University of Wollongong

Research Online

Faculty of Science, Medicine and Health -
Papers: part A

Faculty of Science, Medicine and Health

1-1-2013

Developing 'policy stories' for state health system benchmarking: a small-N quali-quantitative study

Erica Bell
University of Tasmania

Bastian Seidel
University of Tasmania, bastian@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/smhpapers>

Part of the [Medicine and Health Sciences Commons](#), and the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Bell, Erica and Seidel, Bastian, "Developing 'policy stories' for state health system benchmarking: a small-N quali-quantitative study" (2013). *Faculty of Science, Medicine and Health - Papers: part A*. 1039.
<https://ro.uow.edu.au/smhpapers/1039>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Developing 'policy stories' for state health system benchmarking: a small-N qualitative study

Abstract

Background: The benchmarking literature has made important advances and offers many different population health indicators that can be used to compare state health systems. However, there is still a need for qualitative, complexity-oriented approaches that allow policy-makers to develop explanatory 'policy stories' from combining such indicators that are useful to policy solutions. **Methods:** A new qualitative method from the social sciences based on Boolean approaches, called Qualitative Comparative Analysis (QCA), was piloted in a 'real world' policy consultancy to combine Australian state-level indicators of community and health system inputs, interventions, and population health outcomes. Analyses were provided for state inputs and outcomes in a specific area of chronic disease (mental health), along with state profiling for differences in risky health-related behaviours. **Results:** The QCA method suggested that the state of Tasmania may resemble South Australia in terms of having lower community inputs, as well as higher levels of mental health system inputs and interventions (such as prescriptions), than other states with the outcome of higher self-reported psychological distress. Theoretically, employment levels explained state-level differences in self-reported psychological distress. In terms of risky health-related behaviours, the QCA suggested that Tasmania leads other states in both socio-economic challenges and risky health behaviours. Theoretically, smoking explained state-level differences in self-assessed health. **Conclusions:** The QCA method has its weaknesses, but in this study, it allowed for the development of policy stories based on systematic comparisons of different states. It also suggested theoretically plausible explanations for differences in state-level outcomes.

Keywords

study, health, state, developing, quantitative, policy, quali, stories, n, small, benchmarking, system

Disciplines

Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

Bell, E. & Seidel, B. (2013). Developing 'policy stories' for state health system benchmarking: a small-N qualitative study. *International Journal on Disability and Human Development*, 12 (1), 67-75.

Erica Bell* and Bastian Seidel

Developing ‘policy stories’ for state health system benchmarking: a small-*N* quali-quantitative study

Abstract

Background: The benchmarking literature has made important advances and offers many different population health indicators that can be used to compare state health systems. However, there is still a need for qualitative, complexity-oriented approaches that allow policy-makers to develop explanatory ‘policy stories’ from combining such indicators that are useful to policy solutions.

Methods: A new qualitative method from the social sciences based on Boolean approaches, called Qualitative Comparative Analysis (QCA), was piloted in a ‘real world’ policy consultancy to combine Australian state-level indicators of community and health system inputs, interventions, and population health outcomes. Analyses were provided for state inputs and outcomes in a specific area of chronic disease (mental health), along with state profiling for differences in risky health-related behaviours.

Results: The QCA method suggested that the state of Tasmania may resemble South Australia in terms of having lower community inputs, as well as higher levels of mental health system inputs and interventions (such as prescriptions), than other states with the outcome of higher self-reported psychological distress. Theoretically, employment levels explained state-level differences in self-reported psychological distress. In terms of risky health-related behaviours, the QCA suggested that Tasmania leads other states in both socio-economic challenges *and* risky health behaviours. Theoretically, smoking explained state-level differences in self-assessed health.

Conclusions: The QCA method has its weaknesses, but in this study, it allowed for the development of policy stories based on systematic comparisons of different states. It also suggested theoretically plausible explanations for differences in state-level outcomes.

Keywords: benchmarking; chronic disease; state health policy.

Bastian Seidel: Clinical Senior Lecturer, School of Medicine, University of Tasmania, Tasmania, Australia

Introduction

Benchmarking is a relatively recent phenomenon that implies the comparison of often separate but competing agencies as part of what is originally a marketplace approach. Differences in supply, demand, and marketplace conditions are analysed in ways that broaden academic understanding of the strengths, weaknesses and future directions for improvement. These directions can include [1] how productivity should be measured and decided and the efficient deployment of service resources [2]. A critical focus of the benchmarking of state health systems is reducing unequal health outcomes related to socio-economic disparities and providing access to good healthcare quality [3, 4]. With the correct method, benchmarking can help achieve both better outcomes and better processes for quality, in such very specific areas such as hypertensive disorders [5, 6] or communicable disease surveillance systems [7].

National and state systems show wide variation in diverse critical areas, such as health outcomes, quality, access, equity, and efficiency [8]. Attempts at benchmarking state systems have, therefore, often focussed on a range of indicators directed at identifying ‘top performing states’, which are useful to state goal-setting. The Commonwealth Fund (a private foundation) in the USA has produced state scorecards and state child health scorecards, which allow users to manipulate data to view rankings and supporting data for state comparisons (see <http://www.commonwealthfund.org/Maps-and-Data/State-Data-Center/State-Scorecard.aspx>). It has also published analyses of a 2008 comparison of state system performance using 13 indicators of child health system performance grouped into five domains (i.e., access, quality, costs, equity, and measures to do with ‘potential to lead healthy lives’). The Commonwealth Fund also uses the example of ‘high performing states’ to argue that high performance is

*Corresponding author: Erica Bell, PhD, Deputy Director, University Department of Rural Health, University of Tasmania, Hobart, Tasmania 7001, Australia, Phone: +61 (0)3 6226 7377, Fax: +61 (0)3 6226 1952, E-mail: Erica.Bell@utas.edu.au

possible, as well as to identify desirable and less than desirable regional differences in ‘child health system performance’ [9].

On the one hand, closer inspection of these ostensibly ‘high performing states’ yields useful information about their positive features, such as a focus on the continuity of reform and congruency of policies [10]. On the other hand, simplistic use of state scorecards ignores the fact that they may be confounded by market, political, and cultural differences among states; in turn, these factors can influence how well the state systems perform. Therefore, best practice in using population health data in state benchmarking should focus more on learning about the similarities and differences between states (i.e., differences in context that greatly shape interpretability of state outcomes) than developing ‘league tables’ of states [11]. These similarities and differences between states can be about system ‘inputs’ as much as system interventions and outcomes [12].

Therefore, it could be argued that a key challenge in benchmarking is not simply the development of indicators, but rather the *combining* of indicators into suites of indicators about systems inputs, processes and population outcomes that can tell a ‘policy story’. A policy story in this context is defined as a narrative – a qualitative account – of causality that gives insights as to why different state health systems achieve varied population health outcomes. Since the foundational work of Stone [13] and Majone, [14] the health policy literature has long suggested that such policy stories have a powerful role to play in policy development. That is, policy ideas about what causes a situation can shape agreement about the policy solutions [15]. Clearly, there are degrees to which such policy stories may be based on sound evidence. If they are evidence-based, the challenges of systematically combining indicators from different states into a coherent qualitative and causality-oriented account of state differences are considerable. The current study aims to explore how an approach from the social sciences, hereby called Qualitative Comparative Analysis (QCA), can help ensure that policy-makers have better tools for combining diverse indicators about state health systems in ways that tell a policy story about state-level differences. In so doing, the study aims to offer illustrative analyses from a ‘real world’ consultancy that developed a method for using existing data to benchmark (decide what indicators have been met to what level) state health system performance. In this sense, this study is translational: it focuses on how to create a better interface between evidence and policy-makers.

Qualitative comparative analysis

QCA has been developed by social scientists as an adjunct method for comparing small to intermediate numbers of cases (ideally around 50). The cases can be of any kind; in this study a state system is a ‘case’. The key feature of QCA is that it is qualitative: it is about generating qualitative causal descriptions of differences and similarities between cases. The QCA method can be considered an approach to achieving descriptive comparative information in a systematic manner.

Therefore, in QCA, the unit of analysis is the individual case. Every case in QCA has a configuration of possibly causal conditions of interest and an outcome of interest. Groups of cases can have observed configurations that are similar or different in many different respects. QCA uses Boolean-based approaches to summarise all these configurations of cases in the form of ‘logical equations’. The logical equations produced by the QCA method offer shorthand expressions of the different configurations of cases with a particular outcome of interest. Therefore, a QCA ‘logical equation’ appears in the form of a summary configuration (or string of configurations if the cases being described suggest complex causality). The method by which many configurations are reduced to single configurations (often with software) is described by Cronqvist as follows. In this quotation, Cronqvist explains how a variant of QCA he has developed (and which is used in this study) called Multi-Value QCA (MVQCA) builds on the approach developed by QCA founder Ragin:

‘...the most fundamental rule of Boolean reduction as expressed by Ragin can be rewritten for multi-value reduction:

- If two Boolean expressions differ in only one causal condition yet produce the same outcome, then the causal condition that distinguishes the two expressions can be considered irrelevant and can be removed to create a simpler, combined expression’ (Ragin 1987: p. 93).

For multi-value reduction this can be written as:

- ‘If all n multi-value expressions ($c_0\Phi, \dots, c_{n-1}\Phi$) differ only in the causal condition C while all n possible values of c yet produce the same outcome, then the causal condition C that distinguishes these n expressions can be considered irrelevant and can be removed to create a simpler, combined expression Φ ’ (p.9) [16].

By this logic, many configurations are minimised to a single configuration or set of configurations. The QCA

method can be described as a kind of global revolution in small-*N* methods, with textbooks [17], software [18], and websites developed for case-based analysts (<http://www.u.arizona.edu/~cragin/fsQCA/>). A large body of research offering applications, theory and models of research practice has been developed [17–22].

There are two critical aspects of QCA method that are important to this study: a) qualitatively describing observed cases (state health systems) and b) describing what might be theoretically true when all possible missing cases are added to consideration of observed cases. That is, the logical equations produced by the QCA method can be used to describe not simply observed configurations of cases. They can also be used to describe what, in theory, might be true on the basis of an observed set after all possible instances of configurations of case conditions (called ‘logical remainders’ or cases without empirical instances) have been considered. The Boolean minimisation procedure described by Cronqvist thus involves use of ‘simplifying assumptions’ – an expression of the outcome value of a logical remainder [17]. In a state benchmarking exercise, the use of these logical remainders or empirically absent cases allows theory building about state-level differences. This is important for this kind of study in a context in which a) the number of states may be small (i.e., Australia has only six states and two territories), b) data are restricted to a single year because changes in methods for collecting data over time make traditional longitudinal analyses unfeasible, and c) data are often outdated by the time it is available.

Accordingly, QCA is designed to retain the configurational complexity of cases, that is to tread a ‘middle way’ between complexity and parsimony in summarizing individual case features [17, 19, 21–25]. Where cases are complex (i.e., in the case of state health systems), the advantages of having a method to facilitate systematic description of observed cases are apparent. Where the cases are limited (i.e., countries have only a limited number of states or provincial regions and state-level data can be incomplete) finding ways to better manage this limitation and still produce evidence-informed theory is worth considering.

There are also some specific potential advantages of QCA for policy-making. Policy stories about causality are not simply about what has occurred, i.e., the observed set of cases. They are also about what, in principle, might be true, taking into account all possibilities. Real world policy making often involves developing narratives about what is true based on a limited set of cases, as well as what might also be true if all possible cases were considered [15]. In an apparently complementary fashion, Qualitative Comparative Analysts emphasise this dual purpose of

their techniques for analysing individual cases in small-*N* situations [17].

Methods

This paper is based on a real world health policy consultancy, in which the consultant used MVQCA to demonstrate possible methods for using existing indicators to benchmark an Australian state health system (Tasmania) in the area of chronic disease. ‘Benchmarking’ was defined as comparing the state health system with other systems to decide what might represent good or not-so-good health system performance. The method to be developed, therefore, needed to offer a way to help policy-makers decide what groups of indicators suggested about similarities and differences in state performance. Indicators were examined from the following sources: the Australian Bureau of Statistics, including Australia’s National Health Survey [26, 27], Australian Institute of Health and Welfare research for the period 2000–2011, Australia’s National Healthcare Agreement measures developed with assistance from the Australian Institute for Health and Welfare [28], and Australian federal and Tasmanian state government policy documents and websites.

The variables used from the different databases in the illustrative QCA analyses given in this paper are listed in Table 1. They were selected on the basis that they could provide state-level indicators of community and health system inputs, interventions, and population health outcomes in a specific area of chronic disease (mental health), as well as state profiling for differences in risky health-related behaviours. While the variables come from different sources, their selection was designed to help develop a policy story about state-level differences in two areas, namely, mental health and risky health-related behaviours. They were designed in the consultancy to help illustrate the QCA method, not offer a comprehensive description of how Australian state health systems compare. Furthermore, the Northern Territory was self-excluded from this analysis, because key indicators were not available for it.

The indicators in Table 1 have many shortcomings that limit their interpretability. Some argue that they have not been designed for comparing state health systems that have artificial boundaries, which are considered meaningless in interpreting such variables. Certainly, many diverse factors beyond the scope of their documented development have shaped them. Yet this problem of complex causality variously affects all statistical data that might be used in state benchmarking exercises. In a context in which such data have limited value for health practitioner or service-level decision-making, the need to add value (usefulness) to such data for policy persists and is central to the policy consultancy that informs this paper.

The variables in Table 1 were categorised for each state as relatively ‘high’ or ‘moderate’ or ‘low’ using the simple clustering function provided for such threshold setting based on averages in the MVQCA TOSMANA program developed by Cronqvist (publicly available at <http://www.tosmana.net/>). The median value was used to create dichotomous outcome variables (presently TOSMANA does not accommodate multi-value outcomes, only multi-value conditions). Thus, the variables were coded in the TOSMANA file in the following way: higher levels of community and health system inputs and interventions were given a higher ‘score’ (on a scale of 0,1,2); outcome variables were given a 0 (lower) or 1 (higher) score. In the Boolean

Variable name	Type	Source and description
HSHLD IN	Whole of community input (income)	Australian Bureau of Statistics average real weekly equivalised disposable <i>household income</i> (2008). [29]
EMPLOY	Whole-of-community input (employment)	Australian Bureau of Statistics figures for proportion of civilian population in <i>employment</i> 2009–10. [30]
HOUS STRESS	Whole of community input (housing stress)	Australian Institute for Health and Welfare indicator of affordability for low income households in community housing for 2009–2010 ‘proportion of <i>household income</i> left after rent..’ [31]
HIGHER ED	Whole of community input (education)	Australian Bureau of Statistics indicator of level of higher education for persons aged 25–64 years, including postgraduate, masters, graduate diploma, graduate certificate and bachelor degree. [32]
MH STAFF	Health system input (mental health workforce)	Australian Institute for Health and Welfare data for full-time equivalent mental health staff in specialised mental health care facilities (public hospitals to community and residential care facilities) per 100,000 population, 2007–2008. [33]
MH SERV	Health system input (mental health interventions)	Australian Institute for Health and Welfare data for mental health service interventions using numbers of Medical Benefits Scheme subsidised psychiatrist and allied health services for 2008–2009 (i.e., count of services given as rate per 1000 population, not count of patients). [33]
MH PRES	Health system input (mental health prescriptions)	Australian Institute for Health and Welfare data for mental health related prescriptions obtained from figures for patients dispensed with mental health related subsidised prescriptions for 2008–2009 (the rate per 1000 population). This figure includes GPs, non-psychiatrist specialists and psychiatrists. [33]
SA HEALTH	Population health outcome (self assessed health)	Australian Bureau of Statistics (2008 National Health Survey) indicator of self-assessed health using percentage of people who indicated they had fair/poor self-assessed health. [26]
PSY DIST	Population health outcome (self assessed mental health)	Australian Bureau of Statistics (2008 National Health Survey) indicator of self-assessed mental health based on percentage of people who indicated they had high/very high psychological distress. [26, 27]
SMOKE, ALCOHOL, LOW EXER, INAD FRT, OVERWGT	Population health outcomes (self assessed lifestyle risk factors)	Australian Bureau of Statistics (2008 National Health Survey) indicator of self-assessed mental health based on percentage of people who indicated lifestyle risk factors: (current daily smoker, risky/high risk alcohol consumption/sedentary/low exercise level/inadequate fruit or vegetable consumption; overweight/obesity BMI-measured adults). [26, 27]

Table 1 Sample indicators used in QCA study.

equations obtained in TOSMANA summarising the states (the cases), an asterisk * is logical ‘AND’. A plus sign + is logical ‘OR’.

Results

In the illustrative analyses that follow, the QCA method was used to combine two different sets of indicators to suggest how:

Policy stories about state performance in mental health might be developed, based on systematic comparisons between states involving consideration of differences in (sample) community and health system inputs and interventions; and

Policy stories about state-level profiling for differences in risky health-related behaviours might be developed, using sample whole of community inputs as well as self-assessed lifestyle factors.

State performance in mental health

Mental health is a key performance area for state health systems. The logical equation below (‘Logical equation 1’) was obtained when the outcome of higher levels of self-reported distress in a state population was considered alongside selected socio-economic variables (or whole community inputs) and mental health service inputs. The

logical equation produced summarising observed configurations suggests that Tasmania (TAS) may be similar to South Australia (SA) in that it has lower community inputs – as well as higher levels of mental health system inputs and interventions, such as prescriptions – than other states with this outcome of higher psychological distress.

LOGICAL EQUATION 1 (*higher psych distress*): $\text{HSHLD IN}\{1\} * \text{EMPLOY}\{0\} * \text{HOUS STRESS}\{1\} * \text{HIGHER ED}\{1\} * \text{MH STAFF}\{1\} * \text{MH SERV}\{1\} * \text{MH PRES}\{1\} + \text{HSHLD IN}\{1\} * \text{EMPLOY}\{0\} * \text{HOUS STRESS}\{2\} * \text{HIGHER ED}\{1\} * \text{MH STAFF}\{1\} * \text{MH SERV}\{2\} * \text{MH PRES}\{1\} + \text{HSHLD IN}\{0\} * \text{EMPLOY}\{0\} * \text{HOUS STRESS}\{1\} * \text{HIGHER ED}\{0\} * \text{MH STAFF}\{2\} * \text{MH SERV}\{1\} * \text{MH PRES}\{2\}$ (NSW)(VIC)(SA,TAS)

Another logical equation ('Logical equation 2') was obtained for the outcome *lower* psychological distress, again including community and health system inputs and interventions. Unlike the states with high psychological distress, these states suggest higher levels of community socio-economic well-being than Tasmania. Their better mental health well-being result also does not seem to be accompanied by higher mental health system inputs and interventions (at least on these indicators).

LOGICAL EQUATION 2 (*lower psych distress*): $\text{HSHLD IN}\{1\} * \text{EMPLOY}\{1\} * \text{HOUS STRESS}\{2\} * \text{HIGHER ED}\{0\} * \text{MH STAFF}\{1\} * \text{MH SERV}\{1\} * \text{MH PRES}\{1\} + \text{HSHLD IN}\{1\} * \text{EMPLOY}\{1\} * \text{HOUS STRESS}\{0\} * \text{HIGHER ED}\{0\} * \text{MH STAFF}\{2\} * \text{MH SERV}\{1\} * \text{MH PRES}\{1\} + \text{HSHLD IN}\{1\} * \text{EMPLOY}\{2\} * \text{HOUS STRESS}\{1\} * \text{HIGHER ED}\{2\} * \text{MH STAFF}\{1\} * \text{MH SERV}\{0\} * \text{MH PRES}\{1\}$ (QLD)(WA)(ACT)

What should health policy-makers be targeting at the state system level, as a whole, to achieve the outcome of lower levels of psychological distress? An exploration of the answer to this question in a QCA approach involves considering all possible missing cases or 'remainders'. There are 727 and 726 configurations that require reduction to arrive at the two very simple logical equations given as Equations 3 and 4 below for the outcomes of lower and higher psychological distress, respectively. The logical Equation 3 suggests that at the state system level, moderate to higher employment is both necessary and sufficient to achieve lower psychological distress. The logical Equation 4 suggests that lower employment is a sufficient condition to achieve higher psychological distress. This does not at all mean that health system inputs and interventions are irrelevant to the psychological well-being of individuals or groups in society. It does suggest that, on these data and at the level of understanding state comparisons,

the drivers of state differences in psychological distress appear to be strongly about employment levels, at least under a QCA method.

LOGICAL EQUATION 3 (*Lower psych distress*): $\text{EMPLOY}\{1,2\}$ (QLD+WA+ACT)

LOGICAL EQUATION 4 (*Higher psych distress*): $\text{EMPLOY}\{0\}$ (NSW+VIC+SA,TAS)

In summary then, the QCA suggested that an evidence-based policy story about the outcome of higher psychological distress, reported for Tasmania from the National Health Survey [26, 27] needs to be understood in the context of Tasmania's challenges to do with socio-economic stressors such as employment. The Tasmanian mental health system cannot necessarily be ranked behind other states by virtue of such outcomes, achieved even with higher levels of mental health interventions, because its socio-economic context is more challenging.

State profiling of health-related behaviours

Another kind of state-level benchmarking analysis relates to the role of healthy behaviours in achieving good health, which also needs to be understood in the context of socio-economic factors or whole-of-community factors. Logical Equation 5 summarises the state health system configurations for the selected whole-of-community variables using the outcome variable of higher levels of poor self-assessed health. It suggests that while Tasmania and South Australia are similar in terms of these community inputs, Tasmania reports higher levels of smoking and risky alcohol consumption than South Australia. In fact, like Tasmania, New South Wales (NSW) reports high levels for three out of the five risk factors included. In short, Tasmania leads other states in terms of both socio-economic challenges and risky health behaviours (other states with this outcome struggle variously more with the former than the latter).

LOGICAL EQUATION 5 (*higher poor self-assessed health*): $\text{HSHLD IN}\{1\} * \text{EMPLOY}\{0\} * \text{HOUS STRESS}\{1\} * \text{HIGHER ED}\{1\} * \text{SMOKE}\{1\} * \text{ALCOHOL}\{1\} * \text{LOW EXER}\{2\} * \text{INAD FRT}\{2\} * \text{OVERWGT}\{2\} + \text{HSHLD IN}\{0\} * \text{EMPLOY}\{0\} * \text{HOUS STRESS}\{1\} * \text{HIGHER ED}\{0\} * \text{SMOKE}\{1\} * \text{ALCOHOL}\{0\} * \text{LOW EXER}\{2\} * \text{INAD FRT}\{1\} * \text{OVERWGT}\{2\} + \text{HSHLD IN}\{0\} * \text{EMPLOY}\{0\} * \text{HOUS STRESS}\{1\} * \text{HIGHER ED}\{0\} * \text{SMOKE}\{2\} * \text{ALCOHOL}\{1\} * \text{LOW EXER}\{2\} * \text{INAD FRT}\{1\} * \text{OVERWGT}\{2\}$ (NSW)(SA)(TAS)

Logical Equation 6 includes all possible missing configurations; specifically, it suggests that at the level of state differences what generally matters for this outcome of poor health is higher levels of smoking. That is, under this specific QCA, moderate to high levels of smoking (not any of the socio-economic factors or other lifestyle risk factors) are sufficient conditions at the state level for the outcome of higher self-assessed poor health.

LOGICAL EQUATION 6 (higher poor self-assessed health):
 $\text{SMOKE}\{1,2\}$
 (NSW+SA+TAS)

Logical Equations 7 and 8 provide similar analyses for the outcome of *lower* levels of poor self-assessed health, suggesting in a compatible fashion that under a QCA, lower smoking is the sufficient condition, at the level of state comparisons, for lower levels of poor health.

LOGICAL EQUATION 7 (lower poor self-assessed health)
 $\text{HSHLD IN}\{1\} * \text{EMPLOY}\{0\} * \text{HOUS STRESS}\{2\} * \text{HIGHER ED}\{1\} * \text{SMOKE}\{0\} * \text{ALCOHOL}\{0\} * \text{LOW EXER}\{2\} * \text{INAD FRT}\{2\} * \text{OVERWGT}\{2\} +$
 $\text{HSHLD IN}\{1\} * \text{EMPLOY}\{1\} * \text{HOUS STRESS}\{2\} * \text{HIGHER ED}\{0\} * \text{SMOKE}\{0\} * \text{ALCOHOL}\{0\} * \text{LOW EXER}\{0\} * \text{INAD FRT}\{0\} * \text{OVERWGT}\{0\} +$
 $\text{HSHLD IN}\{1\} * \text{EMPLOY}\{1\} * \text{HOUS STRESS}\{0\} * \text{HIGHER ED}\{0\} * \text{SMOKE}\{0\} * \text{ALCOHOL}\{2\} * \text{LOW EXER}\{2\} * \text{INAD FRT}\{2\} * \text{OVERWGT}\{2\} +$
 $\text{HSHLD IN}\{1\} * \text{EMPLOY}\{2\} * \text{HOUS STRESS}\{1\} * \text{HIGHER ED}\{2\} * \text{SMOKE}\{0\} * \text{ALCOHOL}\{1\} * \text{LOW EXER}\{1\} * \text{INAD FRT}\{2\} * \text{OVERWGT}\{1\}$
 (VIC)(QLD)(WA)(ACT)

LOGICAL EQUATION 8 (lower poor self-assessed health):
 $\text{SMOKE}\{0\}$
 (VIC+QLD+WA+ACT)

Accordingly, the policy story suggested by the QCA is that Tasmania may not be doing as well as some states (SA) with similar socio-economic stressors in preventing risky health behaviours that may be linked to poorer self-assessed health. Yet it may be doing better than other states (NSW) that do not have similar levels of socio-economic disadvantage but share similar risky health behavior levels.

Discussions and conclusions

This paper has suggested that QCA can help combine existing indicators into policy narratives based on systematic comparisons between health systems. However,

in order to develop an evidence-based ‘policy story’, what is needed is a suite of indicators that can build a coherent argument about community and health system inputs, interventions, and outcomes. For too many health areas (e.g., types of cancers), such suites of indicators do not exist, are only partially present, or are fully available only for a single year. Thus a comprehensive QCA benchmarking exercise must involve careful collection of the indicators and participative styles of stakeholder involvements. Table 2 provides the steps indicated by the pilot consultancy as needed to progress a more comprehensive QCA-based approach.

QCA appears to have some potential in a context wherein policy-makers lack a method for systematically comparing indicators for different state health systems to develop ‘policy stories’. The QCA method allows description of possibly important contextual state differences in population health inputs that need to be considered in benchmarking outcomes in the different states. The current study suggested theoretically plausible explanations for differences in state-level outcomes that relate to socio-economic disadvantages (unemployment) and risky health behaviours (smoking).

The QCA analyses also suggested multiple explanatory paths for the same outcome (different ways in which the different states may have achieved the same outcomes). Policy-makers are engaged in finding the answers to complex questions such as ‘What combination of which factors in this health system may explain these state outcomes?’ Accordingly, before setting realistic benchmarks, they may need to consider such diverse explanatory paths suggested by different state health systems. Thus, QCA can help add value to existing population health indicators by helping policy-makers develop complexity-oriented policy stories as a foundation for policy options. Such complexity-oriented policy stories can help develop more complexity-oriented policy cultures, which is an ostensible goal of democratic governments such as the UK’s Blair government in the modern age [34].

A key aspect of QCA relates to the setting of thresholds or levels into which the data for an indicator can be categorised. The TOSMANA software used in this study offers a data visualisation facility that can be shared with groups of epidemiologists, QCA researchers, policy-makers, health service administrators, practitioners, and community stakeholders to achieve participative approaches to setting realistic benchmarks for states. The steps for conducting a QCA study set out in Table 2 provide a strong emphasis on participative approaches to both the deciding of what indicators to use and the setting of normative thresholds to categorise data within such indicators (e.g.,

Step one: identify indicators	Step two: consolidate and tabulate indicators	Step three: QCA of indicators	Step four: refine and test the findings	Step five: produce set theoretic findings
<p>Assemble a reference group of epidemiologists, QCA experts, policy-makers, health service administrators, practitioners, and community stakeholders</p> <p>Map and audit available population health indicators: Make preliminary decisions about which indicators may be of relevance, taking into account community and health system inputs, interventions, and population health outcomes</p> <p>Use a wide range of research data, both about this state systems, as well as population health generally, to inform the preliminary choice of indicators</p>	<p>Develop an 'item bank' of relevant indicators for your health system, using input from the stakeholder reference group and diverse experts, and including information from national health benchmarking agreements; this item bank will include the 'provenance' for each population health indicator (how it was developed, limitations, relevant to its meaning)</p> <p>Enter data for all indicators in a software program such as TOSMANA http://www.tosmana.net/ or fsQCA http://www.u.arizona.edu/~cragin/fsQCA/software.shtml</p> <p>Participatively decide on the thresholds (for MVQCA levels such as 0,1,2) for each indicator, taking quantitative data (e.g., confidence intervals) as well as qualitative information (e.g., national policy expectations) into account</p>	<p>Perform the QCA analyses in the QCA software, examining both observed cases, as well as all possible missing configurations (remainders)</p> <p>Perform suites of analyses, including different kinds of community and health system inputs, interventions, and population health outcomes</p> <p>Document the 'logical equations', and other QCA outputs so produced</p>	<p>Participatively discuss the necessary and sufficient (or neither) conditions that may be at work in state comparisons</p> <p>Develop possible explanations for the logical equations so produced</p> <p>Consider multiple explanatory paths (i.e., causal complexity)</p> <p>Triangulate the results of the analysis with reference to other research findings produced by other methods</p>	<p>Produce statements that provide possible causal explanations for state differences</p> <p>Share the findings with the wider community in forums that allow reflection on appropriate recommendations for state policy</p> <p>At regular (yearly or biennial) intervals, review and revise the content of the item bank, and repeat the QCA</p>

Table 2 Steps in developing policy stories using a QCA-based approach to state benchmarking.

as 'low', 'moderately low', or 'high'). For efforts to make population health research more policy relevant, this capacity to accommodate normative judgments could be viewed as a major strength of the MVQCA method.

However, QCA also presents barriers to stakeholder participation. An effort of understanding is required to master QCA concepts and language – not simply the software. Terms such as 'logical equations' and 'remainders' [17]_ENREF_23 are not as familiar to many as 'significance'. Care is needed to ensure that the language of QCA is not used to unequivocally suggest a mythical empirical causality.

More complex questions about QCA lie in the theoretical and analytical rigor of, for example, how QCA integrates missing cases. What precisely is the theoretical and empirical status of a finding obtained by considering, on a non-probabilistic basis, the entire universe of missing configurations, given that the real world may never contain all those cases (at least at any moment in time)? Procedures developed in QCA for sorting through and including or excluding missing cases do not entirely answer such questions at this stage [17].

In the Australian system, as in other systems, there is a need to (1) better identify what population indicators should be used in state benchmarking exercises; (2) examine how to better develop them into robust measures of state-level health system performance; and (3) use these indicators to better tell an evidence-based policy story about state performance. This study has suggested how the QCA method could help achieve the third and often neglected task. It is acknowledged that state benchmarking is a highly politically sensitive issue that has not, and perhaps never will be, entirely driven by evidence. Yet while causal stories in policy-making will often be highly political as Stone and other policy analysts have suggested [13, 35–37], they exist in a spectrum that has at one end of it strongly evidence-informed policy and at the other policy that has little basis in evidence. The challenge addressed by this study was to offer policy-makers support and practical methods for operating at the evidence-based end of the spectrum.

Acknowledgements: The consultancy that informed this paper was funded by the Department of Health and Human Services Tasmania. All statements made in this paper were independently developed by the consultant and do not represent the views of DHHS or any of its employees. The production of this journal paper was supported by a core operating grant for the University Department of Rural Health at the University of Tasmania, received from the Department of Health and Ageing.

Conflict of interest statement

Authors' conflict of interest disclosure: The authors stated that there are no conflicts of interest regarding the publication of this article.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Received October 12, 2012; accepted October 17, 2012

References

- Duszak Jr R, Muroff LR. Measuring and managing radiologist productivity, Part 1: clinical metrics and benchmarks. *J Am Coll Radiol* 2010;7:452–8.
- Economou NA, Tountas Y. Efficiency evaluation in health care. *Arch Hellenic Med* 2007;24:34–47.
- Hussey P, Anderson G, Berthelot JM, Feek C, Kelley E, Osborn R, et al. Trends in socioeconomic disparities in health care quality in four countries. *Int J Qual Health C* 2008;20:53–61.
- Ring IT, O'Brien JF. Our hearts and minds-What would it take for Australia to become the healthiest country in the world? *Med J Aust* 2007;187:447–51.
- Lai MC, Huang HC, Wang WK. Designing a knowledge-based system for benchmarking: A DEA approach. *Knowl-Based Syst* 2011;24:662–71.
- Thornton C, Hennessy A, von Dadelszen P, Nishi C, Makris A, Ogle R, et al. An international benchmarking collaboration: measuring outcomes for the hypertensive disorders of pregnancy. *J Obstet Gynaecol Canada* 2007;29:794–800.
- Reintjes R, Thelen M, Reiche R, Csohán Á. Benchmarking national surveillance systems: a new tool for the comparison of communicable disease surveillance and control in Europe. *Eur J Public Health* 2007;17:375–80.
- Davis K, Stremikis K. Measuring and improving health system performance: what can healthcare management educators do? *J Health Admin Educ* 2008;25:5–15.
- Shea K, Davis K, Schor E, “U.S. Variations in Child Health System Performance: A State Scorecard”. USA: The Commonwealth Fund, 2008.
- Silow-Carroll S, Moody G. Lessons from high- and low-performing states for raising overall health system performance. Issue Brief (Commonwealth Fund) 2011;7:1–11.
- Lemmers O, Broeders M, Verbeek A, Heeten GD, Holland R, Borm GF, et al. League tables of breast cancer screening units: Worst-case and best-case scenario ratings helped in exposing real differences between performance ratings. *J Med Screen* 2009;16:67–72.
- Tawfik-Shukor AR, Klazinga NS, Arah OA. Comparing health system performance assessment and management approaches in the Netherlands and Ontario, Canada. *BMC Health Services Res* 2007;7:25.
- Stone D. Causal stories and the formation of policy agendas. *Polit Sci Quart* 1989;104:281–300.
- Majone G. Evidence, argument & persuasion in the policy process. New Haven, CT: Yale University Press, 1989.
- Bell E. Research for health policy. Oxford: Oxford University Press, 2010.
- Cronqvist L. Using Multi-Value Logic Synthesis in Social Science. 2nd General Conference of the European Consortium for Political Research (ECPR); Marburg: University of Essex, European Consortium for Political Research, 2003.
- Rihoux B, Ragin C, editors. Configurational comparative methods: qualitative comparative analysis (QCA) and related techniques. Los Angeles, CA: Sage, 2009.
- Ragin C, Drass K, Davey S. “Fuzzy set/Qualitative comparative analysis 2.0”. Tucson, AZ: Department of Sociology, University of Arizona, 2007.
- Ragin C. Redesigning social enquiry: fuzzy sets and beyond. Chicago, IL: University of Chicago Press, 2008.
- Ragin C. Set relations in social research: evaluating their consistency and coverage. *Polit Analysis* 2006;14:291–310.
- Ragin C. The limitations of net effects in thinking. In: Grimm H, Rihoux B, editors. Innovative comparative methods for policy analysis. New York: Springer, 2006:13–41.
- Ragin C. Fuzzy-Set social science. Chicago, IL: The University of Chicago Press, 2000.
- Goodman SN. Of P-values and bayes: a modest proposal. *Epidemiology* 2001;12:295–7.
- Poole C. Low P-values or narrow confidence intervals: which are more durable? *Epidemiology* 2001;12:291–4.
- Editors. The Value of P. *Epidemiology* 2001;12:286.
- ABS. 4362.0 National Health Survey: Summary of Results; State Tables, 2007-2008 (Reissue) [database on the Internet]. Australian Bureau of Statistics. 2008 [cited 23 May 2011]. Available from: <http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4362.02007-2008%20%28Reissue%29?OpenDocument>.
- ABS. “National health survey: summary of results, Australia, 2007–08 (reissue)”. Canberra: Australian Institute for Health and Welfare, 2009.
- COAG. “National healthcare agreement”. Canberra: Commonwealth of Australia, 2009.
- ABS. “ABS, Household income and income distribution Australia 2007–08”. Canberra: Australian Bureau of Statistics, 2010.
- ABS. Cat. no. 4102.0, Australian social trends, data cube, work [database on the Internet]. Australian Bureau of Statistics. 2010 [cited 24 May 2011]. Available from: <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Main+Features12008?OpenDocument>.
- AIHW. “Community housing 2009–10”. Canberra: Australian Institute of Health and Welfare, 2011.
- ABS. 1370.0 Measures of Australia's progress, 2010 (2002-2009 Survey of education and work). Canberra: Australian Bureau of

- Statistics; 2011 [cited 2011 23 May]; Available from: <http://www.abs.gov.au/ausstats/abs@.nsf/mf/1370.0>.
33. AIHW. "Mental health services in Australia 2007–2008". Canberra, 2010 August.
34. Government Chief Social Researcher's Office. Magenta Book: Guidance Notes on Policy Evaluation. London: Cabinet Office; 1997 [cited 2008 21 May]; Available from: http://www.nationalschool.gov.uk/policyhub/magenta_book/index.asp.
35. Birnbaum R. Policy scholars are from venus; policy makers are from mars. *Rev Higher Educ* 2000;23:119–32.
36. Black N. Evidence-based policy: proceed with care. *Br Med J* 2001;323:275–9.
37. Nutley S, editor. Bridging the policy/research divide: reflections and lessons from the UK. *Facing the Future: Engaging Stakeholders and Citizens in Developing Public Policy*, 2003, Canberra, National Institute of Governance Conference.