Implementing Patient-Reported Outcome Measures into Clinical Practice Across NSW: Mixed Methods Evaluation of the First Year

Claudia Rutherford[1,2] · Rachel Campbell[1] · Melissa Tinsley[3] ·
Robyn Speerin[3] · Linda Soars[4] · Anna Butcher[5] · Madeleine King[1]

Abstract

Use of patient-reported measures (PRMs) in healthcare is increasing internationally. In New South Wales (NSW) Australia the implementation of a system-wide PRMs program is underway. This program is an enabler for value based healthcare. We report an evaluation (2015/16) of this program's first year. We evaluated the program using mixed methods. People living with chronic conditions from 18 sites across NSW completed PROMIS10 at all sites; DASS21 and CAT administered at selected sites depending on patient cohorts. PRM completion rates and mean scores were calculated. Stakeholder interviews were analysed using thematic analysis. PRM completion rates were high at baseline (69%) but suboptimal at follow-up time-points. Mean scores indicated people with back pain had worse physical health and stress compared to those with other conditions. People with alcohol or drug problems had worse mental health, disease symptoms, depression and anxiety compared to other conditions. Stakeholders reported collection of PRMs could improve health outcomes for patients, enhance service delivery, and reduce avoidable hospital admissions. Routine collection and use of PRMs can be implemented across NSW care settings. Several challenges need to be overcome to realise the full benefits of the PRM program and to improve follow-up completion rates.

Keywords Patient-reported outcomes · Health-related quality of life · Implementation · Co-design · Value based healthcare

✉ Claudia Rutherford
claudia.rutherford@sydney.edu.au

Extended author information available on the last page of the article

🅶 Springer

Introduction

Since the 1990s there has been exponential growth in the development and application of patient-reported measures (PRMs) of health. PRMs can be grouped into Patient Reported Outcome Measures (PROMs) and Patient Reported Experience Measures (PREMs). The former capture the person's perspectives about how illness or healthcare impacts on health and general wellbeing, whereas the latter capture a person's perception of their experience with healthcare systems or services. Such standardized and validated tools have widespread use in clinical trials to assess a person's outcomes such as quality of life or symptoms related to a specific condition. There is now increasing enthusiasm and momentum to use PRMs as a part of routine patient assessment and to evaluate health systems, as part of the quality metric in value-based care(OECD 2010). PRMs can be used at the micro level in the care of individual patients, at the meso level for health service improvement, and at the macro level by health system policy makers. When used as part of routine clinical practice, conclusive evidence supports the use of PRMs for care planning and decision-making to provide timely person-centered care. Their use improves communication between individuals and care providers, and ensures appropriate referrals are based on collaboratively identified patient needs(Chen et al. 2013; Ishaque et al. 2019).

In clinical practice, measures of health status and quality of life are used for needs assessment and for monitoring the effects of chronic illness, treatments, and disabilities. The effects can impact on individual functioning and psychological well-being. Together, anxiety and depression disorders are experienced by more than 20% of Australian adults(Australian Institute of Health and Welfare 2018). One in 10 people report depression or feelings of depression, and one in 13 report anxiety-related conditions (Australian Bureau of Statistics 2019). Being diagnosed with depression and/or anxiety is a stronger predictor of chronic physical health conditions than demographic factors such as age and marital status, and physical activity(Stanton et al. 2019). In Australia in 2017–18, Chronic obstructive pulmonary disease (COPD) was experienced by 598,800 people (2.5%)(Australian Bureau of Statistics 2019). Depression is a major comorbidity among patients with COPD, with an estimate prevalence of up to 80% in severe stages of COPD. Prevalence studies show that patients with COPD are four times as likely to develop depression compared to those without COPD. It is also a major determinant of COPD-related quality of life(Jang et al. 2019). Results from a systematic review show that the rates are equal to, or higher than the rates of depression amongst patients with cancer, AIDS, or heart disease(Solano et al. 2006).

Internationally, PRM implementation initiatives are gaining momentum, including in Australia. However, the routine or systematic use of PRMs within clinical practice may be hampered by a number of practical, methodological and attitudinal barriers (Greenhalgh and Meadows 1999). Implementation issues are healthcare context specific, especially for a state health system providing services across a wide diversity of health conditions, populations, and care settings (metropolitan, rural and remote). A complication in Australia is the mix of state versus commonwealth funding for health service provision, with governance issues that hamper implementation.

In response to the challenges posed by changing health needs and the increasing demands on the New South Wales (NSW) health system the NSW Ministry of Health announced in 2014 the NSW Integrated Care Strategy. This strategy aimed to increase focus on the delivery of coordinated care across diagnostic groups and care settings (Health and Social Policy 2018). The aim was to transform how healthcare was delivered in NSW to improve health outcomes and reduce costs derived from inappropriate and fragmented care, for people across hospital and primary care services. The implementation of PRMs was recognized as an enabler of integrated care and considered a prioritized investment. The Agency for Clinical Innovation (ACI)[1] is responsible for leading and managing the PRM component of the strategy. Through implementation of the PRMs program, ACI aims to 'Enable patients to provide direct, timely feedback about their health related outcomes and experiences of care to their care providers to drive improvement and integration of health care across NSW'(Agency for Clinical Innovation 2019). In doing this, the PRMs program seeks to reduce the burden on clinicians and patients and add value to their interactions.

Guidelines have been developed by the International Society of Quality of Life (ISOQOL), offering a stepwise approach to implementing PRMs into clinical practice(Snyder et al. 2012; Aaronson et al. 2015). Successful implementation requires engagement with key stakeholders, co-design and the identification and resolution of barriers to PRM implementation at the local site level. A number of studies report the setting into which the PRM assessment is to be integrated, has a significant effect on PRM implementation(Porter et al. 2016).

The aim of the larger NSW Health state-wide PRM program is to introduce the systematic collection of PRMs as a driver for clinical improvement, capturing both PROMs and PREMs across a range of clinical settings across NSW including general practices, outpatient specialty clinics, community services and hospitals. The ACI led the co-design, implementation, testing and refining of the PRM proof of concept program – this work has largely contributed to the now broader NSW Health PRM program. The first step was to test the effectiveness of PRMs in 11 geographically located proof of concept locations. This study reports on the feasibility findings from the first year of the three-year evaluation of 18 sites within the 11 geographical locations that collected PROMs (no PREMs). The aim of the one-year pilot (reviewing 2015–2016) was to determine whether PROMs could be routinely collected in NSW and what the completion rates were when such a program was implemented. The NSW Health PRMs program included the collection and implementation of PREMs, which will be reported in a separate paper.

Methods

The one-year pilot evaluation utilized mixed methods including quantitative analysis of PROM data from 18 sites within 11 geographical locations and qualitative interview data from consultations with program stakeholders, clinicians and managers; and a focus group with three patients.

---

[1] Names of authors and funders replaced with XXX for blind manuscript

PRM Program Design

The design of the PRMs program was iterative, utilizing a co-design approach between the ACI and managers, clinicians and people who may access health care services in the proof of concept sites. Co-design is a way of improving healthcare services for people accessing and using care services by bringing together all stakeholders and consumers in partnership, to develop health services that best meet the needs of consumers and carers in the most effective way possible(Greenhalgh et al. 2016).

Initially, sites within four early adopter local health districts (LHDs) were self-selected to develop and implement PROMs in their local service settings. This included sites from Western NSW, Western Sydney, North Sydney and Mid North Coast(Cheung et al. 2019). Key stakeholders from across the identified sites participated in a series of workshops to co-design the program and program logic. Importantly, the design of the program was not a linear process; there were many elements that were co-designed and tested. Aspects of the program that did not meet the needs of the broader program (e.g. sites who were interested in purely collecting PRM data but not utilizing it in real-time) were reconsidered and further developed or excluded for the purposes of the proof of concept program. The scope of geographical locations and sites rapidly expanded soon after the program launched to include 18 sites (general practices, outpatient specialty clinics, community services and hospitals) within 11 geographical locations across LHDs and Specialty Health Networks.

Participants and Procedure

PROMs were available for administration across 18 sites located in NSW Australia, during 2015–2016. Participating sites were selected based on an expression of interest distributed to Local Health Districts and Specialty Health Networks across NSW. Those that were delivering Integrated Care programs were eligible for self-election to participate in the proof of concept program. Decisions about which sites to pilot was based on having a diversity of sites from metropolitan and rural areas and both hospital and general practices and not by prevalence of any one chronic condition.

Patient cohorts were identified in collaboration with participating sites and ACI. Patients over the age of 18, with any chronic or complex health condition, attending a health service taking part in the pilot program were included in the PRMs program. We define a chronic and complex condition as a medical condition that tends to be long-lasting and persistent in their symptoms or development and that requires some form of coordinated care or ongoing management(NSW Ministry of Health 2018). As PROMs were available in English only during the first year, ability to self-complete in English was an eligibility criterion. Following the Integrated Care principles, PROMs supported collection of information through surveys in a way that reflected the whole of a person's health needs; from prevention through to end of life, across both physical and mental health, and in partnership with the individual, their carers and family(NSW Ministry of Health 2018). Individuals at-risk of an emerging complex health condition were also eligible.

Administration of the PROMs was via a stand-alone web-based application (REDCap) implemented at each site designed to collect the PROM data. All sites were provided with hand-held tablet devices to support the routine administration of the

PROMs. PROMs were provided to patients to self-complete in the waiting room or while patients were waiting to see their care provider. Each PROM took approximately 3–5 min to complete. Assistance from a carer, family member or care provider was permitted when completing PROMs; however the responses had to be the patients' own without interpretation. All of the participating sites received training and education in the stand-alone IT system and were provided with user guides and quick reference sheets about how to administer the PROMs. Scoring was completed automatically in the stand-alone IT system. As part of site education and training, it was specified that assistance with PROM completion needed to be without interpretation or bias. Patient demographic and clinical characteristics (e.g. chronic illness type) were collected by site staff upon registration to the study and entered into the study database.

The PROMs were available for patient self-completion at up to three time-points: patients' first (i.e. T0 Baseline), second (i.e. T1) and third visit (i.e. T2) to the site. The exact time interval between the time-points differed per site but was generally 3–6 months between each PROM administration. In rare instances where the electronic format was not available, a small proportion of patients completed PROMs via paper. Data from paper copies was entered into the web-based system by a care provider.

Patient-Reported Outcome Measures

Three PROMs were available during 2015–2016 (noting expansion of PRM collection since this time period) for administration at the participating sites: the Patient-Reported Outcomes Measurement Information System (PROMIS10)(Hays et al. 2009), The Depression Anxiety Stress Scales (DASS21)(Lovibond and Lovibond 1995), and the Chronic obstructive pulmonary disease Assessment Test (CAT)(Jones et al. 2009). PROMIS10 was available at all participating sites, however the DAS21 and COPD-CAT were only available at selected sites that had services for people with mental health or COPD problems.

The PROMIS10 was the chosen generic quality of life measure (based on a rapid scoping review(Chen et al. 2013) and consultation with all key stakeholders) to be completed by all patients; this was important as the patient cohorts had multi-morbid conditions and the PROMIS10 provided a broad overview of the various aspects of a person's life that could be affected by a chronic condition. The DASS21 was added to the PRMs program after consultation with clinician stakeholders involved in the PRM program, who identified mental health as an important outcome that should be collected. Stakeholders often used the DASS21 to assess key indicators of mental health (i.e. symptoms of depression, anxiety and stress) and it was widely acceptable to clinicians within the NSW health care system. In addition, stakeholders identified that a large number of programs, clinics and GP services chose to concentrate their clinics on people living with COPD and therefore we added the COPD CAT measure to the PRM program suite of PROMs. The DASS21 was typically completed, if a person scored high on the emotional well-being item on the PROMIS10; or where a clinician thought it would be of benefit. The COPD-CAT measure was completed at practices or clinics where they were targeting people living with chronic obstructive pulmonary disease (COPD) and therefore measuring their symptom burden was important. All PROMs were scored according to their respective scoring manuals.

PROMIS10

The PROMIS10 is a generic instrument that assesses general perceptions of health during the past week. It consists of 10 items, each evaluating a different aspect of health. Four items (physical health, physical activities, pain, fatigue) generate a global physical health component score. Another four (mental health, emotional problems, satisfaction with social activities and relationships, and quality of life) generate a global mental health component score. The remaining two items assess general health status and role limitations. For the purpose of the present descriptive investigation, only the eight items which scored into the global physical- and mental health component scores were included in the analysis. The physical/mental health component scores range from 0 to 20, with higher scores indicating better functioning. The PROMIS10 was chosen for administration across all sites because it is suitable for use with the general population and with individuals living with chronic conditions(Hays et al. 2009).

DASS21

The DASS21 is the short-form version of the Depression Anxiety and Stress Scale(Lovibond and Lovibond 1995). The instrument consists of 21 items which assess the prevalence of symptoms of depression (7 items), anxiety (7 items), and stress (7 items) during the past week. The seven items corresponding to each subscale are summed to produce a separate score for depression, anxiety and stress that ranges from 0 to 21, with higher scores representing worse functioning.

Cat

The CAT assesses the impact of COPD on the health status of patients with this condition (Jones et al. 2009). It consists of eight items which are rated with respect to the respondent's current condition. The eight items are summed to produce a score from 0 to 40, with higher scores representing worse health status.

Qualitative Interviews

In-depth telephone interviews were conducted with PRM program stakeholders, including NSW Ministry of Health and Cancer Institute NSW representatives, and health service staff and managers from participating sites between 18 April and 16 May 2017. Interviews lasted between 15 and 45 min. Participants were identified by the the ACI and other stakeholders and selected for their ability to comment on the PRM program, with a focus on overall service implementation and potential for patient, system or service level outcomes. One focus group with three participants was held on 30 May 2017 at the the ACI office; one participant attended in person and two attended via teleconference. The focus group lasted approximately 30 min. Participating sites were asked to advertise participation in a qualitative interview to patients who had completed PROMs or were an active person within their care service (health service staff). Consumer participants were selected by the the ACI based on their ability to comment meaningfully on the program experience and outcomes. Interviews were recorded and transcribed.

Statistical Analyses

Quantitative Data Analysis

PROMs were scored according to their respective scoring manuals. In accordance with the PROMIS scoring guidelines, if data was missing for one or more items, the participants' score on that scale (i.e. physical or mental health summary score) was recorded as missing. Given that no guidance was available on missing data in the DASS or CAT scoring manuals, if no more than 50% of items were missing, the missing values were estimated by multiple imputation.

Descriptive Analyses

Were performed on PROM data to examine the demographic characteristics of the sample at baseline; to investigate the frequency of PRM completion at each time-point; and to calculate the mean score on each PRM as a function of the medical conditions most commonly reported at baseline. All analyses were performed using SPSS® Statistics for Windows® version 22.0 (IBM, Armonk, New York, USA) and were conducted using two-tailed tests at five level of significance.

Qualitative Data Analysis

Thematic analysis(Dixon-Woods et al. 2005; Thomas and Harden 2008) was used to synthesize findings from interviews. Transcripts were read iteratively to identify common themes across interviews and to develop a structure of perspectives from different stakeholder groups (i.e. clinicians, site managers and staff, and consumers). This approach allowed for an analysis of themes among participants rather than exact proportions of participants who held a particular perspective.

Results

PROM Data Results - Sample Characteristics at Baseline (I.E. T0)

At baseline, 2321 patients were registered for the PROMs program but when asked if they could complete the questionnaires in English, only 1884 participants (42.8% male) responded yes and were registered for PROM completion across 18 sites. Total mean age was 72.7 years (range 18 to 103 years). The majority of participants were not indigenous with 5.4% Aboriginal and/or Torres Strait Islander. The most commonly reported medical conditions ($N > 50$) were heart disease (21.2%), skin cancer (14.1%), diabetes (11.5%), COPD (10.1%), back pain (4.4%), osteoarthritis (4.2%), and alcohol and other drugs (3.2%). Fifteen percent of participants reported one or more comorbid conditions. Of note, a considerable number of participants ($N = 57.8\%$) failed to record their medical condition. Detailed sample characteristics at baseline are presented in Table 1.

Table 1 Characteristics of the total sample at baseline (i.e. T0), $n = 1884$

Variable	Total N = 1884 (100%)
Age	
Mean age (range)	72.72 (18–103)
Missing	38 (2.0%)
Gender	
Male	807 (42.8%)
Female	893 (47.4%)
Prefer not to say	1 (0.1%)
Missing	183 (9.7%)
Indigenous status	
Only Aboriginal	91 (4.8%)
Only Torres Strait Islander	5 (0.3%)
Aboriginal & Torres Strait Islander	5 (0.3%)
None	1660 (88.1%)
Prefer not to say	114 (6.10%)
Missing	9 (0.50%)
Medical condition*	
COPD	191 (10.1%)
Heart disease	229(12.2%)
Osteoarthritis	80 (4.2%)
Osteoporosis	37 (2.0%)
Diabetes	217 (11.5%)
Back pain	82 (4.4%)
Alcohol & other drugs	60 (3.2%)
High blood pressure	11 (0.6%)
Asthma	8 (0.4%)
Renal disease	20 (1.1%)
Skin cancer	265 (14.1%)
Mental health problems	22 (1.2%)
Other	73 (3.9%)
Missing	1089 (57.8%)
Comorbidity	
Missing	1089 (57.8%)
0	507 (42.2%)
1	149 (7.9%)
2	90 (4.8%)
3	31 (1.6%)
4	12 (0.6%)
5	1 (0.1%)

*Each of the medical conditions was analysed as a separate variable. However, the participants endorsed all of the conditions that applied to them, so the categories are not mutually exclusive.

PROM Completion Rates

PROM completion rates at baseline (i.e. T0), assessment 2 (i.e. T1) and assessment 3 (i.e. T2) are presented in Table 2. Of the 1884 participants who registered at baseline, approx. 1300 (69%) completed at least one of the PROMIS10 items and 1049 (55.68%) completed all PROMIS10 items (i.e. had computable scale scores; See Fig. 1 for a flow chart displaying the completion rates for the PROMIS10 at baseline). At baseline, 130 (6.90%) participants had computable scale scores for the three DASS21 subscales (i.e. depression, anxiety, stress) and 181 (9.61%) participants had computable scale scores for the CAT. Of note, not all participants who completed PROMIS10 were approached to complete the DASS21 or CAT measures.

At visit 2 (i.e. T1), only 168 participants registered for PROM completion, of which, 146 (86.90%), 13 (7.74%), and 14 (8.33%) participants had computable scale scores for the PROMIS10, DASS21 and CAT, respectively. Lastly, at visit 3 (i.e. T2), only 17 English-speaking participants registered for PRM completion, of which, 17 had computable scale scores for the PROMIS10, five (29.41%) for the CAT, and none for any of the DASS21 subscales.

Table 2 Frequency of PROM completion at each time point

PRM and scale	Baseline (T0)	Assessment 2 (T1)	Assessment 3 (T2)
Total English-speaking participants registered at each time-point	1884	168	17
N completed at least one item on PROM			
PROMIS10 Physical health	1300 (69.00%)	146 (86.90%)	17 (100%)
PROMIS10 Mental health	1302 (69.11%)	146 (86.90%)	17 (100%)
N with computable scale scores			
PROMIS10 Physical health[a]	1049 (55.68%)	146 (86.90%)	17 (100%)
PROMIS10 Mental health[b]	1049 (55.68%)	146 (86.90%)	16 (94.12%)
DASS21 Depression[c]	130 (6.90%)	13 (7.74%)	0
DASS21 Anxiety[d]	130 (6.90%)	13 (7.74%)	0
DASS21 Stress[e]	130 (6.90%)	13 (7.74%)	0
CAT[f]	181 (9.61%)	14 (8.33%)	5 (29.41%)

T0 = Baseline assessment, T1 = 2nd assessment, T2 = 3rd assessment

[a] 19.31% of respondents had ≥1 missing item on the PROMIS10 physical health scale. These missing values were set to missing. Ns with computable scales are reflective of this

[b] 19.43% of respondents had ≥1 missing item on the PROMS10 Mental health scale. These missing values were set to missing. Ns with computable scales are reflective of this

[c] 1.53% of respondents had 1 item missing on the DASS21 Depression scale. Missing values were estimated by multiple imputation

[d] 0.77% of respondents had 2 missing items on the DASS21 Anxiety scale. Missing values were estimated by multiple imputation

[e] 0.77% of respondents had two missing items on the DASS Stress scale. Missing values were estimated by multiple imputation

[f] 0.55% of respondents had 1 missing item on the CAT. Missing values were estimated by multiple imputation
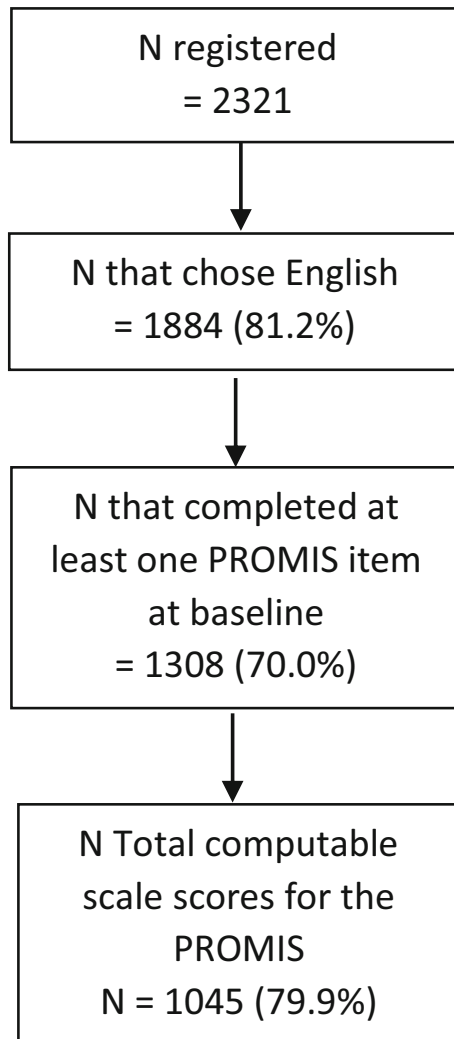
Fig. 1 Flow chart of the completion rates of the PROMIS10 at baseline

Observing completion rates by service type found no patterns in non-completion. Of 18 sites having PROMIS10 questionnaires available for completion, only one metropolitan and one rural community service had >20% of eligible patients not completing baseline PROMIS10 questionnaires. However, all apart from one metropolitan hospital, where available for completion, had >20% of eligible patients not completing baseline DAS21 or COPD-CAT questionnaires (see online supplement).

PROM Mean Scores at Baseline as a Function of Medical Condition

Table 3 presents mean scores for each PROM at baseline for the most commonly reported medical conditions. PROM mean scores, on average, indicate that individuals

Table 3 PRM means by common medical conditions at baseline

Medical condition (sample size per PROM: PROMIS PH, PROMIS MH, CAT, DASS)	PROMIS PH M (SD)	PROMIS MH M (SD)	CAT M (SD)	DASS depression M (SD)	DASS anxiety M (SD)	DASS stress M (SD)
COPD (n = 180, 53, 155, 53)	11.70 (3.25)	12.03 (3.55)	18.50 (8.45)	7.92 (6.06)	7.68 (5.33)	7.94 (5.82)
Heart disease (n = 220, 221, 67, 74)	11.75 (2.98)	12.08 (3.36)	17.84 (8.62)	6.67 (5.66)	6.90 (4.99)	7.35 (5.71)
Osteoarthritis (n = 78, 79, 39, 15)	11.87 (3.08)	12.86 (3.28)	16.05 (8.10)	4.67 (6.09)	6.47 (6.14)	6.00 (6.48)
Diabetes (n = 215, 213, 44, 52)	12.13 (3.40)	12.19 (3.78)	16.39 (7.91)	6.73 (6.90)	6.12 (5.79)	7.13 (6.48)
Back pain (n = 79, 80, 30, 16)	10.49 (2.50)	11.26 (3.43)	17.63 (9.90)	9.25 (5.79)	8.31 (4.59)	10.87 (5.22)
Alcohol & other drugs (n = 59, 59, 11, 7)	11.80 (3.28)	10.78 (3.32)	25.18 (5.34)	11.29 (5.59)	8.86 (5.11)	9.71 (5.56)

PH Physical health; MH Mental Health; M mean; SD standard deviation

Higher scores on the PROMIS10 scales represent better functioning. Higher scores on the CAT and the DASS subscales represent worse functioning

with back pain had lower (worse) scores on PROMIS physical health and higher (worse) scores on DASS stress relative to other medical conditions. Individuals with alcohol and other drug problems had lower (worse) scores on PROMIS mental health and higher (worse) scores on CAT, DASS depression and DASS anxiety. Individuals with diabetes had higher (better) scores on PROMIS physical health and lower (better) scores on DASS anxiety relative to other medical conditions, whereas individuals with osteoarthritis had higher (better) scores on PROMIS mental health and lower (better) scores on CAT, DASS depression and DASS stress.

Qualitative Interview Results

Interviews were conducted with staff at the the ACI ($n = 4$), Ministry of Health (n = 4), Cancer Institute NSW ($n = 1$), site staff and managers ($n = 17$) and consumers ($n = 3$); total 29. The findings are reported according to five broad themes.

Barriers to Successful Implementation of PROMs

A range of challenges to both PROM planning and implementation were identified:

- Impact on clinical workflow due to high level of support required to help patients complete PROMs; limited staff capacity to provide required level of support for implementation; and data entry of paper-based PROMs.
- PROM completion burden on patients.
- Lack of capacity for timely patient referral or follow-up.
- Lack of integrated electronic data management systems.
- Lack of information sharing within and between services, particularly for hospital discharge planning (due to information systems barriers and relationship barriers), specifically for shared care planning.
- Difficulties building relationships with key stakeholders to identify, assess, treat and refer patients (e.g. between acute and primary care settings).
- Need for culture change and shared understanding and ownership by all health professionals.
- Difficulties with linkages to the wider Integrated Care system (e.g. monitoring and evaluation of the program).

Facilitators (Enablers) to Successful PROM Implementation

A range of enablers for successful planning and implementation of PROMs were identified:

- Information access and sharing mechanisms to inform patient care.
- Co-location of medical and allied health services.
- Established care facilitator roles.
- Workforce training and education.
- Patient and Health professional ownership and engagement.
- Efficient communication and pathways.

Potential Benefits of PROMs to Patients, Site Staff, and Health Services

Given this evaluation was at an early stage of implementation, comments made about the program impacts are considered preliminary. However, anecdotally, a number of positive impacts on patients, site staff and health services were reported:

- The program (unintendedly) increased discussion about the patients' experience, and was considered an important first step in building broad acceptance of the collection and use of PROMs across NSW.
- The program facilitated patient-centered and holistic care, with clinicians substantially altering patient care plans based on PROM data. Patients confirmed that their formal care plan had been altered (including introduction of new medication and an emergency response plan) using PROM data. Uncovering patient preferences or unknown health states (e.g. undiagnosed anxiety; difficulty navigating one's home environment in the absence of assistive technology) may have contributed to improved health outcomes through more appropriate care and support provision, and enabling patient engagement in care.
- PROMs were considered a useful triage tool, with data assisting to determine care provision, including timing of care, promoting initial conversations with the patient and carer, and ascertaining whether physical and other supports are required.
- Clinicians used longitudinal PROM data to assist with tracking change in patient's health status over time (e.g. function, mood, quality of life).
- Aggregated PROM data (including comparisons across sites) were used to refine service delivery (e.g. aggregate PROM data at a service used as a 'quality indicator'), ensuring that care was 'timely, efficient, and targeted correctly'. PROM data had, on occasion, led to updated models of service delivery or care.
- PROMs had the potential to reduce unplanned hospital admissions, and associated efficiency gains by ensuring that patients are provided the most appropriate treatment in primary care, uncovering unknown health states and barriers to service access, and enabling integration of care planning (and therefore more integrated care) across primary and acute care service sites.

Barriers to Realizing PROM Benefits

Issues with Information and Communication Technologies (ICT), especially lack of integration with existing data systems, consistently impacted on successful program implementation. ICT issues led to decreased clinician (and sometimes patient) engagement, as potential benefits of PROMs were not considered substantive to justify the additional burden (e.g. time) placed on clinicians (and patients), particularly in fast-paced, business-oriented primary care settings who often see many patients in short periods. Clinicians working in such environments will only embrace new technology when it is easy to navigate, linked to existing systems, does not duplicate other activities, and is likely to benefit patients and clinical staff. Conversely, staff who reported greater access and integration were more likely to use PROM data to guide patient care, and were more satisfied with program implementation.

Implementation at sites had been slowed by lack of inducement to change work practices and limited resources; potential benefits to clinicians and patients were insufficient to promote clinical practice change. The the ACI played a significant role in promoting behavior change (e.g. education, on-call support). There was variation in program understanding and engagement amongst clinical staff. Full implementation (i.e. data collected and used from all eligible patients) was most likely in sites where PROM data collection was comprehensively understood by all staff. Champions were key to program success as they assisted in maintaining program implementation in the absence of the ACI staff (i.e. after completion of initial training). Dedicated time allocated to implementation was essential to success. Finally, program implementation slowed, or halted, when trained staff moved from a practice or service. Despite quick training of new staff, the program sometimes lost momentum in the period between engagement and training.

Future Requirements for Realizing PROM Benefits

- Support for an integrated data management system across all services for collecting, analyzing and real time reporting on PROMs, rather than standalone systems.
- Information targeted for health professionals, summarizing the purpose of PROMs to allow for wider health professional engagement and ownership of the process, specifically for time-poor primary care staff.
- Easy to use, quick access 'how to' text or video guides for clinicians involved in collecting PROMs, accessed via apps or mobile devices.
- Script for administrative staff to explain the purpose of the PROMs and obtain patient consent.
- Training including online resources and assisting service providers who have successfully implemented PROMs to mentor other local providers.
- Flexibility for sites to use different methods (e.g. paper-based questionnaires) to collect PROMs depending on patient preference and accessibility issues.

Discussion

This mixed methods evaluation demonstrated that it is possible to implement PROM data collection and use across NSW care settings, with caveats noted below. Clinicians and site managers reported their services were using PROM data to guide patient care and improve service delivery. Feedback from stakeholders was supported by the PROM data, with 1300 PROMs completed at baseline over a one-year period. However, there were variations in completion rates between sites, with some completing 100% of their PROMs and others completing none. There was also missing PROM data, with only 69% of people eligible to complete PROMs providing responses to all PROMIS10 items at baseline, and low completion of PROMs at follow-up assessment time-points. Some loss to follow-up was expected due to patient drop out (i.e. not all patients were expected to be seen at all three assessments time-points). For some sites were follow-up was unexpected, it was unclear whether the low completion rate was due to patients being unable to complete follow-up PROMs or whether sites were not giving follow-up PROMs to patients to complete. Factors that may have contributed to

drop-out at follow-up times include: patients not returning to clinics or services, or declining further participation; clinicians not administering follow-up PROMs due to inadequate guidance/training or poor systems in place without reminders to prompt clinicians to administer PROMs; staff turnover and changes (which led to lack of continuity at the sites); and lack of funding for some programs and staff, resulting in clinics discontinuing PROMs administration. Medical condition was another key variable that had poor data quality, with 57.8% of sites failing to record this information. This data was optional (rather than mandatory) in the database and may therefore be the reason for the lower completion rate. On reflection, a broader accountability model for collecting and using PROMs and recording clinical data in the database could have been useful for monitoring completion rates in real time and providing support for sites with poor completion rates. Interviews with stakeholders suggested that PROM completion rates had fallen short of initial expectations due to lack of integration with ICT, limited inducement to change and resource availability, variable understanding and engagement amongst staff, and staff turnover.

During program design, the the ACI were aware that a stand-alone IT system would likely be problematic for site staff, however to test the innovative PRM program for collecting and using PROMs electronically, a system was required to test feasibility, system requirements, workflows, system interface, and integration aspects. The stand-alone IT system was not intended as a long-term solution but rather one that enabled the identification of system requirements for future PROM collection.

Depending on PROM scoring algorithms, certain PROMs require that all scale items be completed to enable a scale score to be generated (e.g. the PROMIS10); if one or more items are missing, the whole dataset for that particular scale cannot be used. Feedback from sites indicated that some patients did not understand certain questions or wanted to talk to a healthcare professional about it. Another limitation for the generalisability of completion rate learning is that PROM completion in the pilot was limited to English-speaking patients. In addition, nearly 20% who indicated they were English speakers, preferred PROMs in non-English languages. Our intention was to assess whether PROM collection was feasible across multiple sites and health contexts in NSW. Given the large number of non-English speaking participants identified, the extended PRM program (already underway) is inclusive of non-English speaking participants, and PROMs are available in 10 languages other than English. Moving forward to support the broader NSW Health PRM program, PRMs will be available in 15 languages other than English. We did not collect individual participant characteristics so were unable to assess differences in education level, social status, and financial status between English speaking and non-English speaking patients. However, it is well documented in the literature that minority and disadvantaged groups often have poorer health outcomes and inequity in access to health services(Goldstein et al. 2014; Sze et al. 2015). People from culturally and linguistically diverse communities or non-English speaking people are often excluded from research, health service evaluation, and routine PROM collection(Cella et al. 2015; Glickman et al. 2011), contributing to differences in health outcomes because of disparities in care or barriers to care access.

Missing PROM data is a challenge for data analysis and can compromise the interpretability and value of PRO findings(Mercieca-Bebber et al. 2016). For example, a high proportion of missing PRO data will substantially reduce the statistical power to detect a healthcare intervention effect and inflate standard error, increasing the risk of

type 2 errors, that is, false-negative findings(Fairclough et al. 1998a; Fairclough et al. 1998b). This is an important consideration when aggregating PROM data at the meso level for health service improvement, and at the macro level by health system policy makers. Perhaps more problematic are missing data related to the measured outcome (i.e. pain, depression, HRQOL)(Fairclough et al. 1998a) where non-completers may have more severe impairments and worse PRO scores. In other words, their non-completion may be due to poorer outcomes whereas completers may have better outcomes. This type of missing PRO data is often unavoidable, yet cannot be ignored as it may lead to biased estimates(Fairclough et al. 1998b). It is important to document reasons for non-completion of PRO assessments by patients as some missing data is avoidable, and if detected early, could be rectified such as in the case of a particular site forgetting to administer PROMs.

Looking at the various chronic condition sub-groups, mean trends indicated that people with back pain had worse physical health and stress scores compared to other medical conditions, whereas individuals with alcohol and other drug problems had worse mental health, COPD symptoms, depression, and anxiety compared to other medical conditions. Given the differences in PROs between chronic health conditions, our findings indicate the potential of PRO data for use in aggregate to describe the relative impact of different health conditions on various PROs and domains of HRQOL. With good quality data, we could have undertaken a regression analysis to further understand the effects of specific conditions and the burden of increasing co-morbidity, adjusting for age.

While it is too early in implementation to draw firm conclusions about program impact, the emerging evidence from the literature suggests that collection of PROMs could contribute to positive health outcomes for patients (e.g. HRQL, symptom control and survival), enhance service delivery, and reduce avoidable hospital admissions(Basch et al. 2017; Denis et al. 2017; Velikova et al. 2004; Barbera et al. 2015; Cleeland et al. 2011). Potential positive outcomes for patients and clinicians include use of PROMs in triage and tracking change in PROs over time, and improved engagement of patients in their own care. However, challenges to implementation and to embedding PROMs sustainably within the health system need to be addressed. Specifically, training and ongoing support for clinicians and patients in how to use PRO data and engagement of staff in the process so that it becomes part of daily clinical routine is needed. ICT infrastructure that allows seamless collection and analysis of PRO data and integration within existing patient information systems was considered critical for future PRM program success. A targeted change management approach to ensure that each element of the health system can adapt to accommodate the addition of PROMs into regular clinical routines was also considered beneficial. Our findings are consistent with others that report investment of time and resources in the design of the PROMs strategy, including co-design with key stakeholders and active involvement of professional organizations, and preparing organizations in how to use and interpret PRO data is essential to their success(Foster et al. 2018; Mejdahl et al. 2018; Greenhalgh et al. 2017). Early engagement and design considerations may prevent problems arising when PROMs are then implemented and used in clinical practice. Further, ensuring staff understand the value of completing PROMs over time even if the person improved is critical. PROMs are not only tools to use when a person is at the start of their health journey and adding value to care planning, but also evidence of the success (or failure) of the care received.

The findings from this pilot have implications for the future NSW Health PRM program. We plan to refine components of the program. Planned improvements for the IT solution include fuller integration with other health services (ideally via a single login), access to PROM survey results across care locations and members of care teams, and a patient portal to allow patients to access and view their own reports and recommended self-management information/resources. Further investment will be made in local capability and change-management activities to increase the knowledge of PROMs amongst NSW Health services and primary care – this will help to build sustainability and scalability. Finally, we will ensure that PROMs are available for people in languages other than English to reduce inequity and ensure greater inclusiveness. The potential positive impacts of the NSW PRMs program and a more comprehensive evaluation of consumer views will be undertaken in May 2019 during the summative phase of this evaluation.

Conclusions

This pilot study reveals the potential for PROMs to be collected routinely in NSW. High completion rates at baseline are promising, but strategies are needed to improve completion rates at follow-up assessments. This could be facilitated via more integrated technology solutions and building capability across care settings and stakeholder groups. It is anticipated that better understanding of the importance of routine collection and use of PROMs at various points of care and across care settings will move NSW Health towards Value Based Healthcare.

References

Aaronson, N.K., Elliott, T.E., Greenhalgh, J., Halyard, M.Y., Hess, R., Miller, D., Reeve, B.B., Santana, M., Snyder, C. (2015). User's Guide to Implementing Patient-Reported Outcomes Assessment in Clinical Practice. http://www.isoqol.org/UserFiles/2015UsersGuide-Version2.pdf. Accessed 11 March 2019.

Agency for Clinical Innovation (2019): Patient Reported Measures: Outcomes that Matter to Patients. https://www.aci.health.nsw.gov.au/make-it-happen/prms. Accessed 11 March 2019.

Australian Bureau of Statistics (2019): 2017–18 National Health Survey (NHS). https://www.abs.gov.au/ausstats/abs@.nsf/mf/4364.0.55.001. Accessed 16 Oct 2019.

Australian Institute of Health and Welfare (2018): Mental health services—in brief 2018. https://www.aihw.gov.au/getmedia/0e102c2f-694b-4949-84fb-e5db1c941a58/aihw-hse-211.pdf.aspx?inline=true. Accessed 16 Oct 2019.

Barbera, L., Sutradhar, R., Howell, D., Sussman, J., Seow, H., Dudgeon, D., Atzema, C., Earle, C., Husain, A., Liu, Y., & Krzyzanowska, M. K. (2015). Does routine symptom screening with ESAS decrease ED visits in breast cancer patients undergoing adjuvant chemotherapy? *Supportive Care in Cancer, 23*(10), 3025–3032. https://doi.org/10.1007/s00520-015-2671-3.

Basch, E., Deal, A. M., Dueck, A. C., Scher, H. I., Kris, M. G., Hudis, C., & Schrag, D. (2017). Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine Cancer treatment. *JAMA, 318*(2), 197–198. https://doi.org/10.1001/jama.2017.7156.

Cella, D., Hahn, E.A., Jensen, S.E., et al. (2015): Patient-Reported Outcomes in Performance Measurement. In: Method and Mode of Administration, Data Collection, and Analysis. RTI Press, Research Triangle Park (NC).

Chen, J., Ou, L., & Hollis, S. J. (2013). A systematic review of the impact of routine collection of patient reported outcome measures on patients, providers and health organisations in an oncologic setting. *BMC Health Services Research, 13*, 211. https://doi.org/10.1186/1472-6963-13-211.

Cheung, N.W., Crampton, M., Nesire, V., Hng, T.-M., Chow, C.K., on behalf of the Western Sydney Integrated Care Program Investigators (2019): Model for integrated care for chronic disease in the Australian context: Western Sydney Integrated Care Program. *Australian Health Review*. https://doi.org/10.1071/AH18152

Cleeland, C. S., Wang, X. S., Shi, Q., Mendoza, T. R., Wright, S. L., Berry, M. D., Malveaux, D., Shah, P. K., Gning, I., Hofstetter, W. L., Putnam Jr., J. B., & Vaporciyan, A. A. (2011). Automated symptom alerts reduce postoperative symptom severity after cancer surgery: A randomized controlled clinical trial. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, 29*(8), 994–1000. https://doi.org/10.1200/jco.2010.29.8315.

Denis, F., Lethrosne, C., Pourel, N., Molinier, O., Pointreau, Y., Domont, J., Bourgeois, H., Senellart, H., Trémolières, P., Lizée, T., Bennouna, J., Urban, T., El Khouri, C., Charron, A., Septans, A.-L., Balavoine, M., Landry, S., Solal-Céligny, P., Letellier, C. (2017). Randomized Trial Comparing a Web-Mediated Follow-up With Routine Surveillance in Lung Cancer Patients. *Journal of the National Cancer Institute 109*(9). https://doi.org/10.1093/jnci/djx029

Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., & Sutton, A. (2005). Synthesising qualitative and quantitative evidence: A review of possible methods. *Journal of Health Services Research & Policy, 10*(1), 45–53. https://doi.org/10.1177/135581960501000110.

Fairclough, D. L., Peterson, H. F., Cella, D., & Bonomi, P. (1998a). Comparison of several model-based methods for analysing incomplete quality of life data in cancer clinical trials. *Statistics in Medicine, 17*(5–7), 781–796.

Fairclough, D. L., Peterson, H. F., & Chang, V. (1998b). Why are missing quality of life data a problem in clinical trials of cancer therapy? *Statistics in Medicine, 17*(5–7), 667–677.

Foster, A., Croot, L., Brazier, J., Harris, J., & O'Cathain, A. (2018). The facilitators and barriers to implementing patient reported outcome measures in organisations delivering health related services: A systematic review of reviews. *Journal of Patient-Reported Outcomes, 2*, 46. https://doi.org/10.1186/s41687-018-0072-3.

Glickman, S. W., Ndubuizu, A., Weinfurt, K. P., Hamilton, C. D., Glickman, L. T., Schulman, K. A., & Cairns, C. B. (2011). Perspective: The case for research justice: Inclusion of patients with limited English proficiency in clinical research. *Academic Medicine : Journal of the Association of American Medical Colleges, 86*(3), 389–393. https://doi.org/10.1097/ACM.0b013e318208289a.

Goldstein, D., Bell, M. L., Butow, P., Sze, M., Vaccaro, L., Dong, S., Liauw, W., Hui, R., Tattersall, M., Ng, W., Asghari, R., Steer, C., Vardy, J., Parente, P., Harris, M., Karanth, N. V., King, M., Girgis, A., Eisenbruch, M., & Jefford, M. (2014). Immigrants' perceptions of the quality of their cancer care: an Australian comparative study, identifying potentially modifiable factors. *Annals of Oncology: Official Journal of the European Society for Medical Oncology, 25*(8), 1643–1649. https://doi.org/10.1093/annonc/mdu182.

Greenhalgh, J., & Meadows, K. (1999). The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: A literature review. *Journal of Evaluation in Clinical Practice, 5*(4), 401–416.

Greenhalgh, T., Jackson, C., Shaw, S., & Janamian, T. (2016). Achieving research impact through co-creation in community-based health services: Literature review and case study. *The Milbank Quarterly, 94*(2), 392–429. https://doi.org/10.1111/1468-0009.12197.

Greenhalgh, J., Dalkin, S., Gooding, K., Gibbons, E., Wright, J., Meads, D., Black, N., Valderas, J.M., Pawson, R. (2017): Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. *Health Services and Delivery Research, 5*(2). doi:https://doi.org/10.3310/hsdr05020

Hays, R. D., Bjorner, J. B., Revicki, D. A., Spritzer, K. L., & Cella, D. (2009). Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 18*(7), 873–880. https://doi.org/10.1007/s11136-009-9496-9.

Health and Social Policy (2018): NSW Health Strategic Framework for Integrating Care. https://www.health.nsw.gov.au/integratedcare/Publications/strategic-framework-for-integrating-care.PDF. Accessed 11 March 2019.

Ishaque, S., Karnon, J., Chen, G., Nair, R., & Salter, A. B. (2019). A systematic review of randomised controlled trials evaluating the use of patient-reported outcome measures (PROMs). *Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 28*(3), 567–592. https://doi.org/10.1007/s11136-018-2016-z.

Jang, S.M., Kim, K.U., Na, H.J., Song, S.E., Lee, S.H., Lee, H., Kim, Y.S., Lee, M.K., Park, H.-K. (2019). Depression is a major determinant of both disease-specific and generic health-related quality of life in people with severe COPD. *Chronic Respiratory Disease 16*. https://doi.org/10.1177/1479972318775422

Jones, P. W., Harding, G., Berry, P., Wiklund, I., Chen, W. H., & Kline Leidy, N. (2009). Development and first validation of the COPD assessment test. *The European Respiratory Journal, 34*(3), 648–654. https://doi.org/10.1183/09031936.00102509.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the depression anxiety stress scales* (2nd ed.). Sydney: Psychology Foundation.

Mejdahl, C. T., Schougaard, L. M. V., Hjollund, N. H., Riiskjaer, E., & Lomborg, K. (2018). Exploring organisational mechanisms in PRO-based follow-up in routine outpatient care - an interpretive description of the clinician perspective. *BMC Health Services Research, 18*(1), 546. https://doi.org/10.1186/s12913-018-3352-y.

Mercieca-Bebber, R., Palmer, M. J., Brundage, M., Calvert, M., Stockler, M. R., & King, M. T. (2016). Design, implementation and reporting strategies to reduce the instance and impact of missing patient-reported outcome (PRO) data: A systematic review. *BMJ Open, 6*(6), e010938. https://doi.org/10.1136/bmjopen-2015-010938.

NSW Ministry of Health (2018): NSW integrated care journey. https://www.health.nsw.gov.au/integratedcare/Pages/Our-Plan.aspx. Accessed 11 March 2019.

OECD (2010): Improving value in health care. https://doi.org/10.1787/9789264094819-en.

Porter, I., Goncalves-Bradley, D., Ricci-Cabello, I., Gibbons, C., Gangannagaripalli, J., Fitzpatrick, R., Black, N., Greenhalgh, J., & Valderas, J. M. (2016). Framework and guidance for implementing patient-reported outcomes in clinical practice: Evidence, challenges and opportunities. *Journal of Comparative Effectiveness Research, 5*(5), 507–519. https://doi.org/10.2217/cer-2015-0014.

Snyder, C. F., Aaronson, N. K., Choucair, A. K., Elliott, T. E., Greenhalgh, J., Halyard, M. Y., Hess, R., Miller, D. M., Reeve, B. B., & Santana, M. (2012). Implementing patient-reported outcomes assessment in clinical practice: A review of the options and considerations. *Quality of Life Research: an International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation, 21*(8), 1305–1314. https://doi.org/10.1007/s11136-011-0054-x.

Solano, J. P., Gomes, B., & Higginson, I. J. (2006). A comparison of symptom prevalence in far advanced cancer, AIDS, heart disease, chronic obstructive pulmonary disease and renal disease. *Journal of Pain and Symptom Management, 31*(1), 58–69.

Stanton, R., Rosenbaum, S., Rebar, A., & Happell, B. (2019). Prevalence of chronic health conditions in Australian adults with depression and/or anxiety. *Issues in Mental Health Nursing, 40*(10), 902–907. https://doi.org/10.1080/01612840.2019.1613701.

Sze, M., Butow, P., Bell, M., Vaccaro, L., Dong, S., Eisenbruch, M., Jefford, M., Girgis, A., King, M., McGrane, J., Ng, W., Asghari, R., Parente, P., Liauw, W., Goldstein, D., & Psycho-oncology co-operative research group, C., linguistically diverse, T. (2015). Migrant health in cancer: Outcome disparities and the determinant role of migrant-specific variables. *Oncologist, 20*(5), 523–531. https://doi.org/10.1634/theoncologist.2014-0274.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*, 45. https://doi.org/10.1186/1471-2288-8-45.

Velikova, G., Booth, L., Smith, A. B., Brown, P. M., Lynch, P., Brown, J. M., & Selby, P. J. (2004). Measuring quality of life in routine oncology practice improves communication and patient well-being: A randomized controlled trial. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology, 22*(4), 714–724. https://doi.org/10.1200/jco.2004.06.078.

Affiliations

Claudia Rutherford[1,2] · **Rachel Campbell**[1] · **Melissa Tinsley**[3] · **Robyn Speerin**[3] · **Linda Soars**[4] · **Anna Butcher**[5] · **Madeleine King**[1]

[1]   University of Sydney, Faculty of Science, School of Psychology, Sydney, NSW, Australia

[2]   University of Sydney, Sydney Nursing School, Cancer Nursing Research Unit (CNRU), Sydney, NSW, Australia

[3]   Agency for Clinical Innovation, Sydney, NSW, Australia

[4]   Western Sydney Local Health District, Sydney, NSW, Australia

[5]   Northern Sydney Local health District, Sydney, NSW, Australia