BMJ Open

Feedback of aggregate patient-reported outcomes (PROs) data to clinicians and hospital end users: findings from an Australian codesign workshop process

Olivia Francis Ryan ,[1] Shaun L Hancock ,[1] Violet Marion,[1] Paulette Kelly,[2] Monique F Kilkenny ,[1,3] Benjamin Clissold,[4,5] Penina Gunzburg,[6] Shae Cooke,[7] Lauren Guy,[8] Lauren Sanders,[9,10] Sibilah Breen,[1] Dominique A Cadilhac [1,3]

To cite: Ryan OF, Hancock SL, Marion V, *et al*. Feedback of aggregate patient-reported outcomes (PROs) data to clinicians and hospital end users: findings from an Australian codesign workshop process. *BMJ Open* 2022;**12**:e055999. doi:10.1136/bmjopen-2021-055999

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-055999).

Received 31 August 2021
Accepted 09 June 2022

For numbered affiliations see end of article.

Correspondence to
Professor Dominique A Cadilhac; dominique.cadilhac@monash.edu

ABSTRACT

Objectives Patient-reported outcomes (PROs) are increasingly used to measure the patient's perspective of their outcomes following healthcare interventions. The aim of this study was to determine the preferred formats for reporting service-level PROs data to clinicians, researchers and managers to support greater utility of these data to improve healthcare and patient outcomes.

Setting Healthcare professionals receiving PRO data feedback at the health service level.

Participants An interdisciplinary Project Working Group comprised of clinicians participated in three workshops to codesign reporting templates of summarised PRO data (modified Rankin Scale, EuroQol Five Dimension Descriptive System, EuroQol Visual Analogue Scale and Hospital Anxiety and Depression Scale) using a modified Delphi process. An electronic survey was then distributed to short list the preferred templates among a broad sample of clinical end users. A final workshop was undertaken with the Project Working Group to review results and reach consensus on the final templates.

Primary and secondary outcome measures The recommendation of preferred PRO summary data feedback templates and guiding principles for reporting aggregate PRO data to clinicians was the primary outcome. A secondary outcome was the identification of perceived barriers and enablers to the use of PRO data in hospitals. For each outcome measure, quantitative and qualitative data were summarised.

Results 31 Working Group members (19 stroke, 2 psychology, 1 pharmacy, 9 researchers) participated in the workshops, where 25/55 templates were shortlisted for wider assessment. The survey was completed by 114 end users. Strongest preferences were identified for bar charts (37/82 votes, 45%) and stacked bar charts (37/91 votes, 41%). At the final workshop, recommendations to enhance communication of PROs data for comparing health service performance were made including tailoring feedback to professional roles and use of case-mix adjustment to ensure fair comparisons.

Conclusions Our research provides guidance on PROs reporting for optimising data interpretation and comparing hospital performance.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ We used iterative codesign methods with interdisciplinary clinical or policy stakeholders who were considered the end users of the evidence we were to generate from this study. A modified Delphi process was used to reach consensus on the PRO summary data feedback templates and to establish guidance for the preferred reporting formats.

⇒ Strengths of our study include the engagement of clinical, academic, patient and government representatives who were members of the Working Group for the Delphi process and involvement of a wider group of stakeholders who completed an electronic survey.

⇒ The use of different commonly used PRO measures was a strength, as we trialled both condition-specific and generic PRO measures in an effort to increase the generalisability of our results across different clinical populations.

⇒ A limitation was our inability to include examples of the presentation of baseline and follow-up PRO data due to the sample data being derived from the stroke registry where PROs are only collected at a single time point after the acute event. We acknowledge that these data are important for a range of conditions such as chronic kidney disease or surgical interventions where progression may be important to monitor.

⇒ Due to the use of the condition of stroke as a case study our findings may lack generalisability for other clinical conditions. However, the principles are largely transferable. The next steps are to pilot and refine the recommended templates.

INTRODUCTION

Patient-reported outcomes (PROs) are data collected directly from the patient about their own health, without interpretation by a clinician or any other person.[1] The use of PROs in comparative-effectiveness research and clinical trials is well established.[2] There is increased recognition that PROs have the

potential to enhance clinical practice through directing the need for improvements in the safety and quality of healthcare.[3–5]

PRO data can be used at the individual-patient level as a means to inform clinicians on aspects of their health that are important to them[6] or at the health service level to enable clinicians to compare hospital performance or against achievable benchmarks to identify quality improvement opportunities.[7 8] The interpretation of aggregate, health service-level PROs data by clinicians and other hospital end users is the focus of our research.

Known challenges related to the interpretation and use of PRO data include scoring and scaling differences between measures as well as inconsistencies in the methods used to report PRO data back to clinicians.[9] To address these challenges, our preliminary research involved a scoping review to summarise the existing evidence related to preferred formats for PRO data feedback to clinicians.[10] While a single preferred format or approach to feedback PRO data to healthcare professionals was not identified, we could summarise some general guidance on how to design feedback formats to enhance the interpretation of summarised PRO data.

The aim of the current study was to build on our preliminary research and establish guidance for preferred formats for hospital comparisons of PRO data to support greater utility of these data to improve healthcare and patient outcomes. In terms of quality of care, PROs can have relevance to clinicians if these data are considered in addition to information about performance on standards of clinical care (eg, expected interventions to be provided) and other information such as mortality and readmissions. For example, multicomponent audit and feedback approaches whereby PROs data are used to complement clinical indicator data can have an additive effect and contribute to better care through improving adherence to clinical guidelines.[11] Other authors[12] have acknowledged that there is currently a lack of primary studies and randomised controlled trials related to the effectiveness of aggregate PRO data for benchmarking and quality improvement purposes and the associated impact on patient outcomes.[13] Despite this, PRO feedback is an emerging field and our research constitutes an important development in this area. We also sought to gain preliminary insights into the perceived barriers and enablers to the use of PRO data in hospitals.

METHODS

Study design

This was a codesigned project, whereby the methods included several interdisciplinary end users working together to produce the outcomes.[14] We defined end users as hospital clinicians from a range of disciplines, academics who analyse and report PRO data and government representatives who use PRO data to inform quality improvement and policy decisions. We used a modified Delphi process[15] to reach consensus on PRO summary data templates and establish guidance for preferred reporting formats. The Delphi method is a group communication process, which aims to achieve convergence of opinions on a specific real-world issue.[15] The current study was designed in three stages: (1) a series of three workshops whereby members of a Project Working Group (herein referred to as Working Group) iteratively refined several PRO data feedback templates, (2) an anonymous electronic survey disseminated widely across Australia via clinical and academic networks to enable the short listing of the codesigned PRO summary data templates, (3) a final workshop with the Working Group members to achieve consensus on the recommended set of PRO summary data reporting format characteristics (figure 1).

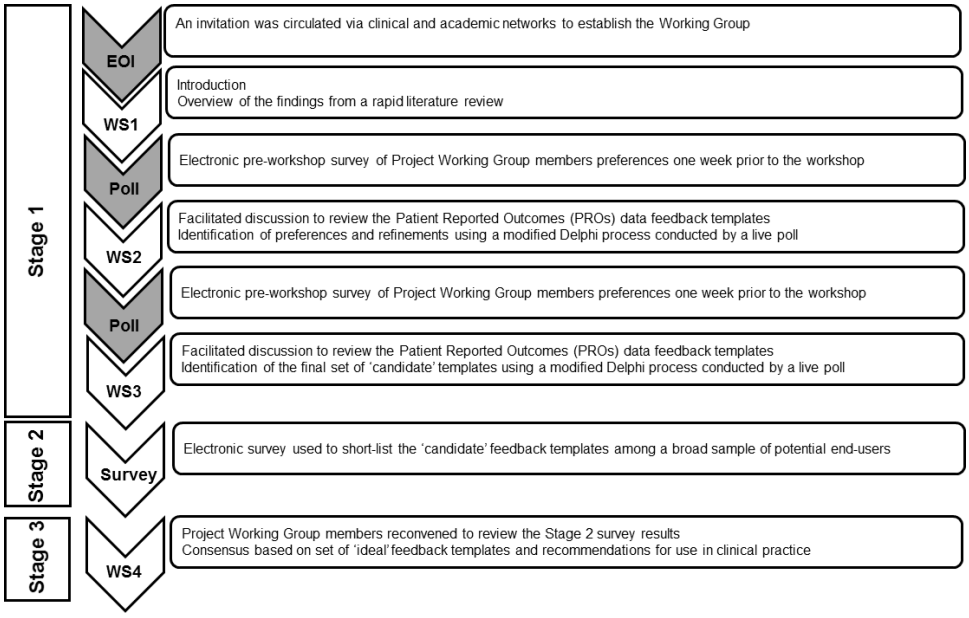


Figure 1 Diagrammatic representation of project stages. EOI: expression of interest; WS: workshop

Setting

This study was commissioned by the Victorian Agency for Health Information (VAHI) as part of a broader strategy to facilitate the comprehensive collection of PRO data in Victoria, Australia. To inform the state-wide strategy, VAHI commissioned consultancy work by Paxton Partners in 2018, whereby the authors identified that regular feedback of PRO data had the potential to add substantial value to improve clinical engagement with these data (Paxton Partners, PRO measures (PROMs): Literature Scan, personal communication, 2019). In response, VAHI appointed a project team comprising representatives from The Florey, Safer Care Victoria and Monash University to build on this work (see online supplemental file 1).

Currently, published evidence related to the systematic use of aggregated PRO data for quality improvement initiatives and policy decisions among clinicians and other stakeholders, such as government representatives, is sparse. The increasing ability to collect PROs and the proliferation of Clinical Quality Registries internationally,[16] has increased the need to establish rigorous evidence regarding the best feedback methods and preferred graphical formats for reporting PROs data to clinicians and other stakeholders . This current knowledge gap exists even though there is a strong policy rationale to support high functioning, mature Clinical Quality Registries with the ability to feedback data (both clinical and PROs)[17] and despite the fact that promoting the collection of PRO data is published as a priority for Australia in the 2020–2025 National Health Reform Agreement.[18]

The current study was initiated to assess the perceived barriers and opportunities for uptake of PRO data among clinicians and to establish best practice feedback approaches to be adopted by Clinical Quality Registries. Since 2012, PROMs (generic and stroke-specific) have been collected by The Florey as part of its role in data management processes of the Australian Stroke Clinical Registry (AuSCR).[19] Therefore, the condition of stroke was chosen as a case study.

Participants

The Working Group was established following the circulation of an expression of interest by the Victorian Stroke Clinical Network. Our objective was to achieve broad, interdisciplinary representation of experts and those with direct interests in the transparent feedback of PRO data. Individuals who were based in metropolitan or regional areas in Victoria and working in stroke care from acute, subacute and community settings were eligible to self-nominate for membership. We also included a consumer representative who was a survivor of stroke and had experience working with the Victorian government on various committees (see online supplemental file 1).

Stage 1: codesign workshops

In stage 1, we conducted three workshops between August and October 2019, with the option for participants to attend face-to-face or via the videoconferencing platform Zoom.[20] The meetings were chaired by authors DAC/SB and the content discussion, for example, data templates were facilitated by an experienced moderator (VM). The first workshop oriented the Working Group members to the aims of the current study and was used to provide an outline of the general reporting principles from the existing literature.[10] The identification of priorities and opinions was facilitated through the use of open-ended questions (eg, '*Who has used PROs data reported for their health service?*', '*Do you think PROs data add value, why/why not?*'). At the conclusion of the first workshop, it was agreed that the Florey project team (authors OFR, SLH, VM, SB and DAC) would develop a set of templates based on three PROMs commonly used in clinical practice. These included: the modified Rankin Scale (mRS),[21] EuroQol Five Dimensions Three Levels (EQ-5D-3L)[22] and the Hospital Anxiety and Depression Scale (HADS)[23] (see online supplemental table I). The PROMs were pragmatically chosen based on the experience of the project team with input from VAHI representatives.[24] All templates used fictitious hospital, state and/or national-level data. For the templates designed to facilitate benchmarking of PRO data, peer hospitals were defined as those facilities with a similar bed size/number of admissions per year and population served. For these types of templates, the data from all eligible peer hospitals were included as a comparison. The number of peer hospitals is specified in the relevant graph axis label, figure legend or instructional footnote.

The initial set of templates developed after the first workshop were circulated to the Working Group prior to the second workshop, via an electronic survey developed using SurveyMonkey software.[25] The Working Group members then ranked the templates according to preference. During the second workshop, members discussed the results of the survey and critiqued the templates to identify the prevailing preferences using Zoom live polling.[20] Refinements to the templates were made, then circulated via an electronic survey prior to the third workshop. The third workshop mirrored that of the second, whereby the survey results were discussed and members could revise their preferences. Agreement was reached for the final set of templates to be shortlisted in stage 2 (see online supplemental files 2–9). All workshops were audio recorded. At the conclusion of each workshop, formal minutes and the aggregated survey results were circulated to all members.

Stage 2: anonymous electronic survey

An anonymous electronic survey, from a broad sample of potential PRO data end users who had not been part of stage 1, was used to evaluate the feedback templates finalised in stage 1. The survey was designed using SurveyMonkey software.[25] Survey respondents were recruited using a convenience sampling approach via invitations sent from several relevant clinical and academic networks. The survey link included instructions and outlined the

voluntary participation requirements. Eligibility was self-determined by individuals who identified as a healthcare professional or academic/clinical researcher interested in the use of PRO data. Where feasible, snowball sampling was also encouraged, whereby respondents could share the link with other colleagues. The survey remained open for four weeks. Reminders to the main dissemination groups were sent one week prior to the closure of the survey.

The survey included questions on the characteristics of respondents, as well as feedback on each of the PRO templates for each specific measure, including mRS, EQ-5D-3L and the HADS. A set of templates that presented PRO data over time were also included. Using previous PRO-related stakeholder survey methods reported by Brundage *et al*,[26] a series of scale response questions (0=least, 10=most) were asked to ascertain: (a) the respondents' rating of the perceived utility: '*How useful do you find this graph?*', (b) the respondents' rating of their ability to understand each sample format: '*How easy it is for you to understand this graph?*'. The respondents then selected the preferred template out of all the options for each category and could provide a free-text comment. Multiple-choice and open-ended questions were included to elicit perceived barriers and enablers related to the use of PRO data. There were three survey questions specific to barriers and enablers, for example, '*What are the main barriers to the use of PROMs in your clinical practice?*'.

Stage 3: final consensus workshop

The Working Group was reconvened for a fourth (final) workshop to review the results of the anonymous electronic survey. The workshop discussion was guided by the same moderator (VM) using open-ended questions to achieve a final consensus on the recommended PRO feedback templates (eg, '*Could this graph be modified?*', '*Do you agree this format is the preferred template for this PROM, why/why not?*').

Statistical analysis

The survey data collected as part of stage 1 (preworkshop and live polls) and stage 2 (electronic survey) were extracted and summarised using Microsoft Excel (V.2016).[27] The quantitative data were summarised using descriptive analyses (totals and proportions). The stage 2 anonymous survey scale data were summarised as medians and IQRs. The stage 2 survey data were analysed overall, as well as according to clinical versus non-clinical respondents. Clinical respondents were defined as those with a medical, nursing or allied health background. Non-clinical respondents were defined as health system administrators, policymakers or researchers. The free-text qualitative data collected as part of the stage 2 survey were analysed using thematic analysis.[28] An inductive approach was used, whereby one member of the project team identified broad theme categories related to each question (VM). A second reviewer (OFR) cross checked the qualitative responses to verify the themes. Both the qualitative and quantitative data were used with the objective of determining prevailing preferences and preferred formatting attributes.

Table 1 Demographic characteristics of the project working group (N=33)*

Characteristic	Clinicians N=19 n (%)	Project team N=13 n (%)
Male	5 (26)	2 (15)
Profession		
Doctor	4 (21)	–
Nurse	5 (26)	–
Allied health	8 (42)	–
Other†	2 (11)	13 (100)
Clinical area of expertise		
Stroke	16 (84)	13 (100)
Other	3 (16)	–
Median years in clinical practice (Q1, Q3)	12.5 (8, 15)	–

*Includes one male consumer representative.
†Other: refers to professional backgrounds including psychology and pharmacy.

Patient and public involvement

The main consumer group for this research were clinicians for which the data summary reports are focused. A patient consumer representative was included as part of the Working Group for this project. No other public representatives or patients were involved in the overall study. Clinician representatives were asked to advise on the interpretation and write up of the results, and several members of our working group were included as coauthors on this article.

RESULTS

The Working Group comprised 33 members (19 clinicians, 1 consumer representative, 13 project team members) (table 1). The majority of clinical representatives were women (74%) and from an allied health (42%) or nursing (26%) profession. Three members were from other professional backgrounds, including psychology and pharmacy. Working Group members from a clinical background had a median of 12.5 years in clinical practice. Stroke management was the field of expertise for 84% (n=16) of the clinical representatives.

Stage 1: codesign workshops

Overall, 31 Working Group members actively participated in one or more workshops (see online supplemental file 1). The number and type of templates presented during each workshop are summarised in online supplemental table II. The templates were refined or retired based on the preferences identified during each workshop.

The formatting preferences included: simple layouts or symbols to reduce complexity, use of definitions and instructions (where appropriate), use of normative population data, CIs and denoting sample sizes (see online supplemental table III). Several templates were determined to be unsuitable for service-level PRO reporting, including spider plots, heat maps and pictographs. The Working Group agreed these formats were more complex and would limit the ability to readily interpret the results, while pictographs were considered too simplistic with insufficient detail for this audience. Following the third workshop, a separate category containing templates used for reporting PROs data over time was supported for wider testing, resulting in six additional templates.

Stage 2: anonymous electronic survey

There were 114 respondents for this survey (table 2). The majority of respondents were women (76%) and aged between 30 and 49 years (56%). Respondents were predominantly clinical representatives (91%), including allied health (38%) or nursing (32%). Overall, 87 respondents (71%) reported stroke as their area of clinical expertise. The respondents had a median of 20 years in clinical practice, and 23% reported using PROs at the time of the survey. The following sections detail the results for each PRO category assessed as part of the survey. The denominators vary since not all respondents completed every question used to evaluate the templates. The overall response rate for the survey was 61%.

Table 2 Characteristics of the stage 2 electronic survey respondents

Respondent characteristics	Survey respondents N=114 n (%)
Female	87 (76)
Age group (years)	
<30	6 (5)
30–49	64 (56)
50–64	40 (35)
65+	4 (4)
Profession	
Clinical:	104 (91)
Doctor	24 (21)
Nurse	37 (32)
Allied health	43 (38)
Non-clinical:	10 (9)
Management	1 (1)
Researcher	4 (3)
Government representative	1 (1)
Health system administration	2 (1)
Quality/safety	2 (1)
Clinical area of expertise	
Stroke	81 (71)
Cancer	5 (4)
Other (ie, rehabilitation, geriatrics etc.)	28 (25)
Current work role*	
Clinical care	82 (71)
Education	27 (23)
Research	47 (41)
Management	26 (23)
Other (eg, quality/safety, integrated cancer service etc.)	6 (5)
Median years in practice (Q1, Q3)	20 (10, 29)
Completed a higher degree (master's or doctorate)	71 (62)
Currently use PROMs in clinical practice	26 (23)

*Multiple responses were permitted for this question therefore percentages do not add up to 100.
PROMs, patient reported outcome measures.

Preferences for graphical formats of summary data

When interpreting mRS data, overall, 41% (n=37/91) of survey respondents preferred format A (see online supplemental files 2 & 3) which reflected a two-bar horizontal stacked bar chart with comparative data (table 3). The prevailing template had the greatest rating for ability to understand the data being presented (median=9), despite only a two-point median difference between the five templates presented (see online supplemental table IV). The responses of non-clinical respondents were similar to clinical respondents, with a preference for an additional table with mRS data values paired with the stacked bar chart.

For EQ-VAS data, 45% of respondents (37/82) preferred format A (see online supplemental file 4) comprised of a bar chart with the inclusion of CIs and stratification by patient sex and age group. Compared with the more complex formats (eg, format C: caterpillar plot), the bar chart resulted in the greatest rating (median=7) for both perceived ease of understanding and usefulness (see online supplemental table IV). The median ratings across the three EQ-VAS templates were similar, with a maximum median difference of two points only. The responses of non-clinical respondents differed to clinical respondents, with a preference for a dot or caterpillar plot.

When interpreting EQ-5D-3L data, format F (see online supplemental files 5 & 6), which comprised a stacked bar chart presenting data for all five dimensions (with categorisation according to the proportion of patients reporting 'no problems' vs those reporting 'problems'), was identified as the preferred format (22/78, 28%). Despite the greatest proportion of respondents preferring format F, there was a diverse range of median ratings across all seven templates. Several anomalies were identified whereby four other templates (eg, formats A and B with vertical stacked bar chart and traffic light colour

Table 3 Respondent preferences for templates displayed in stage 2 survey, overall and by profession

Patient-reported outcome measure	Template reference*	Overall N=114 n/N	Medical N=19 n/N	Nursing N=33 n/N	Allied Health N=43 n/N	Other N=8 n/N
Modified Rankin Scale	Format A	**37/91 (41%)**	**10/19 (53%)**	**15/33 (46%)**	**12/31 (39%)**	0
	Format B	21/91 (23%)	4/19 (21%)	5/33 (15%)	7/31 (23%)	**5/8 (63%)**
	Format C	12/91 (13%)	1/19 (5%)	5/33 (15%)	6/31 (19%)	0
	Format D	11/91 (12%)	1/19 (5%)	4/33 (12%)	5/31 (16%)	1/8 (12%)
	Format E	10/91 (11%)	3/19 (16%)	4/33 (12%)	1/31 (3%)	2/8 (25%)
EuroQol Visual Analogue Scale	Format A	**37/82 (45%)**	**9/19 (47%)**	**18/30 (60%)**	9/26 (35%)	1/7 (14%)
	Format B	29/82 (35%)	6/19 (32%)	10/30 (33%)	**10/26 (38%)**	**3/7 (43%)**
	Format C	16/82 (20%)	4/19 (21%)	2/30 (7%)	7/26 (27%)	**3/7 (43%)**
EuroQol Five Dimension Descriptive System	Format A	15/78 (19%)	2/19 (11%)	**9/30 (30%)**	4/23 (17%)	0
	Format B	10/78 (13%)	1/19 (6%)	5/30 (17%)	4/23 (17%)	0
	Format C	3/78 (4%)	0	3/30 (10%)	0	0
	Format D	7/78 (9%)	4/19 (22%)	1/30 (3%)	2/23 (9%)	0
	Format E	14/78 (18%)	4/19 (22%)	4/30 (13%)	2/23 (9%)	**4/6 (67%)**
	Format F	**22/78 (28%)**	**5/19 (28%)**	7/30 (23%)	**8/23 (35%)**	2/6 (33%)
	Format G	6/78 (8%)	2/19 (11%)	1/30 (3%)	3/23 (13%)	0
Hospital Anxiety and Depression Scale	Format A	14/75 (19%)	2/19 (11%)	**9/28 (32%)**	3/22 (14%)	0
	Format B	**24/75 (32%)**	**7/19 (39%)**	7/28 (25%)	**8/22 (36%)**	2/6 (33%)
	Format C	17/75 (23%)	5/19 (28%)	5/28 (18%)	6/22 (27%)	1/6 (17%)
	Format D	19/75 (26%)	4/19 (22%)	7/28 (25%)	5/22 (23%)	**3/6 (50%)**
Longitudinal data	Format A	7/69 (10%)	1/17 (6%)	5/28 (18%)	1/19 (5%)	0
	Format B	14/69 (20%)	**5/17 (29%)**	5/28 (18%)	4/19 (21%)	0
	Format C	3/69 (4%)	0	2/28 (7%)	1/19 (5%)	0
	Format D	12/69 (17%)	4/17 (24%)	3/28 (11%)	4/19 (21%)	1/5 (20%)
	Format E	**21/69 (30%)**	4/17 (24%)	**9/28 (32%)**	**5/19 (26%)**	**3/5 (60%)**
	Format F	12/69 (17%)	3/17 (17%)	4/28 (14%)	4/19 (21%)	1/5 (20%)

Bolded figures represent the overall preference of ranking for each respondent type (displayed by profession) category according to the survey data. Other: refers to non-clinical professions such as health service/department managers, researchers, etc. The denominators used for each row proportion differ due to survey respondents terminating the survey at different levels of completion.
*See online supplemental files 2-9 for a copy of the templates used for each format.

coding) received a median rating of ≥8 for perceived ease of understanding, which was greater than that of the prevailing template. The preference of non-clinical respondents differed from clinical respondents, with format E comprised of a table of data values paired with a horizontal bar chart prevailing. In addition, the preferred format (24/75, 32%) for presenting HADS data was format B, a bar chart with the inclusion of a dotted line to depict a clinical cut-off point, which indicated possible anxiety or depression (see online supplemental file 7).

A total of 21/69 (30%) respondents preferred format E (see online supplemental files 8 & 9), a line graph template for presenting data over time. The line graph that was short listed as the preference displayed hospital data and comparative state-level hospital data, with additional metrics to stratify results according to patient characteristics (eg, stroke type) as well as a dotted line to indicate normative population estimates. Although this template prevailed, there were two other variations of a line graph (formats A and B) that had identical median ratings (understanding=8, usefulness=7), along with two bar graphs (formats C and D). The consistent median ratings across all formats did not align with the differences observed between the overall rankings of preference. There was no difference for the overall preference between clinical and non-clinical respondents.

Several themes emerged from the qualitative data related to formatting preferences, including: the use of colours, particularly intuitive colours such as green=better/good outcome and red=worse/poorer outcome (see online supplemental table III). Though, the use of dark/shaded colours was cautioned to ensure templates would be interpretable in grayscale (eg, to facilitate printed versions for hospital end users). In addition, the inclusion of clear

labelling (namely graph axes and titles), sample sizes and proportions were also identified as preferred features. Though, the inclusion of too much explanatory text was cautioned. The respondents considered stratification of PRO summary data by age, sex and other variables to be useful, particularly to assist with targeted quality improvement strategies.

The inclusion of benchmarked peer-level, state-level and/or national-level data was also supported to facilitate hospital performance monitoring. The respondent's qualitative responses identified that benchmarks provided information about targets while peer data allowed for comparison against similar hospitals and within jurisdictions. The respondents also preferred the term 'peer hospitals' to be clearly defined. The need to provide case-mix adjusted data was identified by several respondents across various templates, with the importance of this statistical adjustment highlighted for the presentation of longitudinal PRO data (eg, otherwise variation could be due to the difference in patient characteristics each year).

PRO data reporting interval preferences and perceived barriers and enablers

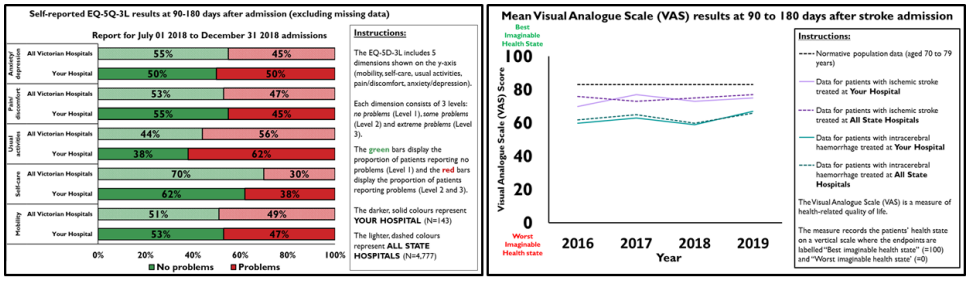
The majority of respondents (60/69, 87%) preferred a 3 to 6-month interval for PRO data reporting. The main perceived barrier to using PRO data in clinical practice from the summarised qualitative data, included resource and time constraints (35/68, 51%) (including staff time to read and interpret the information). Other perceived barriers included a lack of tools available to facilitate meaningful reporting of PROs data, a lack of understanding about how to use the information and lack of organisational support to use the data or a perceived need for culture change. Two of the most common themes identified as enabling factors included: an interest to use PROs as a mechanism to deliver patient-centred care, and the development of enhanced feedback methods to facilitate greater PRO data uptake and future use. Other reported enablers included the need for adequate resources such as time and funding for clinicians to interpret and use the data and enhanced education about how best to interpret and use the data to improve patient care.

Stage 3: Final consensus workshop

The Working Group agreed with the outcomes of the stage 2 survey results and the shortlisted templates. The median ratings for each template, especially for formats where anomalies, were identified in contrast to the overall preference, were used as a supplement for the Working Group to consider. Figure 2 illustrates two templates recommended for presenting: data at a single time point versus data over time, along with a summary of preferred formatting features. An example of a template that was identified to not be appropriate for routine feedback purposes is presented in figure 3. The Working Group agreed that the professional roles of the recipients of these data would require consideration, and that certain recipients (eg, hospital executives) may require further detail such a pairing a graphical display with a table of data values. The Working Group discussed several data quality aspects of PROs (eg, the need for case-mix adjustment) which were supported for incorporation into the final project recommendations (see online supplemental table V).

DISCUSSION

PROs data have the potential to improve healthcare and patient outcomes,[29 30] but in order for this to occur, it is essential that PROs data are presented in a manner that is both useful and understandable.[31 32] The presentation of aggregated PROs data is often done in the absence of best practice guidance.[5] Guided by evidence from our previous literature review,[10] this study aimed to use a codesign process to create a set of PRO summary data



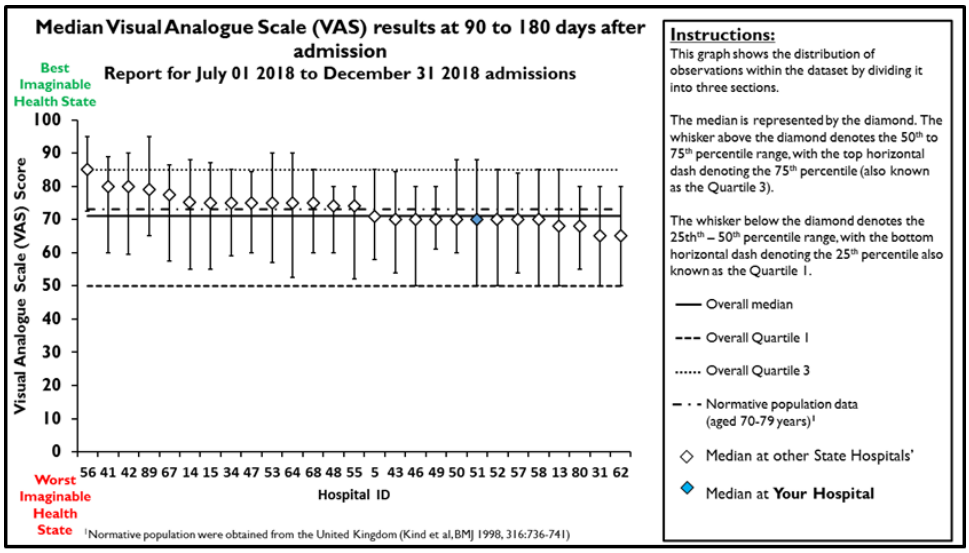
Example of a format determined to be appropriate to display Patient Reported Outcomes data from a single time-point

Recommended template features include:
- Simple format with large font (reduce cognitive burden)
- Clear title, axis labels and instructions
- Allow adequate space between bars
- Denote sample size
- Avoid use of acronyms (spell out in full if used)
- Use of white space
- Intuitive colours (i.e. to facilitate printing in grayscale in a health service setting)

Example of a format determined to be appropriate to display Patient Reported Outcomes data over time

Recommended template features include:
- Normative population data as a reference (fictitious example shown)
- Include comparative data (e.g. state-level hospital data shown)
- Use visual cues to facilitate interpretation of scores (e.g. best/worst imaginable health state labels shown)
- Clear title, axis labels and instructions
- Use of white space

Figure 2 Example illustrating two recommended templates to use for reporting aggregate, service-level patient reported outcomes data to clinicians.



Example of a format determined to be too difficult to interpret in clinical practice

Template features that add complexity include:

- Statistical concepts (e.g. median, interquartile range)
- Inclusion of normative population data as a reference
- Unfamiliar caterpillar plot format
- Multiple explanatory notes that may require additional time to understand
- Less white space compared to other templates

Figure 3 Example illustrating a non-preferred template for reporting aggregate, service-level patient reported outcomes data to clinicians.

templates for health service-level reporting. The recommended templates and preferred formatting attributes identified from our study complement existing mechanisms proposed by Snyder *et al*[31] and contribute much needed evidence for methods to use when communicating service-level PRO data to clinicians.

Despite some mixed opinions among the clinical representatives related to each template, a clear preference emerged to support the provision of these data. Our findings indicate that the simpler formats reduced the cognitive burden for clinicians, and there was a preference for use of explanatory attributes to aid interpretation of the data. For data presented at a single time point, a bar or stacked bar chart with up to four bars prevailed. A caveat to this recommendation was for PROMs with multiple dimensions, such as the EQ-5D, where clinicians preferred to review local and comparison hospital data within the one graph (up to 10 bars). A line graph with a maximum of four lines and minimal additional constructs was identified as the preference for data presented over time. A construct can include statistical or formatting features such as: stratification according to patient demographic/clinical characteristics or providing a reference to normative population data or clinical cut-offs.[33] Our main findings are consistent with the preliminary principles for graphical display of PRO data proposed by Bantug *et al*[5] and builds on the 'less is more' guidance recommended for effective communication of performance monitoring[34] or clinical data in general.[35]

The diversity of opinions among clinicians in the Working Group and broader survey sample in our study is not dissimilar to previous research. Brundage *et al*[1] evaluated the preferences of graphically displayed PRO results among 233 clinicians and found differing preferences for bar and pie charts. In the current study, we also found that clinicians had differing preferences related to the use of pie charts compared with bar charts as well as intuitive colour coding schemes[36 37] (eg, some clinicians indicated that traffic light colours added visual clarity, whereas others disliked them). Furthermore, the discrepancies identified from the median ratings in this current study, in which several templates received a rating of >8 despite not being identified as the prevailing format, highlight the challenge of recommending a single format style for all target audiences. Despite some mixed views, and mostly representation from professionals working in the clinical area of stroke, the recommendations from this current study form an empirical basis for future work in this field. We recommend that there is a need to involve hospital end users, including representatives from quality improvement departments, if aggregate PRO data are being prepared by academics, and that these data should be presented in more than one way to address the needs of specific audience subgroups. The barriers and enablers identified in this current study, including resource and time constraints and education for clinicians, align with the findings of our previously conducted scoping review into PRO feedback.[10]

We also found a preference for the majority of the shortlisted templates to include comparative hospital data. In our examples, we followed the approach used in the AuSCR to provide peer-hospital, state-level and national-level benchmarks. Case-mix adjustment for PROs data is an emerging field[38] and the application of case-mix methods is a prerequisite for quality reporting and benchmarking purposes. Further empirical research is needed to identify patient-related (eg, age, sex, life events, new healthcare episodes) and hospital-related (workload, volume of patients, hospital type, etc) factors requiring adjustment for reporting benchmarked PRO data.[12] It was intriguing that only a single survey respondent considered the importance of case-mix adjustment, and this might suggest that education for clinicians about the importance of case-mix adjustment when reviewing comparisons of PROs at the service level be undertaken. This current study contributes to the discourse in confirming the need for ongoing methodological research to determine the most appropriate analytic methods and the variables needed to enable reliable case-mix adjustment. In the interim, it is recommended that when providing descriptive PROs comparisons without the ability to adjust for differences in patient characteristics, a concise and explicit definition of what constitutes a peer hospital must be included in the template.

Strengths and limitations

The strengths of our study include the engagement of an interdisciplinary sample of clinical, academic and government representatives who were members of the Working Group, and other potential end users who completed the electronic survey. This is one of only a few studies in the field, outside of a cancer setting,[1 9 32] to iteratively develop and evaluate the preferences of clinicians when considering the use of PROs for comparing hospital performance. Applying a modified Delphi process throughout each stage was also a strength as well as the use of multiple engagement tools to optimise participation of Working Group members (eg, live polling). An additional strength was the use of different commonly used PROs measures as part of the templates. While previous work has focused on the development of templates for PRO measures for a single condition,[26] we trialled both condition-specific and generic PRO measures in an effort to increase the generalisability of our results across different clinical populations.

The limitations of our study include that our templates were based on experience reporting PROs as part of the AuSCR and informed by the literature review. Some subjective interpretation of how to display data and how many templates to produce for assessment was unavoidable. Although we sought broad representation for the survey by advertising widely, we acknowledge that the majority of the Working Group were experienced in the field of stroke and were located within a single jurisdiction. A further limitation was our inability to include examples of the presentation of baseline and follow-up PRO data due to the sample data being derived from the stroke registry where PROs are only collected at a single time point after the acute event. We acknowledge these data are important for a range of conditions such as chronic kidney disease or surgical interventions, where progression may be important to monitor. We acknowledge the presence of responder bias in our survey data and that the ability to detect preferences among different professional audience groups (eg, medical/nursing staff vs government representatives/researchers) was limited by the composition of mostly female respondents who self-selected to participate in the survey. Despite this, our findings align with those from preliminary work conducted in other patient populations,[9 39] which suggests that there are general principles for the presentation of aggregate PRO data that are applicable across clinical specialities. We acknowledge that the next steps are to pilot and refine the recommended templates and encourage others to consider evaluating these as well.

CONCLUSIONS

We have illustrated the iterative process and outcomes of a codesigned approach to establishing summary data templates for reporting aggregate service-level PROs data to clinicians. Simple graphical templates, with accompanying instructions and formatting attributes to aid data interpretation, were identified as the preferred format characteristics. This work provides important evidence for Clinical Quality Registries and other organisations that routinely feedback aggregate PRO data, to ensure that the potential of these data to support quality improvement efforts is fully realised.

Author affiliations
[1]Public Health: Stroke Division, Florey Institute of Neuroscience and Mental Health—Austin Campus, Heidelberg, Victoria, Australia
[2]Victorian Agency for Health Information, Victoria Department of Health and Human Services, Melbourne, Victoria, Australia
[3]Stroke and Ageing Research, Department of Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia
[4]Neurosciences Department, Monash Health, Clayton, Victoria, Australia
[5]Neurosciences Department, Barwon Health, Geelong, Victoria, Australia
[6]Physiotherapy Department, Alfred Health, Melbourne, Victoria, Australia
[7]Department of Physiotherapy, Eastern Health, Box Hill, Victoria, Australia
[8]Community Based Rehabilitation, Sunshine Hospital, Saint Albans, Victoria, Australia
[9]Department of Neurosciences, St Vincent's Hospital Melbourne Pty Ltd, Fitzroy, Victoria, Australia
[10]Department of Medical Education, The University of Melbourne - Parkville Campus, Melbourne, Victoria, Australia

Twitter Monique F Kilkenny @KilkennyMonique and Dominique A Cadilhac @DominiqueCad

Acknowledgements We acknowledge the members of Patient Reported Outcome Measures Working Group (see Supplementary 1). We also thank the hospital clinicians, executive and administrative staff, policy makers, consumer representative, researchers and clinical registry special interest groups who contributed to the anonymous electronic survey (see Supplementary 1). We acknowledge the contribution of hospital staff from sites for collecting data based on patients registered in the AuSCR.

Contributors OFR, SLH, VM, PK, SB and DAC were involved in the planning, conceptualisation, design and execution of the project. PK MFK, BC, PG, SC, LG, LS provided critical input to the interpretation of the results and were members of the project working group. All authors (OFR, SLH, VM, PK, MFK, BC, PG, SC, LG, LS, SB, DAC) contributed to the final version of the manuscript, revised it critically for important intellectual content and provided final approval of the version to be published. Authors OFR and DAC had final responsibility for the submission of this work for publication. DAC is responsible for the overall content as guarantor.

Funding This study was funded by the Victorian Agency for Health Information (VAHI). At the time of this study, the AuSCR received funding from the Victorian government, as well as other state governments, and various non-government and industry sources. DAC received Research Fellowship support from the NHMRC (1154273).

Competing interests DAC: is the Data Custodian for the AuSCR. All other authors report no potential conflicts of interest with respect to the research, authorship and/or publication of this article. National and Victorian clinical networks were also drawn upon to contribute to this research.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Ethics approval This study received approval from the Monash University Human Research Ethics Committee (Project ID: 27922).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Raw survey data are available on reasonable request to the corresponding author. Raw survey data of this study are not deposited in any public repository. Qualified investigators and researchers can request access to raw survey data in an anonymised format from this study after ethics clearance and approval is received. All other data relevant to the study are included in the article or supplemental material.

ORCID iDs
Olivia Francis Ryan http://orcid.org/0000-0003-4977-6742
Shaun L Hancock http://orcid.org/0000-0002-2015-2752
Monique F Kilkenny http://orcid.org/0000-0002-3375-287X
Dominique A Cadilhac http://orcid.org/0000-0001-8162-682X

REFERENCES

1 Brundage M, Blackford A, Tolbert E, *et al*. Presenting comparative study pro results to clinicians and researchers: beyond the eye of the beholder. *Qual Life Res* 2018;27:75–90.
2 Calvert M, Kyte D, Price G, *et al*. Maximising the impact of patient reported outcome assessment for patients and society. *BMJ* 2019;364:k5267.
3 Van Der Wees PJ, Nijhuis-Van Der Sanden MWG, Ayanian JZ, *et al*. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *Milbank Q* 2014;92:754–75.
4 Valderas JM, Kotzeva A, Espallargues M, *et al*. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. *Qual Life Res* 2008;17:179–93.
5 Bantug ET, Coles T, Smith KC, *et al*. Graphical displays of patient-reported outcomes (pro) for use in clinical practice: what makes a pro picture worth a thousand words? *Patient Educ Couns* 2016;99:483–90.
6 Klose K, Kreimeier S, Tangermann U, *et al*. Patient- and person-reports on healthcare: preferences, outcomes, experiences, and satisfaction - an essay. *Health Econ Rev* 2016;6:18.
7 Nilsson E, Orwelius L, Kristenson M. Patient-reported outcomes in the swedish national quality registers. *J Intern Med* 2016;279:141–53.
8 Greenhalgh J, Dalkin S, Gibbons E, *et al*. How do aggregated patient-reported outcome measures data stimulate health care improvement? a realist synthesis. *J Health Serv Res Policy* 2018;23:57–65.
9 Snyder C, Smith K, Holzner B, *et al*. Making a picture worth a thousand numbers: recommendations for graphically displaying patient-reported outcomes data. *Qual Life Res* 2019;28:345–56.
10 Hancock SL, Ryan OF, Marion V, *et al*. Feedback of patient-reported outcomes to healthcare professionals for comparing health service performance: a scoping review. *BMJ Open* 2020;10:e038190.
11 Cadilhac DA, Grimley R, Kilkenny MF, *et al*. Multicenter, prospective, controlled, before-and-after, quality improvement study (Stroke123) of acute stroke care. *Stroke* 2019;50:1525–30.
12 Desomer A, Van Den Heede K, Triemstra M. Use of patient-reported outcome and experience measures in patient care and policy Belgian Health Care Knowledge Centre; 2018: 1–151.
13 Boyce MB, Browne JP. Does providing feedback on patient-reported outcomes to healthcare professionals result in better outcomes for patients? a systematic review. *Qual Life Res* 2013;22:2265–78.
14 Sanders EB-N, Stappers PJ. Co-Creation and the new landscapes of design. *CoDesign* 2008;4:5–18.
15 Hsu C-C, Sandford BA. The Delphi technique: making sense of consensus. *Practical Assessment, Research & Evaluation* 2007;12:1–8.
16 Ernstsson O, Janssen MF, Heintz E. Collection and use of EQ-5D for follow-up, decision-making, and quality improvement in health care - the case of the swedish national quality registries. *J Patient Rep Outcomes* 2020;4:78.
17 Australian Commission on Safety and Quality in Healthcare. Economic evaluation of clinical quality registries appendix A- case study summaries, 2016. Available: https://www.safetyandquality.gov.au/sites/default/files/migrated/Economic-evaluation-of-clinical-qualityregistries-Final-report-Nov-2016.pdf [Accessed 07 Feb 2022].
18 Department of Health. Addendum to the National health reform agreement (NHRA) 2020–25, 2020. Available: https://www.health.gov.au/initiatives-and-programs/2020-25-national-health-reform-agreement-nhra [Accessed 07 Feb 2022].
19 Cadilhac DA, Lannin NA, Anderson CS, *et al*. Protocol and pilot data for establishing the Australian stroke clinical registry. *Int J Stroke* 2010;5:217–26.
20 Zoom Video Communication Inc. Zoom meetings, 2019. Available: www.zoom.us/meetings [Accessed 01 Aug 2019].
21 van Swieten JC, Koudstaal PJ, Visser MC, *et al*. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 1988;19:604–7.
22 Brooks R. EuroQol: the current state of play. *Health Policy* 1996;37:53–72.
23 Snaith RP. The hospital anxiety and depression scale. *Health Qual Life Outcomes* 2003;1:29.
24 Breen S, Cadilhac D, Lannin N. The Australian stroke clinical Registry annual report 2018. The. Heidelberg, Victoria, Australia Florey Institute of Neuroscience and Mental Health; 2019.
25 SurveyMonkey Inc. San Mateo, California: USA, 2019. Available: www.surveymonkey.com [Accessed 01 Aug 2019].
26 Brundage MD, Smith KC, Little EA, *et al*. Communicating patient-reported outcome scores using graphic formats: results from a mixed-methods evaluation. *Qual Life Res* 2015;24:2457–72.
27 Microsoft Corporation. Microsoft Excel, 2016
28 Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3:77–101.
29 Ishaque S, Karnon J, Chen G, *et al*. A systematic review of randomised controlled trials evaluating the use of patient-reported outcome measures (PROMs). *Qual Life Res* 2019;28:567–92.
30 Kotronoulas G, Kearney N, Maguire R, *et al*. What is the value of the routine use of patient-reported outcome measures toward improvement of patient outcomes, processes of care, and health service outcomes in cancer care? a systematic review of controlled trials. *J Clin Oncol* 2014;32:1480–501.
31 Snyder C, Brundage M, Rivera YM, *et al*. A PRO-cision medicine methods toolkit to address the challenges of personalizing cancer care using patient-reported outcomes: introduction to the supplement. *Med Care* 2019;57 Suppl 5 Suppl 1:S1–7.

32 Snyder CF, Smith KC, Bantug ET, *et al*. What do these scores mean? presenting patient-reported outcomes data to patients and clinicians to improve interpretability. *Cancer* 2017;123:1848–59.
33 Snyder CF, Aaronson NK, Choucair AK, *et al*. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Qual Life Res* 2012;21:1305–14.
34 Allwood D, Hildon Z, Black N. Clinicians' views of formats of performance comparisons. *J Eval Clin Pract* 2013;19:86–93.
35 National Cancer Institute. Making data talk: A workbook. In: *U.S. department of health and human services*. USA: National Institute of Health, 2011.
36 Kuijpers W, Giesinger JM, Zabernigg A, *et al*. Patients' and health professionals' understanding of and preferences for graphical presentation styles for individual-level EORTC QLQ-C30 scores. *Qual Life Res* 2016;25:595–604.
37 Aiyegbusi OL, Kyte D, Cockwell P, *et al*. Patient and clinician perspectives on electronic patient-reported outcome measures in the management of advanced CKD: a qualitative study. *Am J Kidney Dis* 2019;74:167–78.
38 Oemrawsingh A, van Leeuwen N, Venema E, *et al*. Value-based healthcare in ischemic stroke care: case-mix adjustment models for clinical and patient-reported outcomes. *BMC Med Res Methodol* 2019;19:229.
39 Smith KC, Brundage MD, Tolbert E, *et al*. Engaging stakeholders to improve presentation of patient-reported outcomes data in clinical practice. *Support Care Cancer* 2016;24:4149–57.