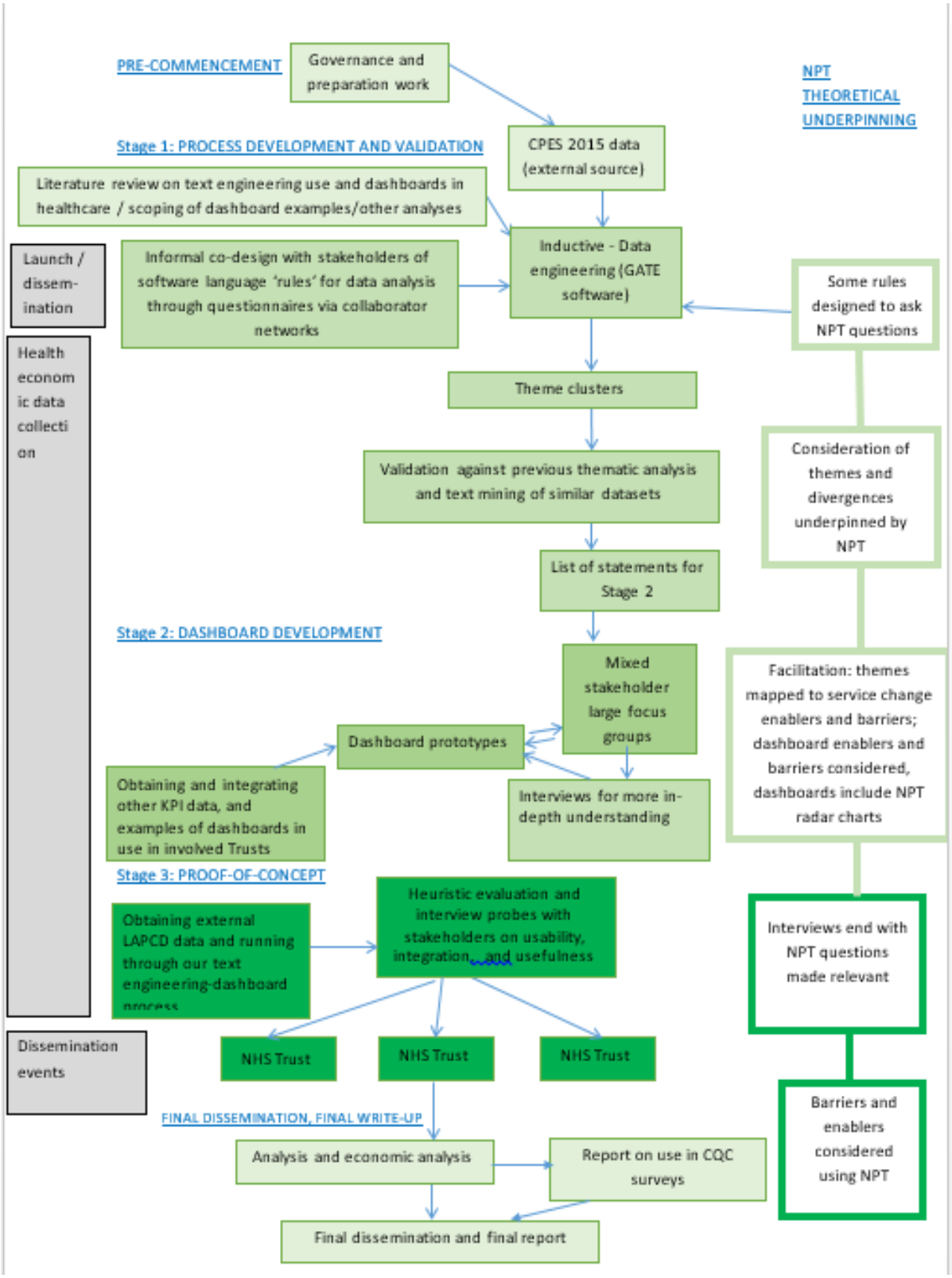
PREMs Patient reported experience measures

PES Patient Experience Survey

Contents

Contents

FIG 1: STUDY PLAN



1. BACKGROUND

1.1 PATIENT EXPERIENCE SURVEYS

Patient reported experience measures (PREMs) are recognised as important indicators of the quality of health service provision (1-6) and service improvement priorities. The Cancer Patient Experience Survey (CPES) Programme, begun in 2010, is widely acknowledged as the most successful national PREM survey in enabling and embedding service improvement (4,5). This has been achieved by providing Trusts with tailored (Trust-specific) statistical feedback of responses to the CPES closed questions, which benchmarks their performance against all other Trusts on a traffic light indicator system.

It is common practice for open-ended questions to be provided at the end of Patient Experience Surveys (PES) for respondents to leave freetext comments (1). CPES respondents, for example, are offered three such questions on their care experience; 1:3 respond (4). CPES is an example of a condition-specific PES. The Care Quality Commission (CQC) coordinates a number of site- or specialty- rather than condition-specific datasets and also uses these to improve healthcare services (3).

The CCQ surveys have similar comment boxes and also invite freetext comments via their website (http://www.cqc.org.uk/). We will be considering both condition-specific and non-specific surveys in our study but we will begin with CPES and then explore transferability to CQC surveys.

1.2 FREETEXT

Currently there is no system to efficiently and usefully analyse and report the freetext responses in PES (1). CPES generates over 70,000 such comments each year, accordingly the conventional approach of manual thematic analysis to make sense of these data (i.e. researchers reading through all the text and assigning topic or theme codes to each comment) can take months and so they cannot be linked to the traffic light system. Hence currently the potential of PES freetext to influence any improvement in patient experience depends on the willingness and capacity of staff to do such analyses. We have previously used text mining, a semi-automated approach to text analysis, as an alternative. But this still requires templates to be developed from manual analysis of a large proportion of the data, that the software can then use (or 'train' on) to analyse the remaining text by comparing it with the templates using decision trees (so called symbolic learning) (6). This means it is unable to handle the unexpected. Statistical word frequency approaches using Qualitative Data Analysis (QDA) software use only basic word searches which limits their use. More consistent use of survey freetext data is asked for by patients and would provide insights into the closed question responses, illustrating processes and experiences that underpin them. It will also enable us to improve health services and the patient experience.

1.3 TEXT ENGINEERING

We propose 'text engineering' as an approach that can address the issues and rapidly extract structured information from unstructured data such as freetext comments with minimal training. Text engineering has been used worldwide in biomedicine, to mine data from electronic medical records and to link these with online information sources (7), develop diagnoses for conditions and illnesses by categorising symptoms (8,9), predict protein structures and biological effects and cluster microarray gene expression data to identify gene groups with similar expression profiles (10,11). It has also been used to analyse comments made about commercial products, for example through analysis by our partner collaborator Nominet; however, it has not been used to analyse healthcare survey freetext data.

Like text mining, text engineering processes text using natural language processing (NLP) rules – the rules that we all use when we talk or write. But text engineering has the advantage of using a more human-like neural network (brain-like) learning approach (using interconnected theme nodes, dynamically weighted links and threshold logic units (12) rather than all or nothing as with text mining). This approach, based on probability (Bayesian statistics) makes it more rapid, sensitive and accurate than text mining and able to analyse the unexpected, and obviates the need to spend months developing templates.

1.4 AIMS AND OBJECTIVES

Our aim is therefore to develop and validate a novel use of 'text engineering' to provide rapid automated thematic analysis of large volume survey freetext (using the Cancer Patient Experience Survey (CPES) as our first case) and to also develop and validate a linked 'dashboard' that will display results in a summary format that can be drilled down to original freetext and can be used by patients and staff alike. The dashboard, and our use of co-design with stakeholders, gives our approach added value over simple thematic analysis in addition to the speed and automaticity that text engineering enables. With linkage to other service indicators and the use of Normalisation Process Theory (NPT) implementation science themes (13,14), providing further value, the display will illuminate service gaps and where patient experience can be practically improved at team, NHS Trust and national level.

As a secondary aim we will explore transferability to the general surveys of CQC. Manuals will ensure transferability to freetext in other health surveys more generally. This means that we will ensure that our text engineering process (that is, the topic-oriented automated freetext analysis) is quickly reproducible and that our dashboard will be designed to be used across topics, and easily modifiable to suit these.

Our objectives will be to integrate co-design and implementation science into our approach, output thematic analyses, validate our approach and ensure transferability. Our work will primarily ask:

1. Is our novel approach a valid, accurate way to analyse large volume CPES freetext responses?

2. Can it be transferred to PES on other topics?

3. Is co-design with mixed UK stakeholders (patients, their partners/carers, NHS managers, clinicians) feasible and effective for our approach?

4. Is Normalisation Process Theory (NPT) (13,14) useful for our approach?

2 METHOD

This is a technology development study, stage 1 in MRC guidance for complex interventions (15) and incorporating implementation science. It comprises three stages.

2.1 STAGE 1: CUSTOMISING TEXT ENGINEERING SOFTWARE

Our approach is underpinned by GATE open source software (12). GATE enables automatic translation ('engineering') of unstructured 'natural' text into structured data (i.e. clusters or 'themes'). Unlike simple word searches, it considers many complex and intersecting features of language to identify words, parts of words and phrases and themes, as well as (or instead of) the word itself. This means specific words can be accurately grouped together with their variations (synonyms, expressions, misspellings etc.). As an important feature, it is sensitive to the effect on meaning of context i.e. near neighbour words (known as co-locations).

While GATE does not require theme templates to process text, and comes ready to use, it is more accurate if purpose-built NLP rules, words and terms are programmed into its software and it is therefore designed to enable this. We will use stakeholder input to work out new rules, words and terms that are specific for healthcare provision, particularly but not exclusively cancer care.

We will manualise what we do so that our method can be repeated by users in other areas of healthcare should they wish to further adapt the software.

To see whether our process works sufficiently well to be used as an alternative to existing methods, we will compare its outputs against those from our previous manual thematic analysis and text mining analysis (17,18).

2.2 STAGE 2: DEVELOPMENT OF AN ASSOCIATED DASHBOARD

Once the basic text engineering process has been validated we will invite stakeholders (patients, carers, service providers, NHS managers, commissioners) to co-design a web based dashboard that can extract, store and display the outputs from stage 1 in a way that can inform service improvement (we will include NPT radar plots to help achieve this). The dashboard will be easily modifiable for other topics. This is a key part of our approach and will be

undertaken using mixed stakeholder concept mapping workshops (also known as 'large focus groups'), which are designed to get different stakeholders to reach consensus through discussion together and negotiation. Some stakeholders will also be invited to interview to further explore some of the discussions in the workshops.

2.3 STAGE 3: PROOF OF CONCEPT PHASE

This stage will check whether we have achieved our aims. We will use data from our 2015/16 Cancer Research UK/Movember-funded prostate cancer survey (with data ready in time for our study) to test our new approach. We will also observe and record different types of NHS manager (n=15) in 3 NHS localities nationally (Wessex, London, and one other chosen for its engagement in service improvement) as they engage in a Structured Walkthrough (19-23) (videoed) of our dashboard to ensure it is usable and can be put into practice. Specifically the walkthrough will involve: (i) Ascertaining the extent to which the dashboard conforms to standard usability principles; (ii) Testing that role relevant tasks can be accomplished by those who use the dashboard; and (iii) Ascertaining the extent to which the dashboard engenders new work practices and processes aimed at improving care, keeping this in mind.

3 MATERIALS

In stage 1 we will use a questionnaire, as attached.

In stage 2 the concept mapping groups will discuss the themes from stage 1, and also a prototype dashboard (which we will develop as part of the study and then refine according to stage 2 feedback). Thus there are no measurement tools, questionnaires, etc. that can be developed in advance for stage 2 group work. The interviews will be semi-structured to allow free flowing talk but they will only consider the discussion that the participant took part in when in the concept mapping group, thus the topic guide for these can only be very loosely defined, as attached.

In stage 3 we will seek answers to a set of questions that are commonly used in heuristic evaluation and cognitive walkthrough, as attached.

4 PARTICIPANTS

FIG 2: RECRUITMENT TO THE DIFFERENT STAGES

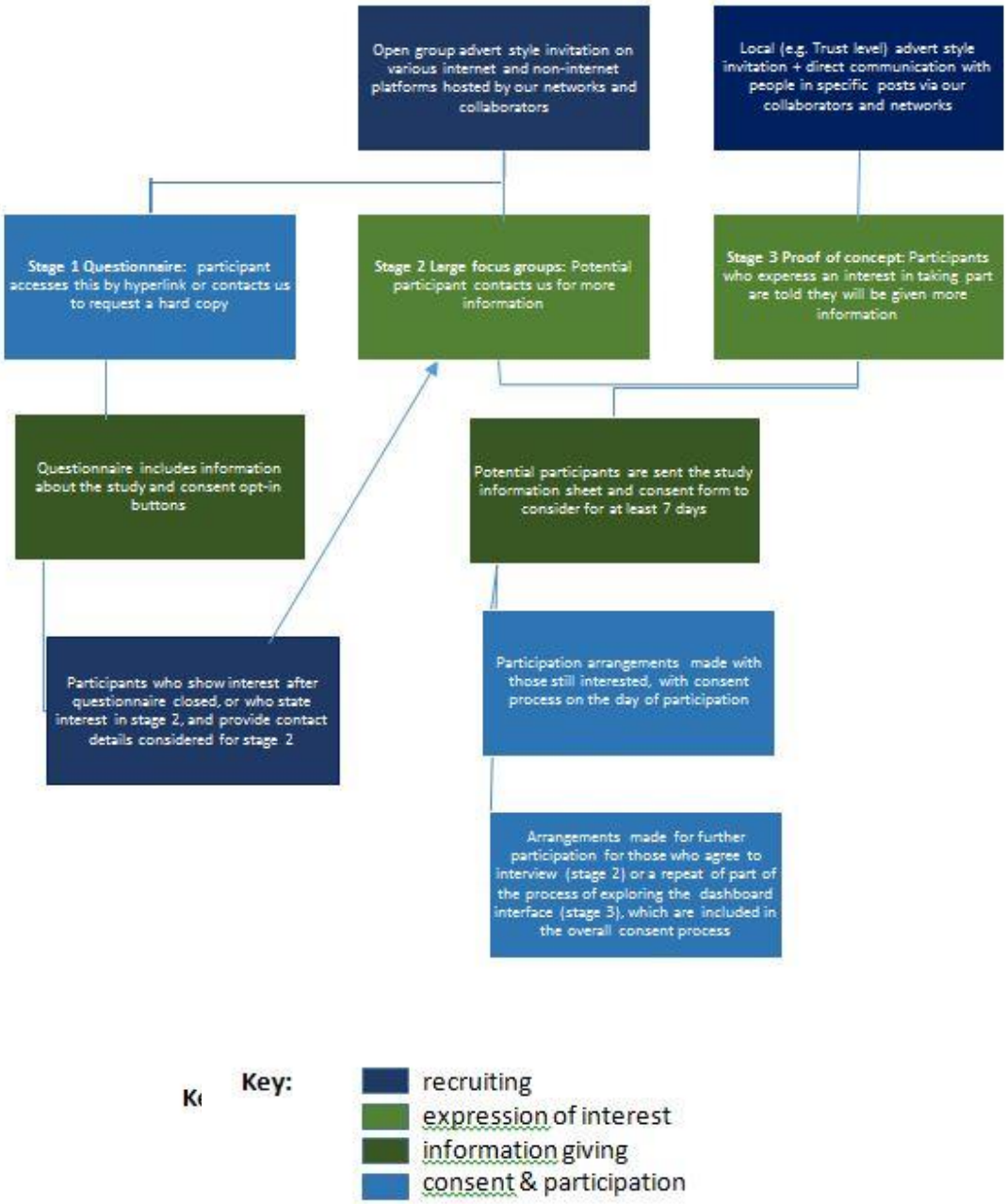


TABLE 1: SUMMARY OF INVOLVEMENT OF PARTICIPANTS

BROAD PARTICIPANT GROUP[a]	STAGE 1 QUESTIONNAIRE NUMBERS	STAGE 2 LARGE FOCUS GROUP NUMBERS	STAGE 2 INTERVIEWS NUMBERS	STAGE 3 PROOF OF CONCEPT NUMBERS
Patient with cancer	20-30	15-25[b]	4[c]	0
Patient previously a surgical inpatient	20-30	15-25[b]	4[c]	0
Carer of a patient with cancer	20-30	15-25[b]	4[c]	0
Carer of a surgical inpatient	20-30	15-25[b]	4[c]	0
Professional involved with healthcare delivery	20-30	15-25[b]	4[c]	15[d]

[a] No participant is necessarily expected to be connected to any other, though carers may come with patients they care for.
[b] May include participants from the previous stage
[c] Must include participants from the previous column (large focus groups)
[d] None from other stages

4.1 STAGE 1

4.1.1 Participants

Inclusion criteria

- For stage 1 and also stage 2, participants will be staff, patients and carers from a specific set of networks and groups
- They should have confirmed diagnosis in cancer or fit the surgical inpatient profile of the Care Quality Commission surveys if patients or if carers, in those they care for
- They must be 18 years or over
- They should have acknowledged that they read the information form (two consent boxes in the questionnaire must be ticked before the participant can proceed).

Exclusion criteria:

- Not fitting any of the inclusion criteria.

We are not excluding private (non-NHS) patient/carer of patient/staff at this stage (though we are at subsequent stages) as we wish to have varied input. But

we will ask whether they are NHS or not as we wish the prioritisation question to be considered for the NHS specifically.

All inclusion and exclusion criteria will be determined through a question at the start of the questionnaire, and if exclusion criteria are met the participant will be politely thanked and told that the current survey is not relevant for them.

We are not excluding people on the basis of language though since the questionnaire is in English we do not expect much non-English language response. The PI is multilingual and will judge any non-English contributions individually, enlisting the help of others for languages she does not understand.

4.1.2 Sites

These will be the various platforms hosted by our networks and collaborators, including Macmillan Cancer Support, Insight and Feedback, NHS England (which hosts CPES) and Healthtalk.org.

4.1.3 Numbers

We will aim for maximum variation in our sample, across cancer types and stage, gender, age, staff role, aiming for 2-4 participants in each sampling group. We thus expect to recruit 100-150 participants or until saturation of themes for each group.

4.1.4 Approaching and recruiting participants

We will not contact participants directly but invite them to fill in our questionnaire via the relevant networks of our collaborators, for a broad UK-wide but targeted approach. This will include internet based and non-internet platforms. Our launch event will bring together UK Trust managers who may provide links to relevant groups additional to those accessed via our collaborators.

We will cascade data requests to ensure we are not overwhelmed; each wave will select for a range of stakeholders and when ideas become saturated we will stop. We do not need large numbers for our crowd-sourcing process which is a modification of an approach already used in GATE (15 participants are sufficient) according to GATE developers).

Our extensive networks and small numbers required and our purposive approach may mean we may not involve all those people who express an interest. We will be sensitive to this, and will thank everyone for their interest and ask if those not included wish to have their contact details kept in case of further similar research.

4.2 STAGE 2

4.2.1 Participants

For stage 2 we will aim for approximately 15 participants per focus group (with 5 groups) from each of the following broad groups: cancer and non-cancer patient and carer, relevant nurse and consultant, any of the 5 groups in stage 3 below.

Inclusion criteria

- Staff, patients and carers from a specific set of networks and groups.
- Involved in cancer or surgical inpatients within the NHS setting in some integral way
- Have a confirmed diagnosis in cancer or fit the surgical inpatient profile of the Care Quality Commission surveys if patients or if carers, in those they care for, to ensure they have sufficient experience of the relevant services
- 18 years or over
- Able to provide full informed consent
- Residing in the UK and have easy travel to the institution where the groups will take place, and same day return.

Participants will be chosen purposively to represent a cross-section of stakeholders, hence we are holding 5 groups. That is, they will be selected according to role in the cancer healthcare process (service providers, policy makers, budget holders and commissioners, then sampled on a convenience basis within role) or if patient/carer, demographic, health and healthcare-related features (gender, age, cancer stage and type and treatment type), as in stage 1. Partners/carers must be caring for/a partner of someone with cancer (not necessarily a patient participant); we define caring as supporting a patient with their daily functioning for at least 20 hours a week.

Exclusion criteria:

- Not able to travel to the groups
- Not fitting any of the other inclusion criteria
- Private (non-NHS) patient/carer of patient/staff.

We are not excluding people on the basis of English language fluency because if sufficient without fluency respond we will hold a special group for them. However 5 of the groups will be held in English.

4.2.2 Sites

Participants will be invited from across the nation but they will need to be able to attend at one of the three insititutions we use as venues for this stage.

These groups will need to take place at a university, training centre, conference centre or similar setting in order to ensure that the appropriate equipment is in place, such as a large monitor to display the concept maps. We will convene the

groups in three locations across the UK, including London and Wessex, and possibly also Leeds, in order to ensure participation across the country.

4.2.3 Numbers

Each large focus group will involve 2-3 participants from each participant group and there will be five focus groups altogether. Each group will therefore comprise 15-25 participants.

4.2.4 Approaching and recruiting participants

Stakeholder participants in the large focus groups will be recruited through the same networks and groups as stage 1 and in the same way except that they will be required to provide us with contact details if interested, rather than filling in a questionnaire. We can then send them the participant information sheet and consent form and follow up to see if they remain interested.

4.3 STAGE 3

For the Structured Walkthrough (heuristic evaluation and cognitive walkthrough with consideration of NPT radar plots included), our choice of sites and participants is shaped by our intention to examine the value and design principles for the dashboard, so that it can have impact as a service improvement tool. To achieve this, and using our extensive and established links we will work with three commissioning systems (localities), from the commissioner outwards to an NHS Trust and local GP services.

4.3.1 Participants

Inclusion criteria

For stage 3 we will recruit from among the following five groups:

- Lead members of multidisciplinary treatment teams (MDT)
- Ward managers drawn from settings where a substantial number of cancer patients have surgery and from a general ward (to consider transferability)
- Associate directors, lead clinicians for clinical groups that include cancer (medical clinician leads as well as nursing and or general managers), Members of the Trust Board
- Lead commissioners in CCGs/Health and Wellbeing Board members
- GP practice leads.

Exclusion criteria

- Not fitting any of the inclusion criteria
- Private (non-NHS) only staff
- Unable to take full informed consent
- Having taken part in stage 1 or 2

- Anyone with an Impediment that makes it impossible to use a dashboard interface.

4.3.2 Sites

We have selected Wessex as one of three localities for this, as this is where we are based, and it has a developed oncology service. A second locality is London because of its complexity and challenges around the Patient Experience; 8 out of 10 of poorest performing NHS Trusts in 2014 (CPES) were from London. Given that NHS England/ NHS IQ has begun a peer to peer buddying initiative for service improvement, with Trusts performing well on CPES being matched with those who are not, we may select a different locality from this initiative that has a record of using CPES to improvement patient experience, to maximise variation and optimise benefits from our evaluation exercise. Our three NHS localities are selected for geographic and socio-demographic variation as well as cancer services.

4.3.3 Numbers

We have chosen 15 participants for the walkthrough since the number of issues found by reviewers undertaking such evaluations shows an incremental rise between 1 and 5 evaluators, with 5 finding 75% of potential issues and diminishing returns from adding further reviewers (24). We will also explore use and integration into practice for several different role types.

As part of the Structured Walkthrough we will determine the relevant knowledge and skillset of the evaluators as well as noting their roles and reasons for potentially using our process. Here we will draw on Diffusion of Innovations Theory and network contagion theory as well as NPT.

4.3.4 Approaching and recruiting participants

Through our networks and links we will make personal contact with potential participants. We will approach people in our own networks directly (on a no-pressure basis) and we will ask our collaborators to request permission from other potential participants to share with us their contact details so we can discuss participation with them. We will also place local announcements e.g. at Trust level. We will recruit purposively from three specific localities.

Participant recruitment will cease after 200 participants or when no new information emerges across the stakeholder groups, whichever is first. Participants who respond once we have closed data collection and before our questionnaire link leads to a ‘Thank you but this survey is now closed’ style of response, and those who contact us despite having seen the 'closed' announcement, will be offered the chance to take part in the next stage or for their contact details to be kept for further research possibilities. This is mainly to avoid their disappointment and to be sensitive to their wish to be involved in research as we do not envisage recruitment problems at any stage of our research.

5 WITHDRAWALS OF PARTICIPANTS

Participants may end their involvement at any time. We will advise them that we plan to keep all data collected from them before they end their involvement, unless they specify otherwise. However, if they require it, we will destroy all their data, which will be done securely as described in the data management protocol. If data have been used in disseminations before withdrawal we will advise participants that this information cannot be withdrawn.

We can only destroy questionnaire data from participants who give us their participant ID as we will otherwise not be able to identify their data.

Participants must tell us by September 2017 if they wish to avoid their data being included in reports and presentations and educational and research materials. This is for practical reasons as we expect to have published some findings by then. After this time, we can still remove data from our archives.

6 PROCEDURE

6.1 STAGE 1

Our collaborators will place notices about the study and invitations to participate on relevant websites, in newsletters and other network platforms. People who access these will be invited to complete a questionnaire by clicking on a web link or requesting hard copy. Questionnaires can be completed at the participants' convenience and wherever they choose. We will ask one of three types of question, building on our previous CPES theme development. (See questionnaire.)

Responses, in the form of ideas, concepts, words and phrases, will be collated by the main researcher and analysed by hand under the supervision of the PI and with reference to the advisory group and patient advisory group. The programmer will use those that are useful for the GATE process, and modify this.

Once the GATE process has been modified, we will use it to run an analysis of CPES 2015 data. The results will be stored in .csv files so that they can be read by other software including that underpinning our dashboard.

We will view the CPES analysis results and develop 100 statements from them, such as the names of the themes themselves. These will be used in stage 2.

Simultaneously we will develop a prototype dashboard from a scoping review of the literature and sampling of preferred dashboards within our networks, undertaken before the study begins.

6.2 STAGE 2

Our collaborators will disseminate stage 2 invitations to participate in the same way they did for stage 1, with participant opt-in as in stage 1. We will convene 5

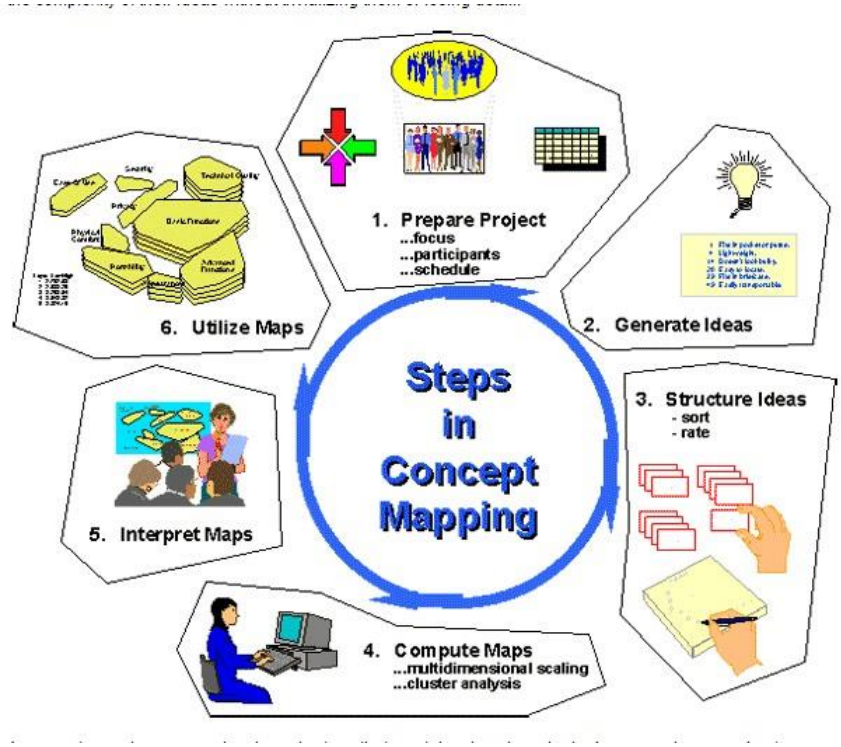
large focus groups (also known as consensus groups or structured group concept mapping workshops). Each will involve 15-25 individuals from a range of stakeholder backgrounds with aim of reconciling differences in opinion through negotiation, facilitated by a trained researcher. Each participant will only attend one large focus group, which will last for a maximum of six hours (probably 10-4pm). Each group will be attended by at least two researchers trained in the process; one will facilitate the discussions, the other will input data into the software as it is generated on the day. One of the researchers will probably be the PI, who has used this method before. Concept Systems Inc. as software suppliers will provide support.

6.2.1 Large focus group process

Early piloting will show whether we need to involve more researchers on the day, to keep pace with the information flow. Small focus groups run beforehand will enable pre-testing of the software and process.

FIG 2: STEPS IN CONCEPT MAPPING

(http://www.socialresearchmethods.net/kb/conmap.php). Stage 1 of the study represents steps 1 and 2 in the figure.



6.2.1.1 Pre-task

We will ensure all participants are at ease before starting the formal session, partly by the provision of warm up tasks, as well as ensuring the environment is appropriate and non-threatening. For example participants may be asked to tell the person next to them about something unusual about themselves, then this

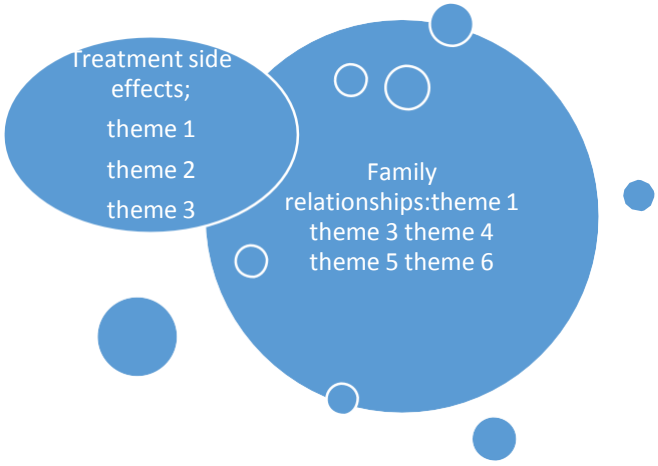
second person has to tell the unusual thing to the group; this process is reciprocal.

6.2.1.2 Consensus negotiation

The first formal task of the group is for each member to develop an individual coding frame in which they group the 100 theme statements that we give them from the stage 1 text engineering process. These are not the individual freetext comments but semantic theme clusters or representations of these, produced by the stage 1 analysis. This first task, in which individuals do not reveal their own codes to others, enables all voices to be represented in the outputs of stage 2. The facilitator/researcher types the codes from each group member (which are written down, and not spoken) into special software which then compares all the participants' coding frames in real time using cluster analysis (a type of Bayesian maths that groups data into clusters according to levels of similarity and is similar to the maths that underlies text engineering). The software then produces cluster maps that represent the group's data, and graphically show levels of consensus for concepts, and the relation of the concepts to each other. With relationships determined mathematically, the process can occur in real time. Each person is anonymised in the maps and the group members can see the maps on a large screen. This facilitates discussions on difference, and consensus forming.

In the second task the groups will discuss divergences and similarities to develop a group consensus and amend previous concepts as appropriate. For example sexual dysfunction may be highlighted as important as a side effect of treatment, a relationship issue or a body change with age – that is in three clusters, not one - depending on perspective and these divergences need to be reconciled, which the second task seeks to do. Discussions might consider whether two related concepts should be merged or should be described differently to make their difference formal and clear.

In the following stylised version for example, participants have coded some theme statements under 'treatment side effects' and some under 'family relationships'. There is some overlap so the clusters are closer together. The size of the clusters shows how many statements have been included within them. Some smaller clusters are contained within the larger one so discussion would consider whether the larger one needs to be broken down – and with three small clusters close together this might be useful – or whether the smaller clusters need to be included within the larger one rather than being distinct.



With this approach, in a single session, input can be efficiently integrated from multiple groups of informants with differing content expertise, interest, beliefs, understandings or experiences, the maps can be discussed and refined by the groups (for example naming or rejecting clusters), clarification sought on points of agreement and disagreement, and recommendations and models formulated from the maps and discussed. These then have the benefit of being rigorously developed from a composite summary grounded in what participants decide.

6.2.1.3 Ratings of action points

Participants can rate feasibility of action points for service improvement developed from the most significant concepts. Importance ratings will show dis/agreement about the significance of themes, thereby prioritising them. Participants will individually and simultaneously make ratings using mini hand-held polling devices connected via an audience/group voting system. Standard host organisation technical support will ensure this works correctly on the day. Group scores will be fed back for discussion, which will begin by focussing on those with high disagreement and then votes re-cast and checked with the group.

Sessions will be audio recorded to maximise their usefulness. Recordings of discussions will be analysed deductively to explore points of convergence and divergence to inform interview topics, interview participant selection (from among large focus group attendees) and process design for refinement or for a future trial. They will also be available to the team for further analysis to augment the main project.

6.2.1.4 Dashboard development: second half of the day

During the second half of the large focus groups day, discussions will consider design of the dashboard that we will use to present themes from the GATE text analysis including the inclusion of other service indicators, and the potential of the dashboard for use in routine work to support improvement efforts. Before the large focus groups we will have developed prototype dashboards from examples in current use in NHS Trusts as collected in a scoping exercise before the study begins. These will facilitate discussion.

6.2.2 Interviews

Interviews will take place within two weeks of the large focus groups where possible. They will be short, probably a half hour, involving only some (15-20) participants, and designed to explore topics where agreement was not achieved or was hard to achieve. They will not ask questions beyond this. They will be undertaken by the main researcher or the PI.

6.3 STAGE 3

We will ask staff in 3 UK NHS Trusts (including London, Wessex and a third site to be determined from CPES 2015 data as we wish to ensure we have poorly performing and better performing Trusts involved) to talk about what they do as they try out our dashboard (a common technique in technology development). We will ask these staff if they would have issues using our approach at work and discuss solutions. These cognitive walkthroughs are expected to take place where the participants work, though they can elect for them to take place elsewhere. They are likely to take approximately two hours per person. At least two researchers will be present, a facilitator and a scribe (who may also operate the video). The feedback is expected to lead to refinements of the dashboard, in which case the participants will be asked to repeat the process and comment on whether we have successfully dealt with any issues they raised the first time. The questions we will use to structure the process are provided as an attachment.

7 STATISTICAL ANALYSIS

There is no real use of descriptive statistics in this study. Data from the text engineering process will be collected in .csv files (which can be imported into Excel and other programmes) using a software routine that has already been written for this project. Therefore statistical analyses can be applied to the data as relevant, but none are planned. The main purpose of using the .csv files is to enable the data to be imported into the dashboard.

For validation of our process in stage 1 we will compare the theme outputs with text mining and with the gold standard way of analysing freetext data, i.e. manual thematic content analysis. To ensure orientation to the same goal, which is to produce a representation of the data intended to drive changes in practice, we have chosen to use our previous study, for which we have already determined sensitivity and specificity (17). The statistical analysis (sensitivity and specificity) will follow simple formulae used routinely in GATE and text mining (6,12,25).

Text mining algorithm performance is measured as sensitivity or recall:

> true positives/(true positives + false negatives)), specificity or precision (true negatives/(true negatives + false positives)

and by the f-score:

(2*sensitivity*precision)/(sensitivity + precision).

Sensitivity describes the ability to identify all relevant comments of a given category, precision defines the ability to exclude non-relevant comments. The f-score describes the overall performance, which represents the harmonic mean of precision and sensitivity.

For the text engineering, we will use the process built into GATE.

In the large focus groups the software we will use enables rapid multivariate statistical methods of multidimensional scaling and hierarchical cluster analysis to produce almost instantaneous concept (cluster) maps from the coding frames. They will provide a statistical and visual summary of concepts, clustering them by significance and likelihood of interdependence, convergence and divergence. The software produces the analyses including all statistical summaries; we only need to input data in real time on the day of each group. As we hold more than one group, the research team will combine the data across groups taking the statistical summaries into account, with input from Mary Kane who co-developed the method. She has already provided support in the drafting of the research design and protocol. Combining the data in this way requires each large focus group to consider the same data in their first formal task.

In stage 3 there is no use of statistics.

8 HEALTH ECONOMICS

There will also be a simple cost-benefit health economic analysis. All significant resource consumption that is expected to differ between analysis options (text mining, manual thematic analysis, our new process) will be estimated within the study regardless of who eventually incurs the cost; that is, the study will take a societal perspective on costs. The results of this calculation will be taken to be the benefit; though service improvements contingent on timely reporting that our analysis enables should lead to substantial cost saving these are beyond the remit of the current health economics analysis.

Resource use during development of our text engineering analysis will be measured to be used in manuals of the process. Modifications to the dashboard for different surveys will be estimated using tallies of costs incurred at specific stages of the dashboard development.

Ongoing resource use for the dashboard (e.g. site maintenance) will be calculated.

Resources will be valued in monetary terms using published or standard unit costs at the time of analysis. Sensitivity analysis will be performed.

Our study might lead to the potential for this type of analysis to be taken in house by the NHS England Insight team and CQC or drastically reduce the cost

when using commercial contractors. This is something we would investigate as part of the economic analysis.

In a future trial, building on this work, willingness to pay and discrete choice experiment (DCE) would be used for outcomes.

9 ETHICAL PRACTICES

9.1 OVERALL

All staff will share the same duty of care to prevent unauthorised disclosure of personal information and to otherwise act according to the Data Protection Act 1998 and good clinical governance. The protocol and data management plan must be followed at all times. Action will be taken in the event of issues (see project management for details).

9.2 RECRUITMENT OVERALL

We will advise all networks and collaborators to ensure participants do not feel pressured to take part.

9.3 STAGE 1

We are analysing public datasets using our modified GATE approach. Although anonymised, information within raw data the comments could identify individuals. Every care will be taken to ensure such data is not put in the dashboard, report or subsequent papers.

9.4 STAGE 2

9.4.1 Enrolling participants

Our extensive networks and small numbers required and our purposive approach may mean we may not involve all those people who express an interest. We will be sensitive to this, and will thank everyone for their interest and ask if those not included wish to have their contact details kept in case of further similar research (this is in the consent form).

9.4.2 During the mixed group work

In the large focus groups, service providers may feel ill at ease discussing things in the presence of service users, or they may feel criticised, and service users may feel disempowered or awkward or reticent or worried about making negative comments, in the presence of the service providers.

We will ensure all participants are at ease before starting the formal session, partly by the provision of warm up tasks, as described in section 6.2, as well as ensuring the environment is appropriate and non-threatening. We will also emphasise that the groups are a 'safe space' and intended to be non-judgmental.

The facilitators will be trained in the processes involved and will ensure all voices are heard and that any issues are dealt with diplomatically and constuctively.

The large focus groups may include discussion of sensitive issues as in any focus group. In particular, as cancer and its treatment will be considered in this study, and also inpatient surgery, some participants or researchers may feel psychological discomfort or stress. We will ensure questions will not be emotive and discussions will not be personalised but will only involve conceptualising services and discussing pre-existing research themes.

We will be alert to any distress and provide reassurance and invite any distressed participant to withdraw from the session or study, in a discrete way. Should any participants continue to feel distressed after reassurance, they will be offered the contact details of a professional who may help them; researchers will carry a list of such contact details. A main contact will be provided in the participant information sheet which the participant should be reminded of (the researcher should carry spare copies of this).

All participants will be asked to respect confidentiality requirements. This must be emphasised at the start and end of the large focus groups.

9.4.3 Use of the dashboard

It is important that service users and potential service users can access the data in the dashboard we produce. Some levels may be restricted however to healthcare professionals and to specific staff within these; selective access will be explored in the large focus groups themselves.

9.4.4 Death or illness progression

Some patients may die or become too ill to take part, between asking for more information and receiving it or between receiving it and taking part. Staff will be briefed beforehand on how to deal with this, using department resources in the first instance. In practice this is likely to be extremely rare; in other qualitative research with a much larger sample of cancer patients there has been an incidence of approximately 3 per 1500.

9.5 STAGE 3

Healthcare professionals in the Structured Walkthrough may worry they are being monitored. We will reassure them that they will not be monitored or assessed and that the nature of the heuristic evaluation is non-judgmental.

10 RISKS FOR RESEARCHERS

Researchers may find themselves overwhelmed by reading/hearing about illness or healthcare even though emotive language and personal experience information will be minimised. Thus the team will have regular debriefing sessions together, timetabled in to coincide with the active data collection periods and as needed at other times.

Researchers may need to deal with participants who feel distressed whilst taking part in the group discussions and interviews. However, they will be briefed about how to deal with this and must carry details and contact numbers for suitable support (though some contacts will be on the information sheets, spare copies of which need to be carried by the researchers when collecting data).

Technical support will be available at the institutions where data collection will take place and researchers should acquaint themselves with the relevant contact details as part of their preparation for data collection.

11 DATA CHECKS AND FILE NAMING

Participant consent forms and eligibility criteria need to be checked before their participation or if this is not possible, as soon as practical after their data are collected, and before data are cleaned (for anonymity) or stored or transferred or recordings transcribed.

Checks will be made of survey outputs as well as original research data, for personal or identifying information. All such information will be removed from all data as part of the cleaning process.

File naming conventions are described in the data management protocol. In particular they will include the letter x if consent is restricted in any way and the researcher will then need to consult a special table to find out the details.

Staff need to check that filenames are correctly assigned, and that the correct file format is used, before storing data in study folders during the study or in longer term storage after.

12 DATA PROTECTION, CONFIDENTIALITY AND ANONYMITY

Further details are available in our data management plan.

12.1 RELEVANT GUIDANCE AND LAW

All research and personal data will be stored and managed in accordance with the Data Protection Act ***1998*** and the Research Governance Framework for Health and Social Care. Our use of the survey data is contingent on our conforming to the Information Standards Board for Health and Social Care (ISB) Anonymisation Standard for the Publication of Health and Social Care Data (33).

12.2 OVERALL STRATEGY

Overall, confidentiality will be maintained through the anonymisation of data, the secure storage of personal data and the explanation of the importance of maintaining confidentiality to all participants in the consent documentation including the consent form, and at the time that the large focus group data are

collected also by spoken emphasis of the need to respect each other's' confidentiality.

When using direct quotations in publications, participants will be identified as Patient, Partner, Carer, or by their relevant role, for example: Manager, Consultant, Cancer Specialist Nurse etc. or by a number or code that does not reveal their identity. Their actual names will never be used or linked to the data.

12.3 RAW DATA

Raw data that the research team collect will not be shared with others by them during the lifetime of the study, although selected short anonymous excerpts may be disseminated at conferences and in publications and study outputs, as made clear on the study consent forms. Considerations of retention and disposal of such data at the end of the study are considered in the data management plan.

12.4 PERSONAL DATA

All linked data will be stored with its own unique identifying code and not with the participant's name and the corresponding linkage details, which will be stored separately. All identifying 'metadata' (names of people and places and other similar information) will be deleted from recordings and transcripts where possible (**see section 10**). Transcript texts will be saved as text files when this has been done, to remove hidden code that could enable original information to be recovered, and will then be reformatted.

No personal data will be shared outside the core research group at any time and nor will data that might be considered potentially sensitive in any way. Video and audio data will not be shared. ***Video recordings will anyway not include participants' faces but the interface between the body and the computer ie their hands.*** Shared data will have been 'cleaned' of sensitive or identifying material.

12.4.1 Publishing of stage 2 research

Since we only plan to recruit two to three people from each sampling group per large focus group in stage 2, we will take extra care in publications not to potentially remove anonymity.

12.4.2 Stage 3 anonymity and dissemination

In the heuristic evaluation our interest is in user role rather than named identity and we will not record names alongside data or link names to the data in any way. Nor will we link individuals to a specific Trust when we disseminate. However it would not be too hard to work out some identities, such as at the top level and we will be aware of this when reporting results and will take care to avoid this possibility as much as possible, for example by developing a 'case study' or summating data. Nonetheless this should not be an issue as this stage examines the technology specifically. We will take care not to disseminate anything that could raise issues for individuals.

Audio and video recordings will be destroyed securely – and mostly audios will be destroyed before the end of the study, after transcriptions have been checked; videos will be destroyed at the end of the study.

12.4.3 The repository

When data are archived in a repository at the end of the study, the research data will be reviewed to consider the value of storing all or part of the raw data and whether access to any of it should be restricted. This is explored further in the data management protocol.

The data sharing policy of the NIHR Journals Library states that:

"Data generated through participation of patients and the public should be put to maximum use by the research community and, whenever possible, translated to deliver patient benefit. Data sharing benefits numerous research-related activities: reproducing analyses; testing secondary hypotheses; developing and evaluating novel statistical methods; teaching; aiding design of future trials; meta-analyses; and helping to prevent error, fraud and selective reporting."

12.4.4 Access to our survey analysis output via the dashboard

Access to the CPES 2015 data analysis produced using our approach will not be restricted except that users will be required to register on our dashboard and consent to our copyright and user etiquette requirements. All data will be anonymised and non-identifiable.

12.4.5 Consultants and collaborators

All consultants and collaborators to the study will sign agreements with confidentiality clauses.

13 PROJECT MANAGEMENT

13.1 OVERALL STAFF OBLIGATIONS

The PI will be responsible for conduct of the study, day-to-day management and decision-making. However all staff will share the same duty of care to prevent unauthorised disclosure of personal information and to otherwise act according to the Data Protection Act 1998 and good clinical governance. The PI cannot be held responsible for breaches by others that are not within her reasonable control. Appropriate action will be taken should any breaches or potential for breaches occur; this might include anything from retraining to dismissal depending on the seriousness. Staff must therefore take care to follow the data management plan and protocol at all times and advise the PI immediately of any potential issues or any need to amend either protocol or data management document. Staff should try to attend all meetings that they are invited to.

14 PATIENT AND PUBLIC INVOLVEMENT

As we propose to analyse patient experience freetext data, PPI is vital to ensure we remain true to the purposes of the data and produce results useful in practice to improve the service experience. This is in addition to our belief in the importance of PPI to ensure that the work we do is relevant to service users, who are at the heart of all our research work as specialists in patient reported outcomes and the patient experience. PPI research group members will complete Macmillan Cancer Support training as needed. We had originally intended to make use of the RDS "Building Research Partnerships" scheme locally but RDS has no intentions of running a training event in the relevant time period. Macmillan Cancer Support also have stopped running their training programme with no further sessions planned in 2015. However they have suggested we use the course content, which is available on their Learnzone to use (users may have to enrol but this is free), and will put us in touch with experienced facilitators to optimise use. We will involve cancer service users and their partners or carers, recruited through our collaborator and co-applicant networks and also the local cancer networks as necessary. Costs for PPI time, direct expenses, and carer costs have been budgeted for.

15 STUDY END

We have planned 19 months for the duration of the active phases of the study. The study will be considered to have ended when the first articles about its findings have been published or within six months after the second round of checks on the dashboard technology in the proof of concept, whichever is sooner. At this time the data storage will be reviewed with the University of Southampton eprints team as described in the data management protocol.

16 IMPACT

We will hold a launch event and a dissemination event in January 2016 and will evaluate dissemination at the end of each stage. We will hold dissemination events at the end of the study. All such events need to consider the range of stakeholders.

We will develop a dedicated website for the study, as well as the dashboard website. This will be set up as soon as possible.

We will maintain a Twitter account, and use any other social media or other platforms deemed appropriate. These will be maintained by the lead researcher or other researcher as agreed at the time.

We will have a study blog in place of a newsletter though we will ensure that this can be printed for sharing with participants who do not have internet access, which we will ascertain from the contact details they have provided and also from asking them directly when they participate in stage 2 and 3.

IPR considerations will be handled by the University of Southampton but staff need to be aware of their potential throughout.

17 PUBLICATIONS AND OUTPUTS

Findings will be disseminated as soon as possible after the end of the study.

1. Our main outputs are a transferable process and taxonomy/text engineering rules. Manuals for their use in cancer surveys and to modify for other topics will be disseminated widely e.g. via collaborator websites/networks: Insight & Feedback, NHS England; Macmillan Cancer Support; Healthtalkonline, and via the Department of Health and the Care Quality Commission (CQC).
2. We will explore transfer to CQC PES in a separate report.
3. We will work with professional/clinical leads at DH and NHS England, the CPES strategy group and our collaborators at national and regional level to disseminate CPES outputs.
4. We will explore the possibility of training workshops in the use and transfer of our approach.
5. We will seek publication of our development process and its outputs including thematic analysis in methodology, research and practice journals to reach the widest possible audience and target open access publications. We will cascade dissemination for maximal impact.
6. We will present at research/professional conferences and seminars locally, nationally and internationally. We have costed in 1 national and 1 international conference within the primary funding; we will seek funding elsewhere for other meetings.

7. We plan 2 launch events in year 1 and 2 dissemination events in year 2 with a view to involving Trust staff in one event in each year and a more mixed audience in the others.
8. We will produce NIHR/DH materials and reports.

The PI will be the primary or last named author in each of these.

Other authorships will be determined at a Writing Committee the first meeting of which will occur during the first data collection phase. Authorships will be determined according to contributions made.

All proposed publications and disseminations including abstracts and talks for conferences need to be sent to the NIHR's representative for approval; as a minimum this needs to be at the same time as submission for publication or at least 28 days before the date intended for publication whichever is earlier. This obligation continues after the end of the Research Period and is contractual.

Disseminations should have the following acknowledgement where possible or a shortened version if not:

> *"This report is independent research funded by the National Institute for Health Research (Health Services and Delivery Research, 14/156/15 - PRESENT: Patient Reported Experience Survey Engineering of Natural Text: developing practical automated analysis and dashboard representations of cancer survey freetext answers). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health."*

18 PROTOCOL AMENDMENTS

Names removed for publication on NIHR site 25/11/15: CR.

19 REFERENCES

1. O'Cathain A, Thomas KJ (2004) "Any other comments?" Open questions on questionnaires – a bane or a bonus to research? *BMC Medical Research Methodology*, 4,25
2. IoM (2001) *Crossing the Quality Chasm: A new health system for the 21st Century*. Washington, DC, National Academy Press.
3. Darzi A (2008) *High Quality care for all: NHS Next Stage Review final report*. London, DoH.
4. Baranski A (2013) *Briefing: The importance of the National Cancer Patient Experience Survey: report,* Macmillan Cancer Support, London
5. Quality Health (2013) *Cancer Patient Experience Survey 2013 National Report*: 30 August 2013, Quality Health: Leeds
6. Eggers S, Huang Z, Chen H, Yan L. Larson C, Rashid A, Chau M. Lin C (2006) Chapter 2 Mapping medical informatics research, In: Chen, H., Fuller, S. S., Friedman, C., & Hersh, W. (editors). *Medical informatics:*

knowledge management and data mining in biomedicine (Vol.8), Springer, pp 35-62
7. Gettinger A, Chaimberg K, Spence B, Nelan J. High-level alerting for rare but critical clinical conditions. *AMIA Annu Symp Proc.* 2008 Nov 6,1065
8. Chowdhury SR, Saha H (2010) Development of a FPGA based fuzzy neural network system for early diagnosis of critical health condition of a patient. *Comput Biol Med*. Feb;40(2),190- 200. doi: 10.1016/j.compbiomed.2009.11.015
9. Salgueiro M, Basogain X, Collado A, Torres X, Bilbao J, Doñate F, Aguilera L, Azkue JJ (2013) An artificial neural network approach for predicting functional outcome in fibromyalgia syndrome after multidisciplinary pain program. *Pain Med*. Oct;14(10):1450-60. doi: 10.1111/pme.12185
10. Goodrich GG, Goodman PH, Budhecha SK, Pritsos CA (2009) Functional polymorphism of detoxification gene NQO1 predicts intensity of empirical treatment of childhood asthma. *Mutat Res.* Mar 31;674(1-2):55-61. doi: 10.1016/j.mrgentox.2008.10.009
11. Mattfeldt T, Wolter H, Kemmerling R, Gottfried HW, Kestler HA (2001) Cluster analysis of comparative genomic hybridization (CGH) data using self-organizing maps: application to prostate carcinomas. *Anal Cell Pathol*.;23(1),29-37
12. Cunningham H, Tablan V, Roberts A, Bontcheva K (2013) Getting more out of biomedical documents with GATE's full lifecycle open source text analytics. *PLoS Comput Biol* 9(2), e1002854. doi:10.1371/journal.pcbi.1002854 — http://tinyurl.com/gate-life-sci/
13. May C (2013) Towards a general theory of implementation. *Implementation Science*, 8, (18), 1-14.(doi:10.1186/1748-5908-8-18). (PMID:23406398).
14. May, CR, Eton DT, Boehmer K, Gallacher K, Hunt K, MacDonald S, Mair F S., May CM, Montori V, Richardson A, Rogers AE, Shippee N (2014) Rethinking the patient: using Burden of Treatment Theory to understand the changing dynamics of illness. *BMC Health Services Research*, 14, (1), 281. (doi:10.1186/1472-6963-14-281).
15. Medical Research Council: *Developing and evaluating complex interventions: new guidance*. London, MRC, 2008.
16. Wagland R, Recio-Saucedo A, Simon M, Bracher M, Foster C, Hunt K, Glaser A, Corner J. Text-mining assisted qualitative analysis of free-text comments relating to quality of patient experiences of care within the National Cancer Experience Survey *BMJ Quality and Safety* (submitted).
17. Corner J, Wagland R, Glaser A, Richards M (2013) Qualitative analysis of patients' feedback from a PROMs survey of cancer patients in England. *BMJ Open,* 3:e002316.
18. Sears A (1997) Heuristic walkthroughs. Finding the problems without the noise[pdf]. International *Journal of Human Computer Interaction* 9(3),.213-234
19. Spencer, R. (2000) The streamlined cognitive walkthrough method, working around social constraints encountered in a software development company. *CHI 2000. Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp353–359. doi: 10.1145/332040.332456
20. Wharton C, Rieman J, Lewis C, Polson P (1994) The Cognitive Walkthrough method: A Practitioner's Guide. In *Usability Inspection Methods*, J. Nielsen and R.L. Mack (Eds.), New York: John Wiley & Sons, pp.105-141.
21. Nielsen J. (1994). “Heuristic evaluation”. In Nielsen, J., and Mack, R.L. (Eds.), *Usability Inspection Methods*, (New York: Wiley)
22. Gerhardt-Powals, J. (1996). Cognitive engineering principles for enhancing human – computer performance". *International Journal of*

Human-Computer Interaction 8(2), 189–211. doi:10.1080/10447319609526147

23. Yourdon, Edward. (1985) *Structured Walkthroughs* (New York: Yourdon)
24. Aggarwal, CC. Chapter 9.Mining text streams. In: *Mining text data*. Editors: Charu C.Aggarwal, Chengxiang Zhai. Kluwer Academic Publishers, London, 297-322