Sequential Research to Evaluate the Impact of Patient and Public Involvement on Cancer Research Outcomes: Using Interviews, Stimulus Material and a Modified Delphi Technique

Raksha Pandya-Wood, PhD

Abstract

Evaluating patient and public involvement (PPI) in healthcare research continues to attract international interest. This article discusses how one exemplar study evaluated the impact of PPI on cancer research outcomes, with user involvement sewn into the design. The research aligned to interpretivist and pragmatist paradigms and resulted in a mixed methods sequential design. Phase 1 involved 23 in-depth interviews to explore perceptions of impact of PPI on cancer research outcomes with patients, researchers and stakeholders. Analysis from Phase 1 formed the basis of a 'stimulus paper' to use in Phase 2. Phase 2 adopted the modified Delphi technique with a virtual panel of 35 experts. This research found several factors shaped the impact of PPI on cancer research outcomes. However, the data itself are not the foci of this article, the methodological process, theoretical decisions, limitations and lessons learned across the research are.

Introduction

PPI has been set as a global policy imperative by the World Health Organisation (WHO). The declaration of Alma Ata (1978) stated that: *'…people have the right and the duty to participate individually and collectively in their health care'* (Declaration of Alma-Ata, 1978 [no page number listed]). PPI has been politicised as a way of servicing a spectrum of needs for democratic governments and for the public. These needs range from increasing peoples' rights in healthcare through to legitimising healthcare services to help address public concerns (Forster & Gabe, 2008). Parallel to this, PPI in research is underpinned by what Snape et al. (2014) define as 'intrinsic values'. The public have an entitlement to be involved in the research process, they have the right to say what research is undertaken and importantly, they have the right to shape how research is used. Many countries involve patients and the public in healthcare governance and decision-making, including

Australia (Todd et al., 2018), North America (Frank et al., 2014), Canada (Boivin et al., 2014) and countries across Europe (Brett et al., 2010). The current study was conducted in the UK, where it is common practice for public funded research studies to have PPI in the design and conduct of the research.

For more than a decade, the appeal to demonstrate impact generated from PPI in research has grown extensively in the UK (Boivin et al., 2018; Russell et al., 2020). There are a plethora of new ideas forming about the evaluation of PPI on research. Most of these research studies have focused on the process of PPI. But whilst implementation science attempts to help close the research and practice gap, (Morris et al., 2011; Butler, 2008;

Independant Researcher, Malaysia

Corresponding Author:
Raksha Pandya-Wood, Independent Researcher, Malaysia.
Email: rpandyawood@gmail.com

Norman, 2010; Bero et al., 1998) *specifically* for PPI in research, there continues to be little evidence on how PPI makes a difference to the outcomes of health research (Brett et al., 2014; Evans et al., 2014; Mathie et al., 2014; Mockford et al., 2011; Shippee et al., 2013; Snape et al., 2014). We have insufficient understanding of the types of issues which affect the implementation of PPI-informed research findings into policy and practice (Staniszewska, Brett, et al., 2011). The present study was therefore designed to address these knowledge gaps.

This article is a social science inductive inquiry that offers an exemplar methodology, in which a combination of interpretivist and pragmatic approaches resulted in a qualitative mixed methods sequential design. The research was the first of its kind to use the modified Delphi technique using stimulus material generated from the interviews, to consolidate understanding of the order of importance across seven themes that affected impact of patient and public involvement. The study used the views and opinions of patients, clinicians, policy makers, politicians and others.

The article starts with an ontological and epistemological grounding, followed by an explanation summarising end user involvement in this research. Then a justification is offered for why cancer was selected as the disease focus in this research. The article moves onto the theoretical grounding behind the methods selected, research tools adopted, and the practical adjustments made to conduct in-depth interviews and a modified Delphi technique. For both phases of the work, the focus of questions, sample frame, sample size and the analysis process adopted are discussed. At the end of the article, there is a discussion on quality, a methodological critique and a conclusion. Throughout the article the term 'participants' refer to interview participants (Phase 1), and 'panellists' for those who took part in the Delphi (Phase 2).

Methodological Stance and Theoretical Orientations

For the current study, the aim was initially to seek understandings about 'perceptions of impact'. The interpretivist position was therefore adopted as it allows for difference and disputes, valuing varied perceptions and opinions (Dyson & Brown, 2005). However, early involvement with end users of this work (described below) had also identified that something tangible should come from the current study to help shorten the gap of understandings about impact of PPI on research outcomes (see Supplemental file-1 section G, H and I which describes a workshop which took place in April 2011). Thus, the current research needed to be pragmatic too. Pragmatism in this research study was embraced as it:

> … accepts, philosophically, that there are singular and multiple realities that are open to empirical inquiry and orients itself toward solving practical problems in the "real world"

> (Feilzer, 2010, p. 8)

Linked to the idea of epistemology is the question about social ontology. For this research, it was felt that the social ontology aligned to the constructivist view because new meaning unfolds every day about aspects affecting the impact of PPI on research outcomes and together these opinions build knowledge. The current study also aligned to the idea of inductive enquiry. Inductive enquiries explore how the data generated relate to theory, allowing researchers to generalise interpretations from data followed by deduction (Bryman, 2012).

Given the study's focus on PPI, end user involvement was a central component and built into this study. In this research, INVOLVE's definition of involvement was adopted. INVOLVE defined involvement as 'Doing research "*with*" or "*by*" people, not "*to*," "*about*" or "*for*" them' (INVOLVE, 2012).

There are two entwined aspects that require explanation here as they impacted upon the thinking behind this study. Firstly, the researcher's own reflexivity on the topic being studied (at the time the researcher worked as a public involvement specialist) and secondly, critical consideration of the involvement of others and their impact on this study. Mindful of this, it was decided that a journal would be kept about key research decisions and an involvement reporting tool would be used called Guidelines for Reporting the Involvement of Patients and Public in research (GRIPP) 2 (Staniszewska et al., 2017). This tool helps researchers to report on the involvement aspects carried out in their research. The involvement of potential 'end users' had an impact on the current study's research design from a pragmatic point of view to help methodological decision-making, and to aid conceptual clarity. End users in this work were a mix of patients and the public, academics researchers, clinicians, policy makers and any other person interested in cancer research. On INVOLVE's (2012) spectrum of involvement, there are three levels which range from 'consultation', 'collaboration' to 'user controlled'. For the current research, the idea of 'consultation' was adopted due to the time and resource constraints of a PhD. Consultation is broadly defined as:

> "when you ask [people] for [their] views and use these views to inform your decision making. Consultation can be about any aspect of the research process – from identifying topics for research through to thinking about the implications of research findings.".

> (INVOLVE, 2012, p. 21)

In summary, involvement in this research occurred during: study aims development; research design for both phases including: piloting questions; analysis and developing the originality. Involvement encounters occurred mostly face-to-face and sometimes remotely (over email and telephone). Overall, 12 key influences of involvement occurred over the course of the study (on average, twice yearly) and over 110 people were involved. A full breakdown of how consultation-based involvement impacted upon this research can be found

in Supplemental file 1. One major design issue that was raised in an early end user involvement workshop with health policy academics was the importance of having an area of disease as a context to focus the research on.

Cancer was selected as the appropriate topic of disease for several reasons in this study. Firstly, more than one in three people will be diagnosed with cancer in their lifetime (Cancer Research UK, 2011). Secondly, cancer research is one area where PPI is already advanced (Hubbard et al., 2007; Stewart et al., 2011) meaning there are a plethora of completed research study examples with PPI. Thirdly, the choice of cancer as the topic would allow for targeted data collection. Finally, the focus on one disease also allowed some rigour and replicability, as empirical work would be richer, fuller and more detailed, building understandings of perceptions and accounts of involvement that professionals, patients and the public offered.

The use of qualitative data collection, then analysis, followed by further data collection was necessary to help develop the pragmatic approach. Traditionally, this has not been compatible for some researchers (Denzin & Lincoln, 1998, p. 8). However, it is now well established that mixing methods offers tangible benefits to health and social research because one of two datasets could serve as an explanation towards the other (Bryman, 2012). There are many examples of published studies in cancer research where qualitative researchers have adopted mixed methods approaches. For example, in a study to establish which outcomes should be included in core outcomes set for oropharyngeal cancer trials, Walters et al. (2014) used qualitative interviews with patients and carers about research outcomes that matter to patients, followed by a consensus study to refine the contents of the core outcome set. The collective methods here proved beneficial as it provided sufficient outcomes knowledge for future trials. Further, a Bladder Cancer study by Bessa et al. (2019) developed a modified Delphi method in a PPI setting after conducting a systematic review and two focus groups with patients and health care practitioners, finally, they held a consensus meeting with both stakeholder groups. This led to achieving a list of unanswered bladder cancer research questions.

Walters et al. (2014) and Bessa et al. (2019) demonstrate that through employing two or more approaches to data collection, data integrity and credibility could be enhanced because findings would help to build further understandings (Burgman, 2008). Thus, data refinement would offer pragmatic understandings that were needed for the richness of the data collected (Creswell, 2003). However, to use two methods, both of which were mainly qualitative, required careful consideration. The qualitative findings could offer context to the external validity or broader variables uncovered through the counted data (Creswell et al., 2008). Counted data in this work concerned voting for order of importance once the themes would emerge from the interviews. In turn, the themes would generate understanding towards a diversity of views across the groups (Bryman, 2012). There is room in qualitative work for counting. Autonomous counting in qualitative research is used when, ranking may help demonstrate the significance and importance of a particular (set of) issue(s) (Hannah & Lautsch, 2011).

Design

The overarching study design is explained here, offering detailed review of the methods used in the study. This work adopted an exploratory sequential design for data collection (Creswell et al., 2008). Data were collected in two phases, through interviews and a Delphi technique, see Figure 1 below. A total of $n = 23$ participants took part in the interviews. Interviews were chosen to form rich accounts of understandings about perceptions of impact of PPI on research outcomes. Patients, researchers and stakeholders provided information about cancer research studies with PPI in the design and conduct and which had finished. Themes identified from the interviews were fed into a 'stimulus paper'. Phase 2 of the research was carried out in the form of a three-round modified Delphi survey (interview data themes formed the preliminary work for the Delphi survey). The Delphi survey's purpose was to offer a sophisticated yet practical understanding of the complex social issues that the interviewees had identified from their accounts of understandings and perceptions concerning the impact of PPI on research outcomes. This phase was carried with $n = 35$ panellists (those working in leading charities and large non-government organisations, policy makers, politicians with current or previous health portfolios, academics, independent consultants, government department leads and 'expert patients'/patient champions). A demographic breakdown of interview participants and Delphi panellists is provided in Tables 1 and 2.

Phase 1 – Qualitative Interviews

The aim of this phase was to understand the experiences of others, inviting them to describe their own perspectives. Semi-structured interviews, through probing, enabled follow-up questions for further clarification and detail. They can allow a space and opportunity to talk, rather than being constrained by pre-identified categories of response, using people's own vocabulary about what they find significant and important to them (Davies & Hughes, 2014). Interviews were therefore carried out face-to-face, rather than remotely.

Preparing an Interview Guide

An interview guide was developed with a list of appropriate and focused questions about people's experiences and knowledge about the impact of PPI on research outcomes. From reading the literature on the topic, gaps were identified. The planned involvement channels of this work also helped with potential questions. National meetings and conferences, focussing on PPI in research, helped the researcher to consider frontline issues that patients, researchers and stakeholders were struggling with in relation to the impact of PPI on
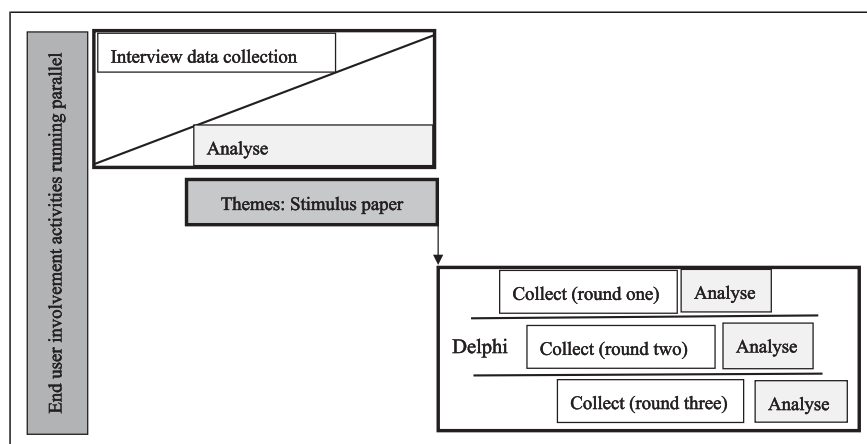
Figure 1. The data collection process.

Table 1. Interview Participants Demographics.

Category	Patients	Researchers	Stakeholders
Sex	F [n = 4] and M [n = 2]	F [n = 3] and M [n = 5]	F [n = 4] and M [n = 5]
Age range	60–65 across group	40–60 across group	22–60 across group
Ethnicity	All White British [n = 6]	All White British [n = 8]	All White British [n = 9]
Experience	Breast cancer [n = 3], oesophageal cancer [n = 1], non-Hodgkin's lymphoma [n = 1], and stomach cancer [n = 1])	Clinical geneticist (doctor) [n = 1], and clinical and academic professor [n = 7]	Broad and varied ranging from junior policy staff in charities, to professors and directors of research centres and a hospice. Two had been awarded the Queen's honour/award. Some had PhD and professor in their title
Type of cancer study they discussed involving PPI	Randomised control trials (RCTs) [n = 2], community research [n = 1], surveys [n = 1], service improvement/redesigning a service [n = 1] and improving cancer experience [n = 1]	Randomised control trials (RCTs) [n = 4], palliative care [n = 1], cancer health provision inequalities [n = 1] and community care [n = 2]	Community research [n = 2], palliative care [n = 2], participatory action research [n = 1] and improving cancer experience [n = 9]

research outcomes. At the time when data collection was being planned, five national studies had been funded on the impact of PPI. These studies became public knowledge on the funders' websites. Efforts were made to ensure the current study remained uniquely focused on the impact of PPI on research *outcomes*. Finally, discussions with colleagues were also considered when generating questions.

The interview guide questions were designed in a way that was suitable for any of the three participant groups being interviewed. Topics for questions followed a logical and chronological structure. The interviews opened with two familiarisation questions: (1) information about the study and (2) what the motivations were for PPI were in the study. Then (3) how patients and public were supported for their roles in the research (mindful that this would also generate research process issues – but necessary, as it added more context). The interview then proceeded to ask questions about the outcomes from the study, including (4) key messages disseminated, (5) what had happened since the study had finished and lastly, (6)

how participants understood impact. The interview guide was not used verbatim but as a topic guide.

Pilot Interviews (End User Involvement)

Once developed, the interview questions and interview process were piloted with two people from each group (researchers, patients and PPI stakeholders). The pilot phase was invaluable for many reasons: it helped the researcher to gain confidence; ensure that the questions followed a logical flow; adjust wording of certain questions and monitor time. Minor changes were made to suit each of the three groups and cues for questioning certain groups. Data from pilots was not used in the analysis.

Sampling Framework

Participants were identified using purposive, non-probability sampling (Tansey, 2007). Using academic and professional

Table 2. Delphi Panellist Demographics.

Category	Patients and Carers	Academic and Clinical Researchers	Policy and Commissioning Work	Stakeholders/Healthcare Professionals and PPI Work
Sex	F [n = 5] M [n = 2]	F [n = 5] M [n = 3]	F [n = 10] M [n = 7]	F [n = 5] M [n = 2]
Age range	55–70 [n = 7] across group	40–65[n = 8] across group	35–60 [n = 17] across group	40–65 [n = 7] across group
Ethnicity	White British [n = 7]	White British [n = 6], Asian British [n = 1] and Black British [n = 1]	White British [n = 15], Asian British [n = 1], and White and Black African [n = 1]	White British [n = 6], Asian British [n = 1]
Experience	Cancer survivor [n = 5], carer [n = 2]	Professor status [n = 5], independent researcher (Dr title) [n = 2] and unknown [n = 1]	Commissioning [n = 3], government [n = 7] and cancer policy work [n = 7]	PPI work [n = 5], communication [n = 1] and quality [n = 1]
Expertise	National role of being an expert patient in international cancer research trials [n = 5] caring for a spouse with cancer [n = 2]	Each had won funding for conducting cancer research with PPI. They had PhDs in a social science discipline [n = 3] clinical doctors and principal investigators in trials [n = 4] and clinical professor [n = 1]	Healthcare commissioner [n = 1], research commissioner [n = 1] health education commissioner [n = 1], national cancer charity policy staff [n = 5], members of parliament [n = 3] policy thinktank of research [n = 1] national patient champion body [n = 1] NHS senior cancer communication staff [n = 4]	National roles involving: Advocating patient voice in research [n = 5] communicating trial results [n = 1], ethics and quality assurance in research [n = 1]

networks, three groups were recruited: patients and the public, researchers and stakeholders.

To help the current study retain its unique focus, the following inclusion and exclusion criteria were applied:

- Researchers and patients interviewed were required to be able to discuss a cancer research study from the last 5 years[1] which had patient and public involvement in the research design and conduct.
- Preference was given to finding participants from the East Midlands (a region in England). Focussing on one regional geographical boundary enabled convenience and snowball sampling approaches to be used by the researcher (Denscombe, 2014).
- Participants needed to be over the age of 18 and able to speak in English. Excluding non-English-speakers was a limitation for this research but was necessary because there were no funds available for translation services. This limitation was magnified given that Leicester, from where most participants were recruited, is a diverse and multicultural city where over 130 languages and dialects are spoken (Census, 2011). Sheldon and Parker's (1992) work on race and ethnicity highlighted problems associated with health research and its limitations for not integrating aspects of race and ethnicity into health research strategies that include all groups.

Interview Sample and Data Collection Process

All potential participants were identified through the researcher's professional networks at the time and were contacted by email. In addition, a professional virtual network called CHAIN was contacted via email advertising this research. CHAIN cuts across health and social care sectors in the UK, and comprised a broad membership list, where members could select which subgroups they wanted to join for targeted information to be sent to them. Interest areas, to list a few, ranged from: 'PPI', 'Cancer', 'Better care without delay' and 'Service improvement' (all of which would work for the stakeholders group to be interviewed).

Data were collected until saturation was reached, that is, when information generated became repetitive and nothing new was being raised (Ives & Damery, 2014). Interviews were carried out with n = 23 participants (see Table 1). Interviews with patients took place in their own homes and were generally 'social' and relaxed in style, typically lasting 90–120 minutes. Interviews with researchers took place in hospital settings and university offices and lasted just under an hour – these interviews were the shortest. Interviews with stakeholders (who were academic health scientists and clinicians concerned with implementation science and policy work) took part in university offices, hospital offices and hospital cafes all over the East Midlands, for a duration of around 90 minutes. After all the interviews were completed, each participant was informed about the next stages of this work. Later, a handwritten card was posted to each participant to express gratitude for their time.

Data Analysis and End User Involvement

This section demonstrates how, from analysis of interviews, two overarching themes with seven subthemes were identified using Braun and Clarke's (2006) model of analysis. The first theme

was the *'Impact of PPI in research processes'* (two subthemes under this were *'PPI processes'* and *'Wanting to make a difference'*). The second overarching theme was '*Impact of PPI on research outcomes'*. Under the latter theme, the main focus of the study, the subthemes generated were: *Networks*; *Leadership and power*; *Resources and the political context*; *Dissemination*; and *Information and Communication Technology*.

All the interviews were recorded and transcribed verbatim. Braun and Clarke's (2006) six stages of data analysis were applied to this process. (1). *Data familiarisation* involved listening to the audio recordings. Once the transcripts were ready for analysis, these files were grouped according to the participants (patients, researchers and stakeholders). For the (2). *Generation of initial codes* phase of the process it meant searching for themes. Despite efforts to minimise data content on the impact of PPI on research processes, more than half of the overall data appeared to be on this topic. Briefly, this large theme was useful because it grounded the data provided. At any opportunity, participants eloquently spoke about their experiences of PPI processes (and in the case of researchers how they were (mostly) pro-involvement of patients in research), they discussed, for example, how opportunities about research were advertised. Patients spoke of the skills they brought to the research process and the sorts of things that motivated them to get involved. Stakeholders described the unique value of research which had involved patients and the public, suggesting that the respective research studies they were describing were 'better quality because of PPI'. It was decided that this information was key to understanding the outcomes of research. After PPI processes were grouped as a theme providing background information to help consider 'contextual issues', they were placed to one side to revisit later. All the remaining codes were themed in the (3). *Searching for themes* phase. For the (4). *reviewing themes* phase, this process involved checking that the themes worked in relation to the coded data and checking that they worked across the entire dataset. To ensure that the analysis so far was an accurate reflection of the emerging themes, a sample of transcripts were sent to the researcher's supervisors to help ensure consistency. This acted as a quality check of the researcher's analysis technique. Themes identified by the supervisors helped to clarify that some themes overlapped, for example, 'networks' and 'dissemination'. (5). *Defining and naming themes,* process was carried out carefully trying to capture depth of theme.

Three 'Making sense of the data' workshops were held with 'end users' see section G on the Supplemental file 1. The end users here were, a group of ($n = 8$) health policy academics, a patient group of ($n = 45$) patients, and attendees at a PPI conference, and a group of social science and health researchers ($n = 12$). The purpose of the involvement was to help broadly categorise the data. Workshop attendees were not shown the coding framework generated on NVIVO, they were only shown Inspiration™ generated visual maps of the themes. This process along with the researcher's own demonstrated that there were seven overarching themes: 'PPI process'; 'Wanting to make a

difference'; 'Networks'; 'Leadership and power', 'Resources and the political context'; 'Dissemination'; and 'Information and communication technology'. The final stage, stage 6, for Braun and Clarke (2006) analysis process involved *producing a report*. In this research, this report constituted a short description of each of the themes to help identify the area of focus for the next stage of the study. The content informed the development of a 'stimulus paper' (see Supplemental file 2.) for the Delphi in Phase 2.

Phase 2 – Modified Delphi Technique

The Delphi technique is a consensus-building method that collects data sequentially through two or more rounds of questionnaires (Campbell et al., 2004). Crucial to the success of the Delphi technique was the use of a 'stimulus paper' which acted as the modification, that is the prior data that was collected (Custer et al., 1999) through Phase 1 in this study.

Stimulus Paper

The literature on Delphi surveys provides no definition of a stimulus paper, but the use of 'stimulus text' in interviews is well documented. Silverman and Brull (1993) suggest a stimulus text offers context, more than just a question or a sentence making a proposition. A stimulus text is a description about an *'outline or story of an event or action, seen or experienced from a viewpoint, uttered by an identifiable or unidentifiable narrator'* (Silverman & Brull, 1993, p. 91–92). The description of the stimulus offering context fits with this research, as the data that were generated in Phase 1 provided initial contextual information about the impact of PPI on research outcomes. For Törrönon (2002), a stimulus text presents important analysis of what has been studied and found:

> The stimulus text [is] expected to articulate the phenomenon under examination to make it perceptible in such a way that …[those]… interpreting the stimulus text, are 'empowered' to express their social experience and cultural knowledge of the issue under question
>
> (Törrönon 2002, p. 345)

Adapting Törrönon's (2002) line of thinking in this work meant that the 'stimulus text' needed to be succinct and articulate, summarising the data themes from Phase 1, with the Delphi questions. In determining how long the paper should be, the researcher followed guidance on developing executive summaries, which suggested that many writers produce a summary under three pages (Custom Writing and Research website, 2013). The paper needed to be short enough to be read by busy professionals but long enough to be a stand-alone document. The final stimulus paper was two pages long and addressed the seven major themes 'PPI process'; 'Wanting to make a difference'; 'Networks'; 'Leadership and power',

'Resources and the political context'; 'Dissemination'; and 'Information and communication technology'.

Phase 2 of the research was concerned with confirming the importance of the themes identified in Phase 1, enhancing the credibility and offering external validity using a diversity of views. Therefore, a modified Delphi was useful for the current study as the themes provided panellists the context required for their opinions (Snyder-Halpern et al., 2000). There are several strengths identified in using this approach.

Panel Members are "Experts"

People who take part in Delphi surveys are often referred to as experts in their field (Snyder-Halpern et al., 2000) as the Delphi technique is reliant upon the use of 'expert' knowledge. The term 'expert' has been critiqued (Green et al., 1999) because it suggests that a 'layperson' may be unacceptable (Meyrick, 2003, p. 10). But Gutierrez (1989) argues that panellists in a Delphi survey should be a group of knowledgeable people, not necessarily 'experts'. In the Delphi survey by Boote et al. (2006), the research teams involved 'lay' as well as 'expert' people to make the study reflective of its focus, demonstrating PPI, and arguably this may have contributed to the study's success. In this research, the panellists were selected to provide relevant input to the process, have the highest authority possible, and or be committed to and interested in the research aims.

Defining Expertise

Six groups of expertise were selected for this study. Group 1 were those working in national charities and large non-government organisations such as the voluntary sector which plays an increasingly large role in cancer funding and provision and delivery of services (Tritter et al., 2003). Group 2 were policy makers, because they provide insight and understanding regarding the broader set of economic, administrative, managerial or policy-related factors that may influence the implementation of cancer care (Cotterell et al., 2011). Group 3 were academics, as they might have insight into why evidence-based healthcare has featured as a policy concern in many healthcare systems, driven by a growing recognition that healthcare delivery does not always reflect what is known to be best practice. Studies suggest that up to 30–40% of patients do not receive care which complies with current scientific evidence (Schuster et al., 1998). Group 4 were independent consultancies researching PPI and service improvement because these types of organisations provide additional business-driven insights into why involvement is important. Group 5 were government department leads and politicians, who could help build further knowledge on how legislation is being used/not being used to support the case for PPI in policy and practice in health and social care (Hughes et al., 2009). Finally, Group 6 was a mix of 'expert patients/ carers and patient champions', to help further understand their knowledge of services affecting them and the extent to which

they can challenge professionals' assumptions towards those with chronic illness (Wilson, 2001). An 'expert patient'/ 'carer'/'patient champion' in the current study was defined as someone who had lived experience of cancer, and now champions the patient voice in research or someone who cared for someone living with cancer.

Delphi Features

The Delphi provides a means of interaction between experts who cannot physically come together but whose participation may increase the credibility of the information gathered (Linstone & Turnoff, 1975). The questions in a Delphi survey are completed anonymously as panellists do not meet face-to-face (Hasson et al., 2000). The Delphi survey is designed to 'obtain the most reliable consensus of opinion of a group of experts…by a series of intensive questionnaires interspersed with controlled opinion feedback' (Dalkey & Helmer, 1963, p. 458). The Delphi technique allows for panellists to interact with each other in a controlled way, that is, the researcher pooling panellists combined knowledge into the controlled feedback (Rowe & Wright, 1999), without physically coming together and not allowing dominant members of a group to taint the views of others (Bolger & Wright, 2011). The Delphi technique reduces chances of powerful professionals with seniority manipulating others (Jairath & Weinstein, 1994). Thus, people taking part would not feel obliged to conform to fellow panellist views (Murphy et al., 1998). The Delphi survey was used successfully by Boote et al. (2006) on the principles and indicators of successful PPI in research. The study found that a common understanding was reached across all stakeholders on manifestations of positive involvement in research.

From a financial point of view, the Delphi was inexpensively facilitated. Another strength of the Delphi was that all communication was carried out via email and using the Blind Carbon Copying (BCC) which meant anonymity was achieved. This anonymity aspect proved useful if something important but controversial was raised by a panellist. Including controversial responses in the controlled feedback was important.

Snyder-Halpern et al. (2000) found that email responses, compared to posted responses were more legible, eased data entry and enhanced communication. Another advantageous feature of the Delphi survey was iteration. Between each questionnaire, controlled feedback was offered, through which the researcher presented a summary of the range of opinions in a numeric way highlighting voting patterns of themes, helping the group see where there was emerging importance, consensus and disagreement, allowing panellists an opportunity to reconsider their views. To form consensus and voting for importance of themes in this research context, the Borda count was used.

Consensus and Voting

The Borda count was developed by a French mathematician and political scientist in 1770 (Emerson, 2013) and is often

described as a consensus-based voting system rather than a majoritarian one. The Borda count allocates points corresponding to the number of options ranked lower. Once all votes have been counted, the option with the most points becomes the winner and the order of preference for the remaining issues being voted on is also achieved for example, 1st, 2nd, 3rd etc. This method was used for the consensus-building-aspect of the study (Lakhanpaul et al., 2014). The method was useful because it determined which of the seven (which later became nine) themes were deemed most to least important.

Sampling Framework

Keeney et al. (2006) argue that a researcher conducting the Delphi must decide on the inclusion and exclusion criteria before the study commences, such as the gender, professional experience, educational background and employment background of the panellists. As argued already, to reduce Delphi limitations, professional and non-professional (lay) people were considered as useful for this study. Panellists would have a broad range of skills and knowledge spanning a range of groups covering policy, practice, academia and patient experience to list a few. Table 2 shows the Delphi panel composition.

Panellists could take part in the study from anywhere in England. They needed access to the internet during the data collection phase of 6 weeks. They needed to be able to read and write English and they were selected based on their specialist expertise for this study from one of six groups listed above.

Sample Size and Recruitment Delphi

According to Reid (1988), there are variations in sample sizes for Delphi surveys depending on the type of research being planned. Sample sizes can range anywhere from 10 to 1500 people (Reyens & Hehn, 2000). Murphy et al. (1998) suggest that larger samples are likely to provide more reliable datasets when research questions have a limited range of answers. This work relied on qualitative responses and therefore too many participants would have become too complicated to manage for one researcher. Any fewer than 20 participants would have been likely to lead to incomplete understandings of this complex research area. Recruiting at least six people from each of the six backgrounds seemed manageable and realistic and also accommodated attrition.

To recruit the Delphi panellists, purposive and convenience sampling strategies were applied (Proctor & Allen, 2006). The researcher approached known academics, cancer charities, consultancies and policy networks, inviting people to participate. Individual letters were sent to local Members of Parliament, Department of Health leads and to members of the European Parliament. Any interested people who came forward were telephoned first to check that they met the criteria for selection, that they were available when the Delphi survey was planned, and that potential panellists understood that they

needed to be committed to the entire 6-weeks process. If they met the criteria, they were then emailed an information sheet, a Delphi process diagram with dates and a consent form. This initial contact was also an important opportunity for potential panellists to ask any questions.

A known problem with Delphi surveys is participant attrition as rounds progress (Mayaka & King, 2002). The researcher mitigated against attrition by sending email reminders midway through a round, and text message reminders for those who had not submitted on the final day of each round. As a result of this thorough strategy of retention, of the 39 people recruited only four people dropped out (a Member of the European Parliament, a representative from an independent political party focused on health, one patient and one academic). The analysis process between rounds was intense and took on average 80 hours per round.

There were three rounds to this work but there could have been more (Keeney et al., 2006), or fewer rounds (Hasson, 2000). After the panellists had read the stimulus paper, this research needed to establish: (1) how relevant the seven themes were to panellists in terms of order of importance, and why; (2) whether anything new should warrant a theme of its own, and why and (3) how impact of PPI on research outcomes could be better understood. Therefore, it was anticipated that three rounds would suffice for the current study.

Delphi Analysis and Controlled Feedback and User Involvement

Each Delphi round was analysed in real time. The system of analysis was similar to Phase 1, in the sense that it used the approach to thematic analysis set out by Braun and Clarke (2006). As most of the data produced were not too long in content, at the end of each round the key points made by panellists about emergent themes were noted to share in the next round. Qualitative data generated were often descriptive and NVIVO was used to manage data. Turoff and Hiltz's (1996, p. 71) technique was used to ensure that clarity, issues of bias, missing information, patterns, hidden disagreements and issues to focus the answers upon were considered throughout. They outline the following:

1) The data analysed and offered in the feedback needed to present a range of views and considerations;
2) That hidden disagreements and judgemental biases needed to be exposed to further clarification;
3) To detect and clarify any missing information or cases of ambiguity in interpretation by different participants;
4) To analyse complex situations only by analysis procedures (such as in the current research using Braun & Clarke, 2006);
5) To detect patterns of information and of subgroup positions (e.g. whether patients took a certain stance in their ranking preference) and

6) To detect critical items that need to be focused upon in the subsequent rounds (e.g. raising further questions about the themes or about impact of PPI on research outcomes).

The first and second points were clarified through round two of questions but point three was clarified by email with panellists as soon as responses started to come in, particularly if responses seemed ambiguous to the researcher. Points four and five used Braun and Clarke's (2006) thematic analysis process to understand patterns.

Once responses were received to each set of questions within the specified deadline, a list of answers was drawn up to keep in mind that the best opinion may have become 'watered down' (Sackman, 1975) or that the survey might generate 'bland statements' (Rennie, 1981). Researcher awareness of these criticisms reinforced the notion that analysis needed rigorous attention to detail concerning each response. Where possible, quotations were offered in the controlled feedback at the end of each round so that the original tone was retained, and any important messages were not misrepresented.

During the Delphi analysis, the input of the researcher's supervisors was key as it offered support in reading a sample of opinions and confirmed or queried the researcher's decisions. Any new questions that the researcher felt needed to be explored in the new rounds of questioning, the supervisors critically appraised. Involvement from an independent academic and a carer helped in this stage too.

Round One and the Controlled Feedback

In round one, 39 panellists read the stimulus paper and answered the questions. Along with ranking the seven factors, panellists were also asked to raise any additional issues that should be added to the seven factors. By the end of the round, the panel had raised two further themes: *PPI in commissioning* and *PPI in implementation,* and suggested two further issues: that the themes sit as *micro, meso and macro issues*, and that they *wanted a definition of 'impact of PPI'*.

Round two, Data Synthesis and Controlled Feedback

In round two, 35 panellists read the controlled feedback and answered the second set of questions. The two new factors were added to the existing seven factors and the panel was asked again to rank the issues but also across concepts such as micro, meso or macro issues. Panellists were asked to define the impact of PPI on research. In the controlled feedback for round two, data were pooled together, and the knowledge generated was shared, helping the next round (Reyens & Hehn, 2000).

Data synthesis can be conducted for different purposes (Mays & Pope, 2008). For the current research, the process of data synthesis served the purpose of formulating a definition of the concept of impact of PPI. The data that was used to form the definition came from an open-ended question asked to the panel: to provide in their own words, a definition of impact of PPI. Based on the 35 answers received a list of typologies were devised using the help of NVIVO software. Characteristics of the impact of PPI were drawn up. A synthesised definition (impact of PPI) was developed by the researcher, capturing the panellists' combined efforts and this definition was shared in the controlled feedback.

Round Three and the Controlled Feedback

In round three, 35 panellists read the controlled feedback and answered the third set of questions. The order of importance was found across the nine themes. The themes were situated at micro, meso and macro levels and this enhanced knowledge about how they affected the impact of PPI. A synthesised definition was offered about impact of PPI on research. Panellists were asked to what extent the collective definition reflected their individual definition and their view of what impact of PPI was. Panellists were asked to comment on whether the findings would apply to other disease areas. Panellists commented on future use and applicability of the findings, along with further research questions the work may have raised for them. When the Delphi process was complete, panellists were informed of the convergence and divergence of opinions that had occurred during the study. The researcher sent a final controlled feedback a week later, summarising the final round responses.

Once the Delphi survey was complete, a thank you card was sent to all panellists. Two panellists contacted the researcher afterwards to say that they felt the research had been conducted very well and efficiently and that the text reminders acted as a personal touch, as did the thank you card.

Quality Considerations

Validity

By employing mixed methods, the credibility of the findings was enhanced because they furthered the internal and external research rigour process (LeCompte & Goetz, 1982). The two datasets helped to serve as an explanation for each other. The qualitative findings provided sufficient accounts of rich and thick descriptions, helping to firmly establish the context of the themes generated from the Delphi process. Similarly, themes were confirmed as valuable in the importance order ranking exercise. Context was offered to the themes when Delphi panellists were asked to rank the themes and each panellist confirmed the themes' validity, linked to personal experience and understanding of the topic. This meant it was highly likely that the qualitative data collected did reflect the diversity of panellists' views. Therefore, the two methods used complemented internal and external validity. To further assess research rigour, Lincoln and Guba (1985) list four areas: *credibility, transferability, dependability* and *confirmability*.

There are several aspects that have helped the credibility of the current study. Through mixing methods, the two phases of the research design helped to contribute towards the validity of the knowledge created and increased understanding about the types of knowledge that people had about impact of PPI on research outcomes. Phase 2 refined the findings through using consensus-building methods. This double layering of data collection acted as a quality measure for internal validity (Morse, 2009). The interviews were carried out across three groups of people, drawing on a variety of viewpoints and experiences. Furthermore, the Delphi survey elicited a range of views from central government to patients.

The concept of transferability implies that the findings have applicability in other contexts and settings. Lincoln and Guba (1985) have suggested that qualitative researchers should be encouraged to produce 'thick description' which provides a strong foundation to make a judgement about transferability of the findings. Thick descriptions were produced during the interviews. These descriptions became the themes studied in the Delphi survey. Furthermore, during the Delphi survey, panellists ranked the information, suggesting that the data themes reflected a sense of reality. Not one panellist questioned the content of the themes' descriptions which were provided. Also, panellists were asked a direct question about the applicability of the current research in other contexts of health and disease (i.e. how transferable the findings were). Their responses demonstrated that, largely, data from this work were transferable beyond the disease of cancer, for studying the impact of PPI on research outcomes. However, particularly for cancer research and the evaluation of PPI there were some unique features this work had found. They raised that cancer research was an example of applied health research which had unique characteristics that differentiated it from other disease areas. They believed that cancer is positioned as a leading priority disease and to its related embeddedness is national research systems and infrastructure. For example, with regards to *resources and the political context,* it was felt that cancer is particularly well placed to benefit from government funding as well as the fact that cancer research charities attract large sums of money from public donations. This was also seen as important in terms of *commissioning*, since this was seen to follow from national priorities. The particular success that cancer charities have in advocating the patient voice was also highlighted as important for the *implementation* of findings. The existence of the National Cancer Patient Experience Survey (conducted across England annually) was felt to be a unique feature of the well-developed *leadership* in this research field within the UK. The well-established *networks*, are also characteristic of cancer research.

Dependability concerns the findings being consistent and reproducible. Lincoln and Guba (1985) suggest that an audit trail be kept by researchers as an aid memoir. During the current research, notes on involvement meetings, fieldwork pilots, supervision meetings, discussions with colleagues and all versions of data collection tools and analytical procedures were kept. This criterion demonstrates transparency and that the decisions made about the research were justifiable.

Confirmability is about the degree of neutrality, which concerns being mindful of the researcher's own identity. At the time the researcher worked for the National Institute for Health Research, Research Design Service, the researcher was in a unique position to conduct this study. Other approaches that have helped with confirmability included having regular involvement meetings in this work. The researcher argues that involvement enables better research and achieves the confirmability feature of Lincoln and Guba's (1985) model.

Research Limitations and Methodological Critique

Phase 1 of the study was designed to be East Midlands region (UK) focused but Phase 2, the Delphi survey was a diverse mix across England represented in the sample. Whilst the findings therefore present a valuable account, the research was limited geographically. Those working in the field of PPI evaluation elsewhere in the world will find inevitable differences. Also, worth noting, during the recruitment stage, active efforts were made to find participants from black and minority ethnic (BAME) backgrounds to help understand if there were any different experiences amongst the participants based on their ethnicity (Dawson, 2018). No participants came forward from this group. One reason for the lack of BAME patient and public participants might be because cancer can be stigmatised in BAME communities (Jones et al., 2015) and because of this stigma they may participate less in research. In Phase 1, all six patient participants came from a white background, and in Phase 2, once again, all seven expert patient panellists identified as white, thus the sample does not reflect the various different subgroups of populations living in the UK.

Being skilful at interviewing required practice and constant reflection (Roulston, 2010). It is well known that interviews with elite participants, such as those holding positional power, senior roles or public office, can shift the power dynamics between researcher and participant (Robson, 2002; Littig, 2009). In this work, the confidence of the researcher was therefore deemed to be an important attribute to overcoming such barriers, during the pilot phase. Also, interviews mostly generate retrospective accounts (Taylor, 2005) and past events can be misremembered, implying inaccurate data might be collected. To mitigate this, only participants with recent experience were sought.

Reducing the stimulus material down to two pages, some might argue may have moderated the data too much, however, given the end user involvement, and making a commitment to listening to the end users views the researcher was advised that people were not likely to read anything longer than two sides, so a difficult decision needed to be made about what to include and what to exclude. Future research teams adopting a similar approach may find it beneficial to be more flexible on the length of the stimulus paper. Also noteworthy, Delphi surveys can become very intense, especially between rounds. The researcher was

mindful that people recruited into the Delphi process were busy. A word limit for each question was not set so that panellists could, if they wanted, provide examples to help further contextualise their response. It was also decided that panellists were unlikely to read controlled feedback which was longer than two sides of A4 paper (applying the same principle as for the stimulus paper length). Whilst the length seemed sufficient, there may have been more opportunities to share more controlled feedback if the feedback was more detailed (longer).

The Delphi survey was carried out via email, apart from when the Members of Parliament (MPs) were involved, where two physical meetings were scheduled to answer the Delphi questions. During the six-week-period, for the major political parties, it was conference season – hence them requesting an interview rather than email. This adjustment to the planned data collection (a face-to-face meeting rather than electronic email response) was necessary as the MPs were too busy to take part in some of the rounds and reading the controlled feedback. To mitigate dropout, the researcher offered verbal controlled feedback and asked the questions directly. This may have impacted on the data collection as others on the Delphi (although each panellist had direct access to the researcher's telephone number to call if they had any questions) did not have the opportunity to discuss their answers.

There was one qualitative researcher in the Delphi survey who raised that they had found the process of the Delphi useful and successful, but interaction not fulfilling enough. This panellist would have preferred face-to-face meeting with the other panellists. This point demonstrates that at least one person did feel able to voice criticism during data collection, putting aside the current researcher's relationship to them. In this situation, it was explained why the Delphi was selected, grounding the answer in the Delphi process' cost effectiveness suitability and strengths, that is, panellists being anonymous and not feeling pressured to answer in a particular way.

Links within the research field meant that some of those with whom there was contact with through work became participants, and their participation sometimes resulted into snowballing (Noy, 2008). This was particularly the case with stakeholders. It needs to be acknowledged therefore that there may have been the potential for the obligation on people to participate (Feeley, 2002). However, an environment was created which tried to ensure that participants felt comfortable with withdrawing from the study. In the Delphi survey, a panellist who was an expert patient did indeed leave the study after round 1. In an email correspondence outside of the Delphi process, the individual conveyed that they felt the current study was not about cancer as initially anticipated and that the focus around impact was too abstract. This was despite attempts to ensure that the study information sheet was clear about the aims and focus. This individual's experiences of cancer research were linked to very specific understandings about cancer trials and finding a cure. This example illustrates that, in at least one case, a patient participant did feel able to withdraw from this study.

Despite these constructive criticisms listed under *Research limitations and methodological critique*, this study was successful because it refined and developed nine themes when evaluating PPI on research outcomes. The Delphi also helped to rank order of importance and furthered understanding about whether the themes were situated at micro, meso or macro levels. Forming a definition of the impact of PPI was not part of the plan for the Delphi, but by 'going with the flow' of what panellists were raising in the first round, the opportunity was there to be taken. With some quick thinking and a helpful discussion with the research supervisors, the panellists were asked to define impact of PPI themselves, and a definition was also achieved through the data synthesis process.

Conclusion

Sometimes, it becomes necessary to mix ideas which might disrupt the traditional research philosophies (McChesney & Aldridge, 2019) but the current research demonstrates that taking a social constructivist stance can allow flexibility. From the outset, evaluating the impact of PPI from an interpretivist position implies a paradox since it is often concerned with assessment, measurement and counting, features of a positivist philosophical position. However, this research article has outlined that a blended approach can bridge an interpretivist and pragmatist paradigm. It is hoped that this exemplar will add to the bank of annotated examples in a growing field because there are very few methodological examples published about evaluating the impact of PPI on research outcomes. Researchers have begun to highlight that this apparent lack of evidence may be due to poor quality reporting of PPI in research (Staniszewska, Brett, et al., 2011). As a result, PPI could become vulnerable to poor practice or tokenistic use. This concern has been described as an enduring challenge (Russell et al., 2020). More specifically, for this qualitative research, evaluating the impact of PPI on research outcomes was not only about having to navigate the researcher's own ontology and epistemology but when end users became involved, it became necessary to simultaneously navigate *their* ontology and epistemology. The academic research community have a duty to untangle these ideas, learn from them, and keep the discussion alive so that globally, patients and the public, PPI advocates, health and social care researchers and policy makers engaging in the PPI process can all objectively and subjectively understand how different views interact.

The approaches used here have helped to form new understandings about the impact of PPI. PPI in research is a complex and multi-layered process and its impact is multi-faceted. To build on the methodological ideas presented here, future researchers may find value utilising implementation science theory alongside PPI frameworks to help evaluate their own PPI on research outcomes. The purpose of sharing this article was to demonstrate that embedded user involvement within a sequential research design, using interviews and

Delphi survey, can create a sound starting point to evaluate PPI on research outcomes.

Author's Note

This article formed part of a PhD methodology. The PhD was funded by De Montfort University, England. During the time the research was taking place, Raksha also worked for the National Institute for Health Research, Research Design Service as a Senior Research Adviser on Patient and Public Involvement and was based at De Montfort University. Raksha relocated with her family to Malaysia in August 2020 and is currently not affiliated to an institution.

Ethical Approval

Ethical approval for phases 1 and 2 was conferred by the Faculty of Health and Life Science Ethics Committee, De Montfort University. Additionally, The Wellcome Trust (WT) *Good Research Practice Guide* (2007) was adopted for use in this study.

ORCID iD

Raksha Pandya-Wood https://orcid.org/0000-0002-4544-6247

Supplemental Material

Supplemental material for this article is available online.

Note

1. According to the Research for Patient Benefit funding stream, it is possible to demonstrate patient benefit between 3 and 5 years of a study finishing.

References

Bero, L. A., Grilli, R., Grimshaw, J. M., Harvey, E., Oxman, A. D., & Thomson, M. A. (1998). Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The cochrane effective practice and organization of care review group. *BMJ (Clinical Research ed.)*, *317*(7156), 465–468. https://doi.org/10.1136/bmj.317.7156.465.

Bessa, A., Maclennan, S., Enting, D., Bryan, R., Josephs, D., Hughes, S., Amery, S., Khan, M. S., Malde, S., Nair, R., Cahill, F., Wylie, H., Thurairaja, R., Chatterton, K., Kinsella, N., Häggström, C., & Van Hemelrijck, M. (2019). Consensus in bladder cancer research priorities between patients and healthcare professionals using a four-stage modified Delphi method. *European Urology*, *76*(2), 258–259. https://doi.org/10.1016/j.eururo.2019.01.031.

Boivin, A., Lehoux, P., Lacombe, R., Burgers, J., & Grol, R. (2014). Involving patients in setting priorities for healthcare improvement: A cluster randomized trial. *Implementation Science*, *9*(1), 24. https://doi.org/10.1186/1748-5908-9-24.

Boivin, A., Richards, T., & Forsythe, L. (2018). Evaluating patient and public involvement in research. *BMJ*. https://doi.org/10.1136/bmj.k5147.

Bolger, F., & Wright, G. (2011). Improving the Delphi process: Lessons from social psychological research. *Technological Forecasting and Social Change*, *78*(9), 1500–1513. https://doi.org/10.1016/j.techfore.2011.07.007.

Boote, J., Barber, R., & Cooper, C. (2006). Principles and indicators of successful consumer involvement in NHS research: Results of a Delphi study and subgroup analysis. *Health Policy*, *75*(3), 280–297. https://doi.org/10.1016/j.healthpol.2005.03.012.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa.

Brett, J., Staniszewska, S., Mockford, C., Herron-Marx, S., Hughes, J., Tysall, C., & Suleman, R. (2014). Mapping the impact of patient and public involvement on health and social care research: A systematic review. *Health Expectations*, *17*(5), 637–650. https://doi.org/10.1111/j.1369-7625.2012.00795.x.

Brett, J., Staniszewska, S., Mockford, C., Seers, K., Herron-Marx, S., & Bayliss, H. (2010). *The PIRICOM Study: A systematic review of the conceptualisation, measurement, impact and outcomes of patients and public involvement in health and social care research*. Royal College of Nursing, University of Warwick.

Bryman, A. (2012). *Social research methods* (4th ed.). Oxford University Press.

Burgman, M. (2008). *Advances in mixed methods research*. Sage.

Butler, D. (2008). Translational research: Crossing the valley of death. *Nature News*, *453*(7197), 840–842. https://doi.org/10.1038/453840a.

Campbell, S. M., Shield, T., Rogers, A., & Gask, L. (2004). How do stakeholder groups vary in a Delphi technique about primary mental health care and what factors influence their ratings? *BMJ Quality & Safety*, *13*(6), 428–434. https://doi.org/10.1136/qhc.13.6.428.

Cancer Research UK (2011). 'Accelerate the translation of research' available online: http://www.cancerresearchuk.org/funding-for-researchers/our-research-strategy/progress-report/accelerate-the-translation-of-research-our-progress (accessed on 6 10 2011).

Census. (2011). The office for national statistics. http://www.ons.gov.uk/ons/guide-method/census/2011/index.html (accessed on 14 2 2012).

Cotterell, P., Harlow, G., Morris, C., Beresford, P., Hanley, B., Sargeant, A., Sitzia, J., & Staley, K. (2011). Service user involvement in cancer care: The impact on service users. *Health Expectations*, *14*(2), 159–169. https://doi.org/10.1111/j.1369-7625.2010.00627.x.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks: Sage.

Creswell, J. W., Plano Clark, V. L., & Garrett, A. L. (2008). Methodological issues in conducting mixed methods research designs. In:. *Advances in Mixed Methods Reseåarch*, 66–83.

Custer, R. L., Scarcella, J. A., & Stewart, B. R. (1999). The modified Delphi technique--a rotational modification. *Journal of Vocational and Technical Education*, *15*(2), 50–58. https://doi.org/10.21061/jcte.v15i2.702.

Custom Writing and Research. (2013). 'The executive summary - common mistakes' available online: http://custompapers.com/summaries-mistakes/(accessed on1312 2013).

Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science*, *9*(3), 458–467. https://doi.org/10.1287/mnsc.9.3.458.

Davies, M. B., & Hughes, N. (2014). *Doing a successful research project: Using qualitative or quantitative methods*. Macmillan International Higher Education.

Dawson, S. (2018). *Patient and public involvement in health services research: Exploring the views and experiences of South asian participants – Shoba Dawson, PhD Candidate*. Institute of Population Health – ongoing at the University of Manchester.

Declaration of Alma-Ata (1978). International conference on Primary health care, Alma- Ata, USSR, 6–12 September available online. http://www.who.int/hpr/archive/docs/almaata.htm (accessed on 6 16 2018).

Denscombe, M. (2014). *The good research guide: For small-scale social research projects*. McGraw-Hill Education.

Denzin, N. K., & Lincoln, Y. S. (1998). *Strategies of qualitative research*. SagePublications.

Dyson, S., & Brown, B. (2005). *Social theory and applied health research*. McGraw-Hill Education.

Emerson, P. (2013). The original Borda count and partial voting. *Social Choice and Welfare*, *40*(2), 353–358. https://doi.org/10.1007/s00355-011-0603-9.

Evans, D., Coad, J., Cottrell, K., Dalrymple, J., Davies, R., Donald, C., Laterza, V., Long, A., Longley, A., Moule, P., Pollard, K., Powell, J., Puddicombe, A., Rice, C., & Sayers, R. (2014). Public involvement in research: Assessing impact through a realist evaluation. *Health Serv Deliv Res*, *2*(36), 1–128. https://doi.org/10.3310/hsdr02360.

Feeley, T. H. (2002). Comment on halo effects in rating and evaluation research. *Human Communication Research*, *28*(4), 578–586. https://doi.org/10.1111/j.1468-2958.2002.tb00825.x.

Feilzer, Y. M. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research*, *4*(1), 6–16. https://doi.org/10.1177/1558689809349691.

Forster, R., & Gabe, J. (2008). Voice or choice? Patient and public involvement in the national health service in England under new labour. *International Journal of Health Services*, *38*(2), 333–356. https://doi.org/10.2190/HS.38.2.g.

Frank, L., Basch, E., & Selby, J.V. (2014). The PCORI perspective on patient-centered outcomes research. *Jama*, *312*(15), 1513–1514. https://doi.org/10.1001/jama.2014.11100.

Green, B., Jones, M., Hughes, D., & Williams, A. (1999). Applying the Delphi technique in a study of GPs' information requirements. *Health & Social Care in the Community*, *7*(3), 198–205. https://doi.org/10.1046/j.1365-2524.1999.00176.x.

Grol, R. (2001). Successes and failures in the implementation of evidence-based guidelines for clinical practice. *Med Care*, *39*(8 Suppl 2), II46-II54. https://doi.org/10.1097/00005650-200108002-00003. PMID: 11583121.

Gutierrez (1989). cited in de Meyrick, J., 2003. The Delphi method and health research. *Health Education*, *103*(1), 7–16.

Hanley, B. (2005). Research as empowerment. In *Report of a series of seminars organised by the Toronto Group*. Joseph Rowntree Foundation.

Hannah, D. R., & Lautsch, B. A. (2011). Counting in qualitative research: Why to conduct it, when to avoid it, and when to closet it. *Journal of Management Inquiry*, *20*(1), 14–22. https://doi.org/10.1177/1056492610375988.

Hasson, F. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x.

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015. https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x.

Hubbard, G., Kidd, L., Donaghy, E., McDonald, C., & Kearney, N. (2007). A review of literature about involving people affected by cancer in research, policy and planning and practice. *Patient Education and Counseling*, *65*(1), 21–33. https://doi.org/10.1016/j.pec.2006.02.009.

Hughes, D., Mullen, C., & Vincent-Jones, P. (2009). Choice vs. voice? PPI policies and the re-positioning of the state in England and Wales. *Health Expectations*, *12*(3), 237–250. https://doi.org/10.1111/j.1369-7625.2009.00559.x.

INVOLVE. (2012). 'Briefing notes two' available online: http://www.invo.org.uk/posttyperesource/what-is-public-involvement-in-research/(accessed on 12 2 2012).

Ives, J., & Damery, S. (2014). Qualitative research. In D. M. Walker (Ed.), *An Introduction to Health Services Research: A Practical Guide*. Sage.

Jairath, N., & Weinstein, J. (1994). cited in Powell C 2003. in the Delphi Technique: Myths and realities. *Journal of Advanced Nursing*, *41*(4), 376–382.

Jones, C. E., Maben, J., Lucas, G., Davies, E. A., Jack, R. H., & Ream, E. (2015). Barriers to early diagnosis of symptomatic breast cancer: A qualitative study of Black African, Black Caribbean and White British women living in the UK. *BMJ Open*, *5*(3), e006944. https://doi.org/10.1136/bmjopen-2014-006944.

Keeney, S., Hasson, F., & McKenna, H. (2006). Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*, *53*(2), 205–212. https://doi.org/10.1111/j.1365-2648.2006.03716.x.

Lakhanpaul, M., Bird, D., Culley, L., Hudson, N., Robertson, N., Johal, N., McFeeters, M., Hamlyn-Williams, C., & Johnson, M. (2014). The use of a collaborative structured methodology for the development of a multifaceted intervention programme for the management of asthma (the MIA project), tailored to the needs of children and families of South Asian origin: A

community-based, participatory study. *Health Serv Deliv Res*, *2*(28). https://doi.org/10.3310/hsdr02280.

LeCompte, M. D., & Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, *52*(1), 31–60. https://doi.org/10.3102/00346543052001031.

Lincoln, Y., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Linstone, A., & Turnoff, M. (1975). cited in Powell C (2003) in the Delphi technique: Myths and realities. *Journal of Advanced Nursing*, *41*(4), 376–382.

Littig, B. (2009). Interviewing the Elite — Interviewing Experts: Is There a Difference? In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing Experts. Research Methods Series* (pp. 98–113). Palgrave Macmillan. https://doi.org/10.1057/9780230244276_5.

Marsden, J., & Bradburn, J. (2004). Patient and clinician collaboration in the design of a national randomized breast cancer trial. *Health Expectations*, *7*(1), 6–17. https://doi.org/10.1111/j.1369-7625.2004.00232.x.

Mathie, E., Wilson, P., Poland, F., McNeilly, E., Howe, A., Staniszewska, S., Cowe, M., Munday, D., & Goodman, C. (2014). Consumer involvement in health research: A UK scoping and survey. *International Journal of Consumer Studies*, *38*(1), 35–44. https://doi.org/10.1111/ijcs.12072.

Mayaka, M., & King, B. (2002). A quality assessment of education and training for Kenya's four-operating sector. *Current Issues in Tourism*, *5*(2), 112–133. https://doi.org/10.1080/13683500208667911.

Mays, N., & Pope, C. (2008). *Qualitative research in health care* (3rd ed.). Blackwell Publishing, BMJ books.

McChesney, K., & Aldridge, J. (2019). Weaving an interpretivist stance throughout mixed methods research. *International Journal of Research & Method in Education*, *42*(3), 225–238. https://doi.org/10.1080/1743727x.2019.1590811.

Meyrick, J. D. (2003) *The Delphi method and health research* (Vol. 103(1), pp. 7–16). Health Education. https://doi.org/10.1108/09654280310459112.

Mockford, C., StaniszewskaGriffiths, S. F., Griffiths, F., & Herron-Marx, S. (2011). The impact of patient and public involvement on UK NHS health care: a systematic review. *International Journal for Quality in Health Care*, *24*(1), 28–38. https://doi.org/10.1093/intqhc/mzr066.

Morris, Z. S., Wooding, S., & Grant, J. (2011). The answer is 17 years, what is the question: understanding time lags in translational research. *Journal of the Royal Society of Medicine*, *104*(12), 510–520. https://doi.org/10.1258/jrsm.2011.110180.

Morse, J. M. (2009). Mixing qualitative methods. *Qualitative Health Research*, *19*, 1523–1524. https://doi.org/10.1177/1049732309349360.

Murphy, M. K., Black, N., Lamping, D. L., McKee, C. M., Sanderson, C. F. B., Askham, J., & Marteau, T. (1998). Consensus development methods and their use in clinical guideline development. *Health Technology Assessment*, *2*(3), 88. https://doi.org/10.3310/hta2030.

Norman, D. A. (2010). The research-practice gap: The need for translational developers. *Interactions*, *17*(4), 9–12. https://doi.org/10.1145/1806491.1806494.

Noy, C. (2008). Sampling knowledge: The hermeneutics of snowball sampling in qualitative research. *International Journal of Social Research Methodology*, *11*(4), 327–344. https://doi.org/10.1080/13645570701401305.

Proctor, S., & Allen, T. (2006). *'Sampling' in the research process in nursing* (5th ed.). Blackwell Publishing.

Reid, N. (1988). The Delphi technique: Its contribution to the evaluation of professional practice. *Professional Competence and Quality Assurance in the Caring Professions*, *23*, 262.

Rennie, D. (1981). Consensus statements. *New England Journal of Medicine*, *304*, 665–666. https://doi.org/10.1056/NEJM198103123041110.

Reyens, K. M., & Hehn, E. J. (2000) *Building consensus using the policy Delphi method, policy politics and nursing practice* (Vol. 1(4), pp. 308–315). Sage Publication.

Robinson, J. (2015). *Participatory research working with adults with Aspergers on the views of adults with Aspergers living in Nottinghamshire of the services that they receive*. De Montfort University.

Robson, C. (2002). *Real world research* (2nd ed.). Blackwell Publishing.

Roulston, K. (2010). *Reflective interviewing: A guide to theory and practice*. Sage.

Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: Issues and analysis. *International Journal of Forecasting*, *15*(4), 353–375. https://doi.org/10.1016/s0169-2070(99)00018-7.

Russell, J., Fudge, N., & Greenhalgh, T. (2020). The impact of public involvement in health research: what are we measuring? Why are we measuring it? Should we stop measuring it? *Research Involvement and Engagement*, *6*(1), 1–8. https://doi.org/10.1186/s40900-020-00239-w.

Sackman, H. (1975). *Delphi critique*. Lexington Books.

Schuster, M., McGlynn, E., & Brook, R. H. (1998). How good is the quality of health care in the United States. *The Milbank Quarterly*, *76*(4), 517–563. https://doi.org/10.1111/1468-0009.00105.

Sheldon, T. A., & Parker, H. (1992). Race and ethnicity in health research. *Journal of Public Health*, *14*(2), 104–110.

Shippee, N. D., Domecq Garces, J. P., Prutsky Lopez, G. J., Wang, Z., Elraiyah, T. A., Nabhan, M., Brito, J. P., Boehmer, K., Hasan, R., Firwana, B., Erwin, P. J., Montori, V. M., & Murad, M. H. (2013). Patient and service user engagement in research: A systematic review and synthesized framework. *Health Expect*, *18*(5), 1151–1166. https://doi.org/10.1111/hex.12090.

Silverman, D. G., & Brull, S. J. (1993). The effect of a tetanic stimulus on the response to subsequent tetanic stimulation. *Anesthesia & Analgesia*, *76*(6), 1284–1287. https://doi.org/10.1213/00000539-199376060-00017.

Snape, D., Kirkham, J., Preston, J., Popay, J., Britten, N., Collins, M., Froggatt, K., Gibson, A., Lobban, F., Wyatt, K., & Jacoby, A. (2014). Exploring areas of consensus and conflict around values underpinning public involvement in health and social care research: A modified Delphi study. *BMJ Open*, *4*(1), e004217. https://doi.org/10.1136/bmjopen-2013-004217.

Snyder-Halpern, R., Thompson, C. B., & Schaffer, J. (2000). Comparison of mailed vs. Internet applications of the Delphi

technique in clinical informatics research. *Proceedings/AMIA (American Medical Informatics*.

Staley, K. (2009). *Exploring Impact: Public involvement in NHS, public health, and social care research*. INVOLVE.

Staniszewska, S., Adebajo, A., Barber, R., Beresford, P., Brady, L.-M., Brett, J., Elliott, J., Evans, D., Haywood, K. L., Jones, D., Mockford, C., Nettle, M., Rose, D., & Williamson, T. (2011). Developing the evidence base of patient and public involvement in health and social care research: The case for measuring impact. *International Journal of Consumer Studies*, *35*(6), 628–632. https://doi.org/10.1111/j.1470-6431.2011.01020.x.

Staniszewska, S., Brett, J., Mockford, C., & Barber, R. (2011). The GRIPP checklist: Strengthening the quality of patient and public involvement reporting in research. *International Journal of Technology Assessment Health Care*, *27*(4), 391. https://doi.org/10.1017/S0266462311000481.

Staniszewska, S., Brett, J., Simera, I., Seers, K., Mockford, C., Goodlad, S., Altman, D. G., Moher, D., Barber, R., Denegri, S., Entwistle, A., Littlejohns, P., Morris, C., Suleman, R., Thomas, V., & Tysall, C. (2017). GRIPP2 reporting checklists: Tools to improve reporting of patient and public involvement in research. *Research Involvement and Engagement*, *3*(1), 13. https://doi.org/10.1186/s40900-017-0062-2.

Stewart, D., Wilson, R., Selby, P., & Darbyshire, J. (2011). Patient and public involvement. *Annals of Oncology*, *22*(Suppl_7), vii54–vii56. https://doi.org/10.1093/annonc/mdr427.

Tansey, O. (2007). Process tracing and elite interviewing: A case for non-probability sampling PS. *Political Science and Politics*, *40*(4), 765–772. https://doi.org/10.1017/s1049096507071211.

Taylor, M. C. (2005). Interviewing. In I. Holloway (Ed.), *Qualitative research in health care* (pp. 29–55). Open University Press.

Todd, A. L., Nutbeam, D., Martin, B., Douglas, P., Lower, T., Rolfe, M., Monaghan, N., Hooker, C., Capon, A., Hess, I. M., & Li, B. (2018). Involving consumers in health research: What do consumers say? *Public Health Research & Practice*, *28*(2), 2821813. https://doi.org/10.17061/phrp2821813.

Törrönen, J. (2002). Semiotic theory on qualitative interviewing using stimulus texts. *Qualitative Research*, *2*(3), 343–362.

Tritter, J. Q., Barley, V., Daykin, N., Evans, S., McNeill, J., Rimmer, J., Sanidas, M., & Turton, P. (2003). Divided care and the Third Way: User involvement in statutory and voluntary sector cancer services. *Sociology of Health & Illness*, *25*(5), 429–456. https://doi.org/10.1111/1467-9566.00353.

Turoff, M. & Hiltz, S. R. 1996. in Adler and Ziglio, Gazing into the oracle Jessica Kinglsey Publishers London.

Waters, A. M., Smith, C. T., Young, B., & Jones, T. M. (2014). The CONSENSUS study: Protocol for a mixed methods study to establish which outcomes should be included in a core outcome set for oropharyngeal cancer. *Trials*, *15*(1), 1–9. https://doi.org/10.1186/1745-6215-15-168.

Wellcome Trust a good research practice guide. (2007). 'Guidelines on good research practice' available online: https://wellcome.ac.uk/funding/guidance/guidelines-good-research-practice (accessed on 169 2018).

Wilson, P. M. (2001). A policy analysis of the Expert Patient in the United Kingdom: self-care as an expression of pastoral power? *Health & Social Care in the Community*, *9*(3), 134–142. https://doi.org/10.1046/j.1365-2524.2001.00289.x.