2014

The use of citizens' juries in health policy decision-making: a systematic review

Jackie M. Street  
*University of Adelaide*

Katherine M. Duszynski  
*University of Adelaide*

Stephanie Krawczyk  
*University of Adelaide*

Annette J. Braunack-Mayer  
*University of Wollongong*, abmayer@uow.edu.au

**Publication Details**

The use of citizens' juries in health policy decision-making: a systematic review

Abstract
Deliberative inclusive approaches, such as citizen juries, have been used to engage citizens on a range of issues in health care and public health. Researchers engaging with the public to inform policy and practice have adapted the citizen jury method in a variety of ways. The nature and impact of these adaptations has not been evaluated.

We systematically searched Medline (PubMED), CINAHL and Scopus databases to identify deliberative inclusive methods, particularly citizens' juries and their adaptations, deployed in health research. Identified studies were evaluated focusing on principles associated with deliberative democracy: inclusivity, deliberation and active citizenship. We examined overall process, recruitment, evidence presentation, documentation and outputs in empirical studies, and the relationship of these elements to theoretical explications of deliberative inclusive methods.

The search yielded 37 papers describing 66 citizens' juries. The review demonstrated that the citizens' jury model has been extensively adapted. Inclusivity has been operationalised with sampling strategies that aim to recruit representative juries, although these efforts have produced mixed results. Deliberation has been supported through use of steering committees and facilitators to promote fair interaction between jurors. Many juries were shorter duration than originally recommended, limiting opportunity for constructive dialogue. With respect to citizenship, few juries' rulings were considered by decision-making bodies thereby limiting transfer into policy and practice.

Constraints in public policy process may preclude use of the 'ideal' citizens' jury with potential loss of an effective method for informed community engagement. Adapted citizens' jury models provide an alternative: however, this review demonstrates that special attention should be paid to recruitment, independent oversight, jury duration and moderation.

Keywords
review, systematic, decision-making; juries, policy, citizens', health

Disciplines
Education | Social and Behavioral Sciences

Publication Details

This journal article is available at Research Online: http://ro.uow.edu.au/sspapers/3786
The use of citizens’ juries in health policy decision-making: A systematic review

Jackie Street\textsuperscript{a,},* Katherine Duszynski\textsuperscript{b}, Stephanie Krawczyk\textsuperscript{a}, Annette Braunack-Mayer\textsuperscript{a}

\textsuperscript{a}School of Population Health, The University of Adelaide, Australia
\textsuperscript{b}School of Paediatrics & Reproductive Health, Discipline of Paediatrics, The University of Adelaide, Australia

\section*{1. Background}

Deliberative inclusive approaches, as a vehicle for citizen engagement, have particular appeal both because of the fiscal importance of health policy and because health matters touch the lives of citizens very personally. Such approaches aim to bring together diverse citizens, supported by a range of resources, to discuss issues of public concern (Davies \textit{et al.}, 2006, p.4). Some deliberative inclusive approaches methods have been well described in the theoretical literature including citizens’ juries (Parkinson, 2004; Pickard, 1998; Smith and Wales, 2000), consensus conferences (Dryzek and Tucker, 2008; Hendriks, 2005), planning cells (Hendriks, 2005) and deliberative polling (Fishkin, 1991). Others, such as World Cafe (Brown, 2001), remain primarily outside academic peer-review and critique. Some deliberative methodologists advocate combining methods in order that “the weaknesses of one would be overcome by the strengths of another” (Carson and Hartz-Karp, 2005, p.121), while others argue that their application, as originally described, is unworkable in real-world settings (Pickard, 1998).

Citizens’ juries, in particular, have undergone a process of evolution and adaptation. Developed in the 1970s, the term is a registered trademark of the Jefferson Centre (2004, p10) purportedly to “preserve the integrity of the process”. The Centre has described, essential characteristics of a citizens’ jury and, within the USA, the term has been tightly regulated. Elsewhere, it has been used much less precisely, as researchers have variously adapted the...
citizens’ jury approach. However, the nature and impact of these changes has not been documented.

Citizens’ juries offer a useful tool for engaging citizens in health policy decision-making: they are small enough to permit effective deliberation, relatively inexpensive compared to the larger deliberative exercises of planning cells and consensus conferences, yet sufficiently diverse that the citizens engaged are exposed to a broad range of public experience and perspectives. In this paper, we review use of citizens’ juries for community engagement in health research, focussing on methodological aspects. We have taken an interpretation of citizens’ juries that accepts that the term is now used more loosely and covers a broader array of activities than originally described. We examine ways in which researchers have adapted the citizens’ jury model and explore the fit between these methodological adaptations of the historic citizens’ jury and principles associated with deliberative democracy: inclusivity, deliberation and active citizenship. In this comparison, we draw on the conceptualisation of these principles described by Smith and Wales (2000).

2. Methodology

2.1. Sources

Published documents identifying studies under the broad heading of deliberative inclusive methods in health-related areas, including health care and services, were sourced for the years 1995–2010. The choice of the year 1995 reflects the onset of health authorities’ and researchers’ use of deliberative processes for citizen engagement on health issues (Abelson et al., 2003; Parkinson, 2004). Databases, Medline (PubMed), CINAHL and Scopus, were selected for their coverage of health-oriented research, as well as political and social science materials.

2.2. Search strategy development

Searches were carried out by one researcher (SK) based on criteria developed by all authors (Table 1). Search terms were selected to identify deliberative inclusive methods deployed in health research. Terms centred on public participation, as opposed to expert discussion, with the term, citizen, and its synonyms included. Terms related to deliberative processes were included, specifically names given to variations of deliberative methods (Abelson et al., 2003). The final search strategy was revised with the assistance of a research librarian. The full search is provided online (Appendix 1).

2.3. Article screening and criteria

Relevant search results (Fig. 1) were combined in a citations database. Abstracts were scanned by one author (SK) using Table 1

<table>
<thead>
<tr>
<th>Inclusion criteria:</th>
</tr>
</thead>
<tbody>
<tr>
<td>English language</td>
</tr>
<tr>
<td>Article describes use of a deliberative forum (e.g. citizen jury, citizen panel, planning cell, consensus conference) which aimed for deliberation, inclusivity and influence in policy or practice</td>
</tr>
<tr>
<td>Forum was applied to topics, activities or projects that impacted on public health, health care and health services</td>
</tr>
<tr>
<td>Participants were lay citizens or lay citizens were included in the deliberative forum</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insufficient detail provided to gauge nature of the forum</td>
</tr>
<tr>
<td>Forums not explicitly addressing a health issue</td>
</tr>
</tbody>
</table>

2.4. Analysis

Data were managed with a Microsoft Access database using a framework comprising eight domains (Table 2). The jury research questions were classified into categories (Table 3). Where questions addressed more than one category, the primary category (as assessed by the authors) was selected.

Analysis was informed by the description of citizens’ juries by Smith and Wales (2000). This framework was chosen for its focus on how key tenets of deliberative democracy, namely inclusivity, deliberation and active citizenship, play out in the implementation of the citizens’ jury model. Smith and Wales describe how citizens’ juries approximate the inclusivity ideal “by aiming for a broadly representative jury selection” (Smith and Wales, 2000, p.56) where inclusivity describes participation of “all citizens’ in public dialogue”, with all viewpoints given “equal right to be heard” (Smith and Wales, 2000, p.53). Furthermore, citizens’ juries have implemented deliberation by establishing “rules of conduct” between jurors (Smith and Wales, 2000, p.58), with this tenet described as discussion that “encourages mutual recognition and respect and is oriented towards the public negotiation of the common good” (Smith and Wales, 2000, p.53). Citizens’ juries have advanced an active value of citizenship by encouraging citizen participation in decision-making processes. Realisation of active citizenship may be bolstered by good facilitation and pre-jury contracts binding commissioning organisations to respond to jury recommendations (Smith and Wales, 2000, p.60). Attention to fulfilling the tenets of inclusivity and “egalitarian, uncoerced, competent” deliberation “free from delusion, deception, power and strategy” permits the practice of active citizenship where individuals can engage with the diverse “knowledge, experience and capabilities” of others (Smith and Wales, 2000, p.53–54). Such engagement “has the potential to transform the values and preferences of citizens in response to encounters with others” (Smith and Wales, 2000, p.54).

We thus examined overall process, recruitment, evidence presentation, documentation and outputs (e.g. reports), to explore the relationship between deliberative democratic tenets, as described by Smith and Wales, and the practice of citizens’ juries in health research.

In particular, we wished to know where and how the citizens’ jury model has been adapted to meet the needs or restrictions of a research or policy context and how such methodological changes have impacted on the outcome. The word ‘outcome’ describes the findings or verdict of the jury, reached by consensus or vote, including any record of dissent and underlying reasons for the
verdict reached. We used the following questions to inform our analysis.

1. What topic areas have been considered?
2. Where have juries been conducted?
3. What recruitment strategies have been used and how did these strategies impact on the representativeness of the demographic composition of the jury?
4. What was the duration of the juries and how were timelines structured? How did jury duration and timing impact on deliberation and the ability to provide a usable outcome in the policy or research context?
5. How were the juries conducted particularly with respect to steering groups, moderation and structured deliberation?
6. How were expert witnesses selected and expert testimony presented?
7. How have jury recommendations been formulated, reported and disseminated?

3. Findings

Our findings indicate that, although the citizens’ jury method originally described a tightly prescribed method, this precision has been lost over time and currently the term describes a broad array of methodological approaches. In all, 37 papers were identified describing 66 citizens’ juries within 28 studies. The term, ‘citizens’ jury’, or a variant thereof was explicitly used to describe the method in 22 studies. Of the remaining six, the methods employed were similar, although four (Abelson et al., 2007a; Huston, 2004; Maxwell et al., 2003; Willison et al., 2008) also drew on the work of Yankelovich (1991). Of 66 juries, 30 involved examining the same question at multiple sites, 17 related to a single jury and question, 11 involved multiple questions conducted at multiple sites in a single study and 8 were jury pairs conducted at the same site examining the same question.

Two thirds of studies (n = 18) were overseen by a steering committee or advisory group although the composition and influence of these groups varied widely. Usually the group consisted of key stakeholders, particularly policy-makers, but could also include discipline experts, advocacy group representatives, clinical practitioners, deliberative methodologists, patients and caregivers. The role of the groups was variously described as: to prevent bias in expert presentation (Pidgeon and Rogers-Hayden, 2007), to guide question development and evidence presentation (Dunkerley and Glasner, 1998; Parkin and Paul, 2011; Pickard, 1998; Rogers et al., 2009; Toni & von Braun, 2001), general oversight (Dunkerley and Glasner, 1998; Maxwell et al., 2003), to engage stakeholder representatives (Haigh and Scott-Samuel, 2008; Huston, 2004; Lenaghan, 1999; Lenaghan et al., 1996) and to disseminate or implement findings (Gooberman-Hill et al., 2008; Kashefi and Mort, 2004).

Five juries reported jury costings (Einsiedel, 2002; Elwood and Longley, 2010; Iredale et al., 2006; Lenaghan, 1999; Maxwell et al., 2003).

3.1. Topic areas

A large proportion of the 66 juries (9 studies comprising 25 juries) addressed population based ethical issues but a range of other issues were discussed (Table 3).

3.2. Location

More than half (38/66) of the juries were conducted in Canada, 16 in the UK, seven in Australia or New Zealand, three in the USA.
and one each in Brazil and Italy. Nearly half were conducted within four Canadian studies characterised by multiple juries addressing a single question (Einsiedel, 2002; Huston, 2004; Maxwell et al., 2003; Willison et al., 2008).

3.3. Recruitment strategies

Around 20 different recruitment strategies were used with the most common (11 studies/29 juries) being stratified random sampling through the electoral roll (Bennett and Smith, 2007; Haigh and Scott-Samuel, 2008; Oliver-Weymouth, 2000; Parkin and Paul, 2011; Paul et al., 2008), random digit dialling (Secko et al., 2009) or commercial database of registered telephone numbers (Menon and Stafinski, 2008) or national polling institute (Carson, 2006). Three studies (9 juries) used non-stratified random sampling through electoral roll, random digit dialling or survey response (Kim et al., 2009; Paul et al., 2008; Willison et al., 2008). Recruitment by a market research company was used in six studies (11 juries) (Dunkerley and Glasner, 1998; Elwood and Longley, 2010; Huston, 2004; Kashefi and Mort, 2004; Lenaghan, 1999; Rogers et al., 2009). One jury (Kashefi and Mort, 2004), employed a professional recruiter who directly recruited individuals at public sites. Newspaper advertisements were used to recruit women with incontinence (Herbison et al., 2009) while word-of-mouth or advertising through networks was used to recruit youth, aged, caregiver or marginalised population groups (Iredale et al., 2006; Kim et al., 2010; Pidgeon and Rogers-Hayden, 2007; Timotijevic and Raats, 2007). Other juries recruited through community organisations, government departments or existing citizens’ council (Abelson et al., 2007a; Gooberman-Hill et al., 2008; Mort and Finch, 2005; Toni & von Braun, 2001).

Most studies explicitly stated that they aimed to recruit a jury, descriptively representative of the community (Button and Mattson, 1999; Carson, 2006; Dunkerley and Glasner, 1998; Einsiedel and Ross, 2002; Elwood and Longley, 2010; Haigh and Scott-Samuel, 2008; Huston, 2004; Lenaghan, 1999; Lenaghan et al., 1996; Menon and Stafinski, 2008; Rogers et al., 2009; Secko et al., 2009), providing a cross-section of community perspectives (Einsiedel and Ross, 2002; Maxwell et al., 2003) or incorporating diverse voices (Bennett and Smith, 2007; Gooberman-Hill et al., 2008). Where criteria were stated (15 studies), both age and sex were used as stratifying variables in all studies, bar one (Haigh and Scott-Samuel, 2008), which used geographic area only. Race/ethnicity and education were each used in five studies; and at least one of employment status, housing tenure, religion and occupation was used in three studies. Geographic location, socioeconomic status, income, social class, car access, health parameters, children and language were used in only one or two studies.

No juries specifically recruited patients although Herbison et al. (2009) recruited women with incontinence: half had...
sought medical help and therefore could be considered patients. Four studies (five juries) recruited particular age groups (Iredale et al., 2006; Kim et al., 2009; Paul et al., 2008; Timotijevic and Raats, 2007), two studies (six juries) representatives from community organisations (Abelson et al., 2007a; Mort and Finch, 2005) and one each recruited carers (Kim et al., 2010), poor/unemployed citizens (Toni & van Braun, 2001) and ‘marginalised’ persons (Pidgeon and Rogers-Hayden, 2007). In one jury, the recruitment profile, developed by a steering group to provide diverse community voices, resulted in a jury composed of primarily unemployed and disadvantaged participants (Kashefi and Mort, 2004).

In six studies, conflicts of interest were gauged during recruitment and particular groups actively excluded: on the grounds of previous or current employment (Maxwell et al., 2003; Menon and Stafinski, 2008; Mooney and Blackwell, 2004; Timotijevic and Raats, 2007; Toni & van Braun, 2001), previous disease experience (Elwood and Longley, 2010; Paul et al., 2008), prior topic knowledge, or membership of a patient advocacy group (Menon and Stafinski, 2008). In one jury, to minimise recruitment of participants with strongly held prior beliefs, participants were not told of the topic until the day of the jury (Gooberman-Hill et al., 2008).

Nine juries (nine juries) (Bennett and Smith, 2007; Elwood and Longley, 2010; Paul et al., 2008), prior topic knowledge, or membership of a patient advocacy group (Menon and Stafinski, 2008). In one jury, to minimise recruitment of participants with strongly held prior beliefs, participants were not told of the topic until the day of the jury (Gooberman-Hill et al., 2008).

In shorter juries, some participants complained of insufficient detail to determine if facilitation occurred (Haigh and Scott-Samuel, 2008; Kim et al., 2010; Maxwell et al., 2003). Most studies recruited to achieve the recommended composition of 12–25 participants and most succeeded. Six studies reported juries with 9–11 participants (Huston, 2004; Mort and Finch, 2005; Parkin and Paul, 2011; Paul et al., 2008; Timotijevic and Raats, 2007; Toni & van Braun, 2001), primarily due to late drop-out of participants. Four studies did not report individual jury numbers (Haigh and Scott-Samuel, 2008; Huston, 2004; Mooney and Blackwell, 2004; Willison et al., 2008).

Most studies reported neither the number of individuals contacted for participation, nor the size of the population from which the sample was drawn. Therefore, in most cases, it was not possible to assess participation or attrition rates and relate these to recruitment strategy. Nine studies did report withdrawal numbers, which ranged from 1 to 10 individuals. Attrition is an important factor in bias since those dropping out tend to reflect those harder to recruit initially (e.g. Timotijevic and Raats, 2007).

3.4. Duration and timing

Nearly two-thirds of the juries took place over the equivalent of 1–2 days, usually on a weekend, considerably fewer than the 4–5 days recommended by the Jefferson Centre (2004) or UK Institute for Public Policy Research (Coote and Lenaghan, 1997). The longest jury (Kashefi and Mort, 2004) was conducted over five consecutive weekdays with all but two participants unemployed or retired. In most, the days were consecutive but one jury was held over two weekends (Secko et al., 2009), in another, through brief sessions over five weeks (Pidgeon and Rogers-Hayden, 2007) and one jury met 11 times over 16 weeks (Gooberman-Hill et al., 2008).

In shorter juries, some participants complained of insufficient time to explore the issues but brief daylong juries still delivered outcomes (Abelson et al., 2007a; Carson, 2006; Huston, 2004; Kim et al., 2010; Maxwell et al., 2003; Mooney and Blackwell, 2004; Mort and Finch, 2005; Willison et al., 2008). Jury length did not appear to impact on recruitment bias; that is, longer juries were balanced in terms of the selection criteria, providing measures were taken to support recruitment of hard-to-engage groups. Longer juries did permit participants greater control over the ensuing report (Dunkerley and Glasner, 1998; Iredale et al., 2006; Lenaghan, 1999; Parkin and Paul, 2011) and provided opportunity to engage with different forms of evidence. For example, one five-day jury (Kashefi and Mort, 2004) included video, witnesses and a creative writing exercise. Juries held in staggered fashion encountered problems associated with the disjointed nature of multiple sessions held over a long period. For example, in one case, experts presenting at different times created confusion when their evidence conflicted (Pidgeon and Rogers-Hayden, 2007).

3.5. Moderation

Only three studies indicated that the jury was not facilitated (Herbison et al., 2009; Pickard, 1998; Willison et al., 2008); in two, jury members lead discussion (Herbison et al., 2009; Pickard, 1998) whereas the other was guided by a workbook (Willison et al., 2008). Three studies provided insufficient detail to determine if facilitation occurred (Haigh and Scott-Samuel, 2008; Mooney and Blackwell, 2004; Mort and Finch, 2005). Even where facilitation was described, the nature of the facilitation – particularly the facilitators’ independence, training and experience – was not always clear. Nine studies described using trained, skilled or experienced facilitators (Dunkerley and Glasner, 1998; Huston, 2004; Iredale et al.,

Toni & van Braun, 2001), caregivers (Kim et al., 2010) and women with symptoms of incontinence (Herbison et al., 2009).

Only three studies reported juries with more than 25 participants, each with around 40 participants (Carson, 2006; Kim et al., 2010; Maxwell et al., 2003). Most studies recruited to achieve the recommended composition of 12–25 participants and most succeeded. Six studies reported juries with 9–11 participants (Huston, 2004; Mort and Finch, 2005; Parkin and Paul, 2011; Paul et al., 2008; Timotijevic and Raats, 2007; Toni & van Braun, 2001), primarily due to late drop-out of participants. Four studies did not report individual jury numbers (Haigh and Scott-Samuel, 2008; Huston, 2004; Mooney and Blackwell, 2004; Willison et al., 2008).

Most studies reported neither the number of individuals contacted for participation, nor the size of the population from which the sample was drawn. Therefore, in most cases, it was not possible to assess participation or attrition rates and relate these to recruitment strategy. Nine studies did report withdrawal numbers, which ranged from 1 to 10 individuals. Attrition is an important factor in bias since those dropping out tend to reflect those harder to recruit initially (e.g. Timotijevic and Raats, 2007).

3.4. Duration and timing

Nearly two-thirds of the juries took place over the equivalent of 1–2 days, usually on a weekend, considerably fewer than the 4–5 days recommended by the Jefferson Centre (2004) or UK Institute for Public Policy Research (Coote and Lenaghan, 1997). The longest jury (Kashefi and Mort, 2004) was conducted over five consecutive weekdays with all but two participants unemployed or retired. In most, the days were consecutive but one jury was held over two weekends (Secko et al., 2009), in another, through brief sessions over five weeks (Pidgeon and Rogers-Hayden, 2007) and one jury met 11 times over 16 weeks (Gooberman-Hill et al., 2008).

In shorter juries, some participants complained of insufficient time to explore the issues but brief daylong juries still delivered outcomes (Abelson et al., 2007a; Carson, 2006; Huston, 2004; Kim et al., 2010; Maxwell et al., 2003; Mooney and Blackwell, 2004; Mort and Finch, 2005; Willison et al., 2008). Jury length did not appear to impact on recruitment bias; that is, longer juries were balanced in terms of the selection criteria, providing measures were taken to support recruitment of hard-to-engage groups. Longer juries did permit participants greater control over the ensuing report (Dunkerley and Glasner, 1998; Iredale et al., 2006; Lenaghan, 1999; Parkin and Paul, 2011) and provided opportunity to engage with different forms of evidence. For example, one five-day jury (Kashefi and Mort, 2004) included video, witnesses and a creative writing exercise. Juries held in staggered fashion encountered problems associated with the disjointed nature of multiple sessions held over a long period. For example, in one case, experts presenting at different times created confusion when their evidence conflicted (Pidgeon and Rogers-Hayden, 2007).

3.5. Moderation

Only three studies indicated that the jury was not facilitated (Herbison et al., 2009; Pickard, 1998; Willison et al., 2008); in two, jury members lead discussion (Herbison et al., 2009; Pickard, 1998) whereas the other was guided by a workbook (Willison et al., 2008). Three studies provided insufficient detail to determine if facilitation occurred (Haigh and Scott-Samuel, 2008; Mooney and Blackwell, 2004; Mort and Finch, 2005). Even where facilitation was described, the nature of the facilitation – particularly the facilitators’ independence, training and experience – was not always clear. Nine studies described using trained, skilled or experienced facilitators (Dunkerley and Glasner, 1998; Huston, 2004; Iredale et al.,

Toni & van Braun, 2001), caregivers (Kim et al., 2010) and women with symptoms of incontinence (Herbison et al., 2009).
The facilitator role was variously defined as: drafting a proposal for common ground (Huston, 2004); being neutral in content but active in process (Kim et al., 2010, 2009); ensuring discussion stayed on-topic (Dunkerley and Glasner, 1998; Lenaghan, 1999; Lenaghan et al., 1996) and assisting question formulation and reaching for consensus (Toni & von Braun, 2001). In most cases, however, the facilitator role was undefined. Two juries included a resource person whose role was to investigate questions raised by the jury (Kashefi and Mort, 2004; Willison et al., 2008).

Structured elements, used to stimulate and guide discussion, were rarely discussed: 17 studies described small group work and 10 used scenarios or hypotheticals. Other techniques included: scoring methods (Menon and Stafinski, 2008), priority setting (Herbison et al., 2009; Menon and Stafinski, 2008), workbooks (Maxwell et al., 2003; Secko et al., 2009; Willison et al., 2008), dialogue guide (Huston), voting (Bennett and Smith, 2007; Pickard, 1998), physical model (Burgess et al., 2008) and a courtroom format (Bennett and Smith, 2007).

3.6. Selection of expert witnesses and nature of expert testimony

Most studies included a range of expert testimony, described either the nature of the expert evidence or the presenters’ expertise and indicated that participants could engage with presenters and challenge the evidence (See online Appendix 2). In most cases, neither the mode nor the reasons for expert selection were discussed although four studies described involving the steering group in these decisions (Dunkerley and Glasner, 1998; Gooberman-Hill et al., 2008; Iredale et al., 2006; Parkin and Paul, 2011). Three Canadian studies (Huston, 2004; Maxwell et al., 2003; Willison et al., 2008) and possibly Abelson et al. (2007a) did not use expert witnesses but instead utilised a workbook to provide balanced relevant information. Two studies undertook local research to present to the jury (Haigh and Scott-Samuel, 2008; Kashefi and Mort, 2004). Written material was provided to jurors in 14 studies with seven juries provided with material in advance. One study required participants to prepare questions (Kim et al., 2010, 2009). Seven studies included ethical analysis (Bennett and Smith, 2007; Einsiedel, 2002; Huston, 2004; Kim et al., 2010, 2009; Parkin and Paul, 2011; Rogers et al., 2009; Secko et al., 2009). Very few studies specifically described aiming for balanced evidence presentation (Iredale et al., 2006; Mooney and Blackwell, 2004; Paul et al., 2008; Pidgeon and Rogers-Hayden, 2007; Rogers et al., 2009), although this may be surmised from the discussion in other studies.

3.7. Output formulation, reporting and dissemination of jury recommendations and findings

Various strategies were used to collect data. Proceedings were audio-recorded in half the studies. Other methods included: contemporaneous notes by organisers (Abelson et al., 2007a; Rogers et al., 2009; Willison et al., 2008) or participants (Einsiedel, 2002) workbooks (Huston, 2004; Maxwell et al., 2003), video-recording (Button and Mattson, 1999; Iredale et al., 2006), whiteboard scribing (Herbison et al., 2009; Rogers et al., 2009), flip charts (Dunkerley and Glasner, 1998; Herbison et al., 2009; Huston, 2004; Secko et al., 2009), voting (Pickard, 1998), participant diaries (Iredale et al., 2006), participant hand-held video-recording (Iredale et al., 2006), questionnaires (Iredale et al., 2006) and interviews (Iredale et al., 2006). Data were analysed qualitatively in six studies using ‘content analysis’ (Menon and Stafinski, 2008; Timotijevic and Raats, 2007), ‘close and repeated reading’ (Bennett and Smith, 2007), discourse analysis (O’Doherty and Burgess, 2009), ‘coding’ (Willison et al., 2008) or without method description (Haigh and Scott-Samuel, 2008).

Recommendations were developed through consensus (Huston, 2004; Menon and Stafinski, 2008; Parkin and Paul, 2011; Rogers et al., 2009), consensus with minority opinion (Bennett and Smith, 2007; Kim et al., 2010, 2009; O’Doherty and Burgess, 2009; Pidgeon and Rogers-Hayden, 2007) or voting (Paul et al., 2008). In one study, no decision choice dominated (Willison et al., 2008). Some studies failed to describe how recommendations were reached and whether decisions were unanimous. Jury reports were variously written by researchers based on participants’ recommendations (Rogers et al., 2009), by jurors alone (Parkin and Paul, 2011) or with facilitator assistance (Dunkerley and Glasner, 1998; Elwood and Longley, 2010; Herbison et al., 2009) or by researchers in consultation with participants (Elwood and Longley, 2010; Kashefi and Mort, 2004; Lenaghan, 1999).

In addition to peer reviewed articles, 13 described a grey-literature report. Dissemination through other outlets (13 studies) included: media coverage (Abelson et al., 2007a; Dunkerley and Glasner, 1998; Haigh and Scott-Samuel, 2008; Iredale et al., 2006), direct presentation to decision makers (Abelson et al., 2007a; Dunkerley and Glasner, 1998; Iredale et al., 2006), direct community engagement (Kashefi and Mort, 2004) and through planning meetings (Abelson et al., 2007a; Haigh and Scott-Samuel, 2008; Kashefi and Mort, 2004). Only three described commitment by a decision-making body to consider recommendations (Lenaghan, 1999; O’Doherty and Burgess, 2009; Pickard, 1998). One jury fell within formal health technology assessment (Menon and Stafinski, 2008).

4. Discussion

Citizen juries permit citizens to engage with evidence, deliberate and deliver recommendations on a range of complex and demanding topics. This review demonstrates that, provided the research question is tightly defined and concrete in nature, even brief one-day citizens’ juries can deliver useable outcomes to inform policy and practice.

Many juries described in this review were conducted to meet instrumental aims, that is, to deliver recommendations usable in policy and practice decision-making. It was often not possible, in our review, to establish the extent to which juries succeeded in this respect, primarily because the final link between the jury and decision-makers is not well established. However, in the tradition from which citizens’ juries arise, instrumental goals are less important and goals related to empowering citizens have greater prominence.

Smith and Wales’ (2000) work on citizens’ juries supports further inquiry into the relationship between health research and philosophical aims. Given there is no consensus on what constitutes a deliberative process ‘in theory or in practice’ (Blacksher et al., 2012, pp14), Smith and Wales provide a framework to consider the extent to which health researchers engage with deliberative theory in their work on citizens’ juries. By examining how citizens’ juries implement inclusivity, deliberation and citizenship, the following discussion reveals a tension between research aims and deliberative ideals. We further suggest that improvements in reporting the practice of citizens’ juries could produce insights relevant to the ‘macro-political uptake’ of deliberative processes (Goodin and Dryzek, 2006), and strengthen dialogue between deliberative practitioners and theoreticians.

4.1. Inclusivity

Inclusivity refers to efforts to ensure that citizens’ juries represent a wide variety of experience and backgrounds (Smith and
Most of the juries aimed to recruit a jury descriptively representative of the community from which the jurors were drawn, suggesting that, at least in intent, inclusivity was valued. However, a significant minority of studies did not sufficiently report on the jury’s composition to allow a judgement about inclusivity to be made, implying either that inclusivity was taken for granted or, more likely, concern for inclusivity was not built into the study design.

The extent to which recruitment strategies succeeded in creating an inclusive environment varied considerably. In general, studies systematically seeking to include people from a wide variety of backgrounds — by using stratification, recruitment using market researchers, and honoraria, tended to be more successful in recruiting diverse voices than those relying on random sampling. Ironically, this finding suggests that the rationale for random sampling — the idea that each person from a population has an equal chance of being selected, with the randomly-selected group understood to represent the population in microcosm (Carson and Hartz-Karp, 2005) — appears not to work for the small sample size of a jury. Even in studies which appear to have recruited a jury intuitively representative of a broader population, the lack of information about the populations from which jurors were drawn, makes it difficult to draw conclusions about the extent to which these juries were truly inclusive. In addition, having a jury of fewer than 12 persons, may impact on the ability of the jury to support the diverse range of community perspectives and experience required for inclusivity.

The comments above presume that inclusivity turns only on jury composition. However, scholars have suggested that meeting the criterion of inclusivity also requires attention to other aspects. For example, inclusivity may be implemented through presentation of diverse witnesses (Smith and Wales, 2000), both for content and diverse viewpoints, through participants’ presentation of narrative that captures relevant personal experiences (Burkhalter et al., 2002; Young, 1996) and by allowing juries opportunity to challenge presented evidence and request additional information. Our findings indicate that at least some studies attempted to meet these conditions but it is also apparent that relatively little attention was paid to this criterion, such that, in many cases, it was difficult to judge if the criterion of inclusivity had been met.

4.2. Deliberation

Deliberation refers to “communication that induces reflection on preferences, values and interests in a non-coercive fashion” (Mansbridge et al., 2010, p.2) but which in the deliberative democracy context reaches for consensus. The concept is drawn from deliberative democratic theory, which explores the possibility of democracy through the mode of the “public forum... oriented towards the common good” (Hendriks, 2002, p8) with forum proceedings equating to “reasoned public discussion amongst equals” who are in pursuit of “reaching understanding” such that “legitimate decisions” can be made (Hendriks, 2002, p8).

The citizens’ juries reported in this paper created a positive environment for deliberation in a range of ways. First, two-thirds of the studies were overseen by a steering or advisory committee. Smith and Wales (2000) suggest that “the very integrity” of the process is dependent on decisions made early in the development process and that such decisions must be seen to be unbiased. However, even when steering group composition was described in detail, bias was difficult to judge since the group could replicate imbalance in society.

Second, 21 of the 28 studies indicated that their juries were facilitated by a person other than a jury member and only three explicitly indicated the juries were not facilitated. However, not all studies described the facilitators as ‘trained’ or ‘independent’, both factors which may be important for providing an unbiased and supportive process. In addition, the role of the moderator was generally not well defined.

Thirdly, all but three of the juries described a jury size of fewer than 25. Keeping group size small appears to provide greatest opportunity for participants to receive equal and adequate speaking time (Smith and Wales, 2000). The way in which decisions were reached and communicated also helped to support deliberative dialogue. The dominant method through which recommendations were developed was consensus or consensus with minority opinion, suggesting that considerable care had been given to ensuring that jurors were able to participate effectively in the jury.

Smith and Wales (2000) note that the short length of juries can limit the opportunity for jurors to develop and maintain open and constructive dialogue. Almost two-thirds of the juries reported in this paper were conducted over 1–2 days, arguably, too short a time for jurors to develop the collaborative environment needed for deliberative dialogue.

4.3. Active citizenship

Citizens’ juries have been promoted as vehicles to “reassert the importance of a more active form of citizenship” (Smith and Wales, 2000, p60). It is clear that the term, citizen, as used by Smith and Wales, and interpreted as such in most of the studies, more closely approximates the Athenian ideal rather than its popular use synonymous with nationality. In only six of the juries was juror selection dependent on formal requirement for ‘citizenship’ of the country in which the study took place. Recognition of the active citizenship role, within a jury, might require that, at a minimum, there be formal recognition of jurors’ perspectives or recommendations by appropriate authorities.

Defined in this way, only a small number of the identified studies met the citizenship criterion. Only three studies described commitment by a decision-making body to consider recommendations (Lenaghan, 1999; O’Doherty and Burgess, 2009; Pickard, 1998). Given the relationship researchers have with governments and other decision-makers, this is not surprising. Most researchers recognise that it is beyond their power to enact participants’ recommendations (Carson and Hartz-Karp, 2005). However, they also recognise the importance of addressing juror-policy maker interactions: citizens have reported the need for greater ‘accountability’ by decision-makers who support deliberative forums so that citizens can feel that “their input is wanted and is going to be needed” (Abelson et al., 2004, p209–10). This review shows that researchers use a range of techniques to promote the findings of the citizens’ juries including academic literature, media channels and direct engagement with decision-makers.

There are other ways in which juries can enhance the citizen role, particularly in capacity building for empowered citizens (Abelson et al., 2007b). From our own experience (Braunack-Mayer et al., 2010; Rogers et al., 2009; Watt et al., 2012), jurors are often impressive, both in their commitment and their capacity to grasp complex topics. Our experience was mirrored in some included juries: for example, Gooberman-Hill et al. (2008, p.280) comment that “deep engagement” of the jurors “manifested itself in commitment to the process and continued interest in research”.

5. Conclusion

This review demonstrates that the citizens’ jury model has been extensively adapted from its earliest forms. Inclusivity has been operationalised with sampling strategies that aim to recruit representative juries, although these efforts have produced mixed
results. Various strategies have been implemented to support deliberation, including enlistment of steering committees and facilitators to promote fair interactions between jurors. However, citizens’ juries generally appear to be conducted over a shorter length of time than originally recommended, thereby limiting opportunities for reflection on the preferences, values and interests of others. Finally, fewer studies have honoured the citizenship criterion in jury designs, with only a handful of juries’ rulings considered by decision-making bodies.

In light of these findings, the ‘ideal’ process promulgated by Smith and Wales (2000), Crosby (The Jefferson Center, 2004) and others might be considered to be undermined by deliberative experiments in health research. Such a view pays inadequate attention to the administrative, financial and political constraints that shape the design and implementation of many community engagement strategies in real world settings. Indeed, restricting the use of citizens’ juries to a narrowly defined set of parameters may preclude their use in policy processes or to inform practice reform. Purity about the nature of the ideal deliberative process, such that the methods are untenable within the constraints of the decision-making process, may impose a considerable loss to community engagement. Strict adherence to, and, in particular, legal regulation of a methodology, through patent or trademark, could be counterproductive for knowledge production since it is through testing and adapting methodologies that new ideas are developed and our understanding expands.

Regardless of the gap between ideal and more pragmatic approaches to the use of citizens’ juries, it is still important to understand how various methodological decisions can shape jury processes and outcomes. It is clear that some factors are essential to provide an unbiased inclusive deliberative process. This review indicates that, in adapting the citizen jury to instrumental aims, particular care and attention should be paid to recruitment methods, independent oversight by a steering committee, jury duration, moderation and respect for the participant volunteer. The use of adapted deliberative inclusive processes for instrumental means, and, in particular, to inform health policy, is a developing field. Careful attention to reporting and improved evaluation of process and outcomes can only assist in ensuring that these methods are best designed to meet both democratic and instrumental goals.

Acknowledgements

We gratefully acknowledge the work of Monira Osman and Yee Ling Kok in undertaking the full-text review of selected articles and Vicki Xafis for discussions assisting the review. Jackie Street is supported by an NHRMC Capacity Building Grant 565501 and an Australian National Preventive Health Agency Fellowship.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.socscimed.2014.03.005.

References


