Developing improvement measures using driver diagrams and a modified Delphi technique

ABSTRACT

Purpose A new method to engage salient organisational stakeholders in identifying and ranking measures of healthcare improvement programs is described. The method is illustrated using Executive WalkRounds (EWRs) in a multi-site Australian Health District.

Methodology Subject matter experts (SMEs) conducted document analysis, identified potential EWRs measures, created driver diagrams and then eliminated weak measures. Then a panel of executives skilled in EWRs ranked and ratified the potential measures using a modified Delphi technique.

Findings Document analysis revealed three main EWRs aims: two-way communication and empowered staff, patient and staff satisfaction and positive and open workplace cultures. Ten of 28 potential measures were eliminated by the SME review. After repeated Delphi rounds the executive panel achieved consensus (75 per cent cut-off) on seven measures. One outcome, one process and one implementation fidelity metric were selected from these seven measures to measure and monitor the impact of EWRs in the health district.

Practical implications Perceptions of weak relationships between measures and intended improvements can lead to practitioner scepticism. This work offers a structured method to combine the technical expertise of SMEs with the practical knowledge of healthcare staff in selecting improvement measures.

Originality/value This research is the first to combine driver diagrams with a modified Delphi technique to leverage formal and practical types of expertise in selecting measures that are strongly linked to local quality improvement goals.

Keywords: Methodology, Performance measurement, Mixed methods, Quality improvement, Rounding, Driver Diagrams, Delphi Technique, Healthcare Executives
Introduction

Measurement is critical to improvement (Shewhart and Deming, 1939; Teare, 2014). In healthcare settings measurement issues include inconsistent collection, inability to effectively monitor performance and unclear alignment with strategy (Mesabbah and Arisha, 2016). Furthermore, too many or irrelevant measures waste effort and reduce measurement efficacy (Weeks et al., 2008; Teare, 2014), leading to calls to reduce the number of healthcare performance measures in some countries by 75 per cent (Berwick, 2018). Hospital quality improvement project evaluations (Thor et al., 2010) and national quality indicator reviews (Riain et al., 2015) confirm the importance of judicious measurement selection.

This work supports matching metrics to context-specific improvement goals (Njuangang et al., 2017; Carlin and Chesters, 2019), and provides a method to accomplish this objective. The structured method developed combines driver diagrams with a modified Delphi technique to support engagement of key organisational stakeholders. These stakeholders include technical experts in healthcare quality improvement and the staff implementing the initiative. The method can be applied to a variety of health service initiatives to clarify, ratify and select appropriate improvement metrics. The specific EWRs measures selected are likely to interest readers with questions about measuring rounding outcomes.

This introduction has two sections. First, the EWRs literature is reviewed with a focus on measurement and methods. Second, the need for a new method for measuring improvement is explained and justified. The method is illustrated using EWRs in the context of South Western Sydney Local Health District’s (SWSLHD’s) transformational program.

Executive WalkRounds

EWRs (also known as Senior Leader Rounding, Leadership WalkRounds or Patient Safety Leadership WalkRounds) refer to healthcare executives routinely visiting places where care is delivered and having informal, semi-structured conversations with patients and staff. Reported aims for EWRs include: hearing concerns and developing relationships with frontline staff, (Frankel et al., 2005; Frankel et al., 2008; Studer, 2008; Morello et al., 2013; Singer and Tucker, 2014) promoting patient safety and patient safety culture, (Frankel et al., 2005; Frankel et al., 2008; Adelman, 2012; Singer and Tucker, 2014; Sexton et al., 2014) promoting accountability, (Frankel et al., 2005) demonstrating organisational commitment,
generating and implementing improvement ideas (Adelman, 2012; Sølvtofte et al., 2017) and allocating resources wisely (Frankel et al., 2005).

There is some consensus that EWRs are beneficial, following the model of Management by Walking Around and the Lean Systems Thinking ethos of ‘Going to Gemba’ (Singer and Tucker, 2014). However, studies to date show uncertain returns on the commitment of executive time. Research assessing EWRs outcomes in healthcare settings variously report improvement (Thomas et al., 2005; Sexton et al., 2014; Sexton et al., 2017), no impact (Singer and Tucker, 2014; Winter and Tjong, 2015) or declines (Rotteau et al., 2014; Tucker and Singer, 2015) in safety climate, staff engagement or staff well being measures.

Generally, as the rigour of EWRs evaluation studies has increased over time, results have become more ambiguous. Poorly defined relationships between outcome measures and what EWRs are expected to achieve in a particular care context can lead to scepticism about the value of executives spending time at the point of care (Beil-Hildebrand, 2006; Tucker and Singer, 2015). Methodological issues in the literature, including heterogeneity among the studies’ measures and settings (Morello et al., 2013), limit the ability to make conclusions about the effectiveness of EWRs. What is certain is that EWRs require investments of executive and front line staff time which could be devoted to other activities. Furthermore, when poorly executed, EWRs can produce unintended, damaging effects on staff engagement (Martin et al., 2014; Rotteau et al., 2014).

*Measures of Executive WalkRounds for patient safety*

The literature specific to EWRs intended to improve patient safety reveals a lack of objective performance measures to verify improvement (Singer and Tucker, 2014). Many studies are unclear about the rationale why specific measures were selected. Few indicate if metrics were chosen in consultation with quality improvement personnel (Martin et al., 2014) or if executives were even aware of the measurement and analysis regime. Different studies report improvement, no impact, or declines in *safety climate measures*, (Singer and Tucker, 2014). Executives have been observed to diverge from the facility’s EWRs script to add measures that satisfied their own goals (Martin et al., 2014). This may reflect weak links between measures selected (e.g. safety climate) and what executives expect EWRs to achieve in their facility. Determining which measures accurately capture the changes intended from implementing EWRs in a particular context could help clarify EWRs’ impact.
Measures of Executive WalkRounds with aims other than patient safety

Few published EWRs studies measure goals except patient safety. These retrospectively report EWRs as co-existing with high performance (Adelman, 2012; Aboumatar et al., 2015) or use staff perceptions of performance as the primary metric (Tucker and Singer, 2015). Another (Sexton et al., 2017) measures employee engagement, burnout and work-life balance. One study (Martin et al., 2014) set a priori outcome measures that included improved relationships between “ward and board”. The paucity of rigorous evaluations of EWRs intended to address issues other than safety means validated measurement tools may not be available for context-specific aims.

Justification for developing a new method

The EWRs literature is dominated by measures of safety culture and safety climate, but these were inconsistent with SWSLHD’s intent. The impetus for this research was the practical need to select EWRs measures with strong links to SWSLHD’s EWRs aims. Hence, a new approach was required. This led to consideration of existing approaches, then development of a new method.

A new method for developing measures for improvement in healthcare

The SWSLHD’s Steering Committee had directed that a systematic approach be used to determining EWRs measures. Reliance on EWRs co-existing with high performance, as reported in early case studies, without specific EWRs measures would not meet the criteria for a systematic approach.

Extensive consultation had been conducted with staff in development of the “Transforming your Experience” (TYE) program and the objectives of each program element had been extensively documented. Furthermore, subject matter experts (SMEs) in quality measurement and in EWRs existed in the district’s staff, but had not worked together to clarify EWRs aims or measures. District executives who were skilled in EWRs represented another important source of knowledge. These documents, skills and resources could be systematically combined to develop EWRs measures.

Mixed methods in health services research

Health services organisations consist of individuals from a variety of professions located within shared national and societal contexts. The ability to effectively identify, isolate and
accurately measure both dependent and independent variables within such natural settings is limited, and supports a qualitative research approach (Lincoln and Guba, 1985). However, the need to provide a numeric scale to systematically rank potential EWRs measures required a combination of qualitative and quantitative tools (Abildgaard et al., 2016). Therefore a mixed methods approach was appropriate, incorporating qualitative tools such as document analysis and focus groups, and quantitative tools such as Delphi questionnaires.

The ability to combine SMEs’ technical knowledge (to ensure fidelity to the intent of EWRs) with the practical knowledge of SWSLHD executives skilled in conducting EWRs supported a mixed-method, multiple-step approach to selecting measures. Combining driver diagrams with the Delphi technique helps to address the shortcomings and biases typical of each tool if used in isolation, and support executives’ involvement in measurement selection through minimising the time required for effective participation.

Combining driver diagrams and Delphi technique to identify, rank and ratify measures

This work now explains how context-specific aims were identified and driver diagrams were combined with a modified Delphi technique to select measures explicitly linking SWSLHD’s EWRs aims with process, outcome and fidelity measures.

The method of combining document analysis, creating driver diagrams and applying a modified Delphi technique was specifically designed to address weaknesses associated with the Delphi technique (Hsu and Sandford, 2007) such as the time demanded of panel members and unequal levels of expertise and practical experience in the topic of interest. Creating driver diagrams engaged SMEs in assessing connections between SWSLHD’s EWRs goals and valid and reliable potential measures. This was necessary because measures reported in the EWRs literature included many not relevant to the district’s aims, and did not consider locally available measures, such as existing annual staff surveys, which could help measure and monitor progress.

The use of driver diagrams followed by a Delphi technique to identify quality improvement measures has not been previously reported in the literature, suggesting the method described in this paper will be of interest to quality improvement scholars and practitioners. While the example reported here involves EWRs, the method to identify, rank and ratify measures congruent with claims made for improvement initiatives in a specific setting can be applied to other interventions.
Method

Site

SWSLHD is located in the State of New South Wales and provides public healthcare to approximately one million Australians living within its 6,000 square kilometre (2,300 square mile) catchment. It encompasses six hospitals in addition to outpatient, community, drug, oral, mental health and corporate services. EWRs are one of the interventions that comprise the TYE initiative to improve the quality and safety of care, patient and staff experiences and leadership at SWSLHD.

Personnel

SWSLHD SMEs participated in selection of the EWRs measures as members of the EWRs Measurement and Monitoring Workgroup. The workgroup consisted of the TYE Director, District Quality Manager and two EWRs SMEs. It was convened to improve EWRs measurement and data collection plans. The workgroup was facilitated by a Lean Six Sigma Black Belt and an academic.

SWSLHD executives with practical experience of EWRs in the district were recruited to the Delphi panel. Inclusion criteria for the EWRs Delphi panel were membership of a service or facility’s executive leadership team, completion of EWRs training, reliably rounding and being assessed by SMEs as effectively conducting EWRs. Purposive sampling was used to identify all executives in SWSLHD currently conducting EWRs, and 28 judged by the SMEs to have a high degree of EWRs skill were invited to participate in the multi-disciplinary expert panel. Potential panel members had all received training regarding: EWRs objectives and routines, the electronic system used to record the EWRs results and ensuring action resulted from identified issues. The time required, and the need to commit to completing both rounds two and three of the Delphi process, were explained and emphasised prior to executives being accepted on the panel. The recruitment objective was to achieve a 12-15 member panel. The executive panel was purposefully multi-disciplinary and included clinical and non-clinical executives. Panel members included general managers and directors representing medical services, nursing and midwifery and corporate functions from the district’s hospitals and community health units.
**Overview of method**

Figure 1 shows an overview of the three steps used to combine the technical expertise of SMEs with the practical knowledge of executives successfully conducting EWRs in the district.

---

**Figure 1**: Overview of three step method to develop improvement measures

Each step is detailed in the following sections.

**Step 1: Conduct document analysis to identify SWSLHD Executive WalkRounds aims and subject matter expert focus group to identify potential measures**

The EWRs Measurement and Monitoring Workgroup first reviewed existing SWSLHD documentation to clarify and confirm the EWRs aims of the health district. SWSLHD documents, including the Seven Safety Essentials Toolkit, Transforming Your Experience Implementation Plan 2017-2021, Transforming Your Experience: Our strategy to transform South Western Sydney Local Health District 2017-2021, and the Transforming Your Experience: Monitoring & Evaluation Plan (South Western Sydney Health District, 2018) were analysed. Sources of potential quantitative and qualitative measures of staff and consumer perceptions, including the annual People Matter survey (NSW Government Public Service Commission, 2017) and the Adult Admitted Inpatient Survey (NSW Bureau of Health Information, 2017) were discussed and debated by the workgroup members. These surveys are routinely administered and the survey questions have been validated. The list of potential measures from the People Matter and Adult Admitted Inpatient surveys was expanded with other measures the SMEs judged relevant to SWSLHD EWRs aims. Efforts were made to identify existing measures with routine data collection schedules already in place. These included quantitative performance measures from NSW Health Service Performance Agreements (Health System Information & Performance Reporting Branch, 2017).
Step 2: Develop driver diagrams linking documented SWSLHD Executive WalkRounds aims with potential measures

After determining SWSLHD’s aims for EWRs, the EWRs Measurement and Monitoring Workgroup developed driver diagrams. Driver diagrams visually show logical links between an improvement aim and the general and specific actions (primary and secondary “drivers” of change) that support achievement of that aim. Driver diagrams can include prioritised change ideas and relevant measures of performance (Clinical Excellence Commission, 2017). Their value lies in translating broad improvement objectives into specific measurements that are congruent with program aims. Step-by-step guides to developing driver diagrams are readily available (Bennett and Provost, 2015; Clinical Excellence Commission, 2017; Institute for Healthcare Improvement, 2017).

Identifying outcome, process and fidelity measures in a driver diagram is beneficial because each type of measure provides different kinds of information and complements the other two measurement types. Outcome measures indicate progress towards a target, process measures provide diagnostic capabilities should outcomes vary from expectations (Mant, 2001) and fidelity measures support ongoing verification that the intervention is being implemented as intended (Pérez et al., 2016).

The excerpt from an EWRs Driver Diagram shown in Figure 2 illustrates how one high-level aim, that EWRs produce two way communication and empowered staff, is linked to potential measures. The diagram shows one of the primary EWRs drivers for change at SWSLHD is to conduct a standard, not ad-hoc, discussion. The middle portions of the diagram show the secondary drivers or “how” the aim is expected to be achieved (in this case the minimum questions for discussion). Adhering to a set schedule for rounding would be another secondary driver. Potential measures appear on the right hand side (e.g. number of opportunities for improvement identified through EWRs). Thus the driver diagram excerpt shows logical links between potential measures to monitor the intended EWRs impact in the health district.

Insert Figure 2 around here
Step 3: Perform modified Delphi technique

The Delphi technique is a structured group communication process combining questionnaires to elicit anonymous responses with controlled feedback to arrive at a collective consensus (Hasson et al., 2000). It has been used for over 50 years and combines the individual contributions of experts to produce group perspectives not otherwise attainable (Hasson et al., 2000; Stitt-Gohdes and Crews, 2004). The Delphi technique allows participation based on both professional experience and expert knowledge (Njuangang et al., 2017; Carlin and Chesters, 2019). However, no single criterion is used to identify consensus and different researchers have applied divergent data collection and analysis tools (for example, ni Riain et al., 2015; Ab Latif et al., 2016; Njuangang et al., 2017; Carlin and Chesters, 2019).

Nevertheless Delphi methods have proved useful in healthcare settings, for example, to produce evidence and consensus-based indicators for cancer (Gagliardi et al., 2005) and to determine system-level indicators (Hassan and Barnett, 2002; ni Riain et al., 2015). Detailed guides to the variations in applying Delphi techniques and analysing results in healthcare settings are available (Hassan and Barnett, 2002; Pereira and Alvim, 2015; Ab Latif et al., 2016; Njuangang et al., 2017).

A modified Delphi approach was chosen to review and ratify EWRs measures because the problem does not lend itself to precise analytical techniques but can benefit from subjective judgments on a collective basis. Additionally, more individuals were needed than could effectively interact in a face-to-face exchange, and the panel members were geographically dispersed. Furthermore, the heterogeneity of the participants’ responses needed to be preserved by avoiding domination from professional hierarchies or force of personality (Hasson et al., 2000; Stitt-Gohdes and Crews, 2004). The modified Delphi technique used provides an iterative mixed methods approach (Carlin and Chesters, 2019) and is shown graphically in Figure 3, followed by a detailed description of each Delphi round.

---

**Figure II.** Excerpt from driver diagram showing potential Executive WalkRounds measures

---

**Figure III.** Schematic of modified Delphi Method used to review and ratify measures
**Step 3, Delphi round 1:**

*Subject matter expert review to eliminate weak measures from the EWRs driver diagram*

Delphi round one consisted of the EWRs Measurement and Monitoring Workgroup reviewing and eliminating weak potential measures from consideration. The criteria used for rejection or retention of a measure were a subset of previously identified desirable attributes for measurement (Lohr, 2002). Specifically, measures judged to have a strong, logical link to the concept of interest, which were reliable, valid, responsive to change and easy to collect received the highest ranking. The SMEs independently rated 28 potential measures using a three point scale (0 = reject from consideration, 1 = potentially useful as a secondary, diagnostic measure, 2 = strong measure) to create the list of potential measures to be assessed by the executive panel in the following step. The potential measures that were retained were presented to the executive panel in a questionnaire in Delphi round two.

**Step 3, Delphi round 2:**

*Executive panel’s first review of potential measures*

During a face-to-face meeting with a SME the executives invited to be part of the Delphi panel were familiarised with their task by receiving a summary of the Delphi technique, the objective of the panel and how the Delphi review would be conducted. Round two data collection was completed in the same meeting. Delphi panel members rated each potential measure on a seven point Likert scale: 1=unacceptable measure, 2=very poor measure, 3=poor measure, 4=acceptable measure, 5=good measure, 6=very good measure, 7=excellent measure, 0=don’t know.

**Step 3, Delphi round 3:**

*Executive panel’s second review of potential measures*

The round three questionnaire was emailed to all panel members with a date for the questionnaire to be completed. In addition to the questions used in round two, the round three questionnaire showed each panel member’s individual round two responses, the combined executive panel’s average ratings in round two, comments written by panel members in round
two and if any additional measures had been nominated for consideration. Panel members used the same seven-point Likert scale to identify measures they perceived as most important and relevant for EWRs. As a last step the Measurement and Monitoring Workgroup reviewed the Delphi panel’s ratings and the degree of consensus about each potential measure to identify three EWRs measures for the health district.

Data analysis method

Results were analysed using consensus, median scores and interquartile ranges (Hasson et al., 2000; von der Gracht, 2012). Agreement levels were set prior to data collection as: strong consensus = > 75 per cent, Moderate consensus 60-74 per cent and absence of consensus less than 60 per cent (Pereira and Alvim, 2015).

Results

Time requirements for each step of the method and data analysis are first reported. Then the EWRs measures selected by SWSLHD are tabulated showing data from each Delphi round and levels of panel consensus.

The three steps (identify EWRs aims and potential measures, develop driver diagrams and conduct modified Delphi technique) were completed between September 17, 2017 and March 8, 2018, an elapsed time of 24 weeks.

Time requirements for each step

Step 1. Twelve hours of document analysis and 20 hours of focus groups and meetings were required to determine SWSLHD’s specific aims and list potential measures. Hence, step one required 32 hours of SME time in total.

Step 2. Developing driver diagrams required a total of ten hours from the SMEs.

Step 3 round 1. Ranking potential measures and eliminating those not meeting the criteria set for SWSLHD measures required a total of 20 SME hours.

Step 3 rounds 2 and 3. Recruitment of panel members, familiarising panel members with the Delphi technique, running two Delphi rounds, and analysing data required:
Recruitment and scheduling meetings, five hours of administrative support

Familiarising panel members with the Delphi technique took five minutes for each of the 17 panellists (1 hour and 25 minutes of executives’ time in total) and three hours of a SME’s time.

Data collection: Four panellists did not complete the third Delphi round. Therefore, five hours of executive time was required for data collection, (10 minutes x 17 panellists for the first iteration and 10 minutes x 13 panellists answering the second iteration of the questionnaire). Four hours of administrative time was required to remind executives to complete the emailed round three Delphi questionnaire.

Data analysis and tabulation of the results required seven hours (four hours from SMEs and three hours administrative support).

The total time for all personnel involved was 87 hours and 25 minutes. SMEs contributed 79 per cent (69 hours), administration 14 per cent (12 hours) and executives 7 per cent (6 hours and 25 minutes) of the time required to conduct the method. Forty-six percent of the SMEs’ time was devoted to determining the district’s EWRs aims.

*Step 1. Executive WalkRounds Results*

Triangulating data from document analyses with SWSLHD EWRs and quality improvement SMEs produced the following EWRs aims: two-way communication and empowered staff, positive and open workplace cultures and improved patient and staff satisfaction. The SMEs added three implementation fidelity aims: adherence to rounding routine, rounding coverage across all locations and repeat visits in areas that required more support. The SWSLHD EWRs aims derived from document analysis did not have explicit links to safety, in contrast to the predominance of safety oriented EWRs reported in the scholarly literature.

*Step 2. Executive WalkRounds Results*

Development of Executive WalkRounds Driver Diagrams addressing the documented aims produced a list of 28 potential measures for review and ranking

*Step 3, Delphi round 1. Executive WalkRounds Results*

Twenty eight unique potential measures were reduced to 18 during the SMEs’ review of the of driver diagrams.
Step 3, Delphi rounds 2 and 3. Executive WalkRounds Results

Seventeen SWSLHD executives completed the second Delphi round, and thirteen completed both the second and third rounds providing an acceptable panel size (Ab Latif et al., 2016). The retention rate between the two executive panel rounds was 76 per cent. Three executives used the rating of 0 (don’t know) a total of four times; these four ratings were excluded from calculations of central tendency and consensus. The full range of ratings in the Likert scale were used by the individual executives, from 1 (unacceptable measure) to 7 (excellent measure).

Executives wrote comments supporting, describing weaknesses or identifying needs for clarification for approximately a fifth of the potential measures, for example:

“Difficult areas” would need to be defined.

and

Sometimes this is not feasible [in relation to monitoring the number of improvements implemented each month].

The potential employee survey measure ‘I believe senior managers listen to employees’ elicited comments such as:

Need to classify who people identify as senior managers.

and

Excellent measure of openness, but will depend on individual.

Executives also commented on the proposed frequency of data collection and other logistical matters. No new measures were proposed by the executive panel.

Data Analysis. Consensus, median scores and interquartile range

Panel members displayed strong consensus on seven potential measures and moderate consensus on four measures. All measures that reached the strong consensus cut-off had interquartile ranges (IQR) of one or less and median scores of five or higher indicating assessments of ‘good’, ‘very good’ or ‘excellent’. None of the 18 measures presented to the executive panel had median scores less than four, indicating all were viewed as acceptable, or better, measures. Median scores, Delphi round three measures of consensus and interquartile ranges for the potential EWRs measures are shown in Table 1.
Three measures were chosen to assess the performance of EWRs against SWSLHD’s aims for the program. The outcome measure selected was the response to the People Matter survey question “I would recommend my organisation as a great place to work.” The fidelity of implementation measure selected was the proportion of each facility’s departments involved in EWRs in the preceding month. The EWRs process measure selected was the number of staff rounded with, reported for one month. These three measures all had interquartile ranges of 1.0 combined with 69 per cent or more of the panel rating the measure as “good” or better on the Likert scale, indicating an acceptable level of consensus (von der Gracht, 2012).

Discussion

Combining driver diagrams with a modified Delphi technique to develop and ratify measures that explicitly linked health district EWRs goals with process, outcome and fidelity measures succeeded in combining SMEs and executives’ insights. In addition to benefiting from the executive panel’s experience and practical knowledge, the method provided an opportunity to reinforce the health district’s EWRs objectives with executives.

While the method met the objective of demanding little time from executives, the elapsed time was longer than expected. This was largely due to data collection spanning the summer holiday period. Data collection could be completed more quickly in other months of the year. Furthermore, most of the SME hours were dedicated to identifying SWSLHD’s EWRs aims. In organisations with clearly defined improvement aims, step one could be eliminated from the method.

Ranking and ratifying EWRs measures provides opportunities to engage executive and clinical groups in the process of metric selection and accommodate the diversity of professional and occupational cultures that characterise health services (Nembhard et al., 2012). The number of comments written on the Delphi rounds rating spreadsheet signified good executive engagement with the process and that the metrics selected to measure EWRs were of interest to them.
Furthermore the comparatively low rankings given to some measures that had passed the SME review suggest the executives, some of whom had clinical as well as managerial responsibilities, added value to the selection of EWRs measures. Acceptance of the method by the executive panel was good. It has since been used to identify the measures for other patient-focussed elements of the transformation program, such as safety huddles and clinical handover.

Given the importance of engaging healthcare professionals in quality and performance improvement, a method that involves and leverages their expertise in selecting appropriate measures with minimal demands on their time is important. In the EWRs example presented here executives required 25 minutes each to play an informed and active role in measurement selection. Furthermore, the selected EWRs measures had strong links to the concept of interest, were easy to collect and were regarded by SMEs as valid and reliable measures. The measures selected using the combination of driver diagrams and Delphi technique matched the aims of EWRs across SWSLHD, and will support more robust EWRs evaluations than the descriptive case study methods that dominate EWRs appraisals to date.

**Limitations**

All measurement techniques are subject to researcher and subject bias, for example in the questions asked and the potential for respondents to select responses that will reflect favourably upon them. The sequential mixed method design mitigates these biases by recruiting multidisciplinary Delphi panel members from multiple sites and balancing formal and experiential sources of knowledge. The semi-anonymous feedback between Delphi panel rounds also reduces bias due to profession and personality-based influence.

It must be acknowledged that relatively low individual rankings given to some measures that had passed the SMEs review may represent individual executive’s attempts to avoid onerous EWRs measures. To manage bias panel members were encouraged to write comments about measures in round two which were then read by the rest of the panel in round three.

The measures selected are specific to the context of SWSLHD and its aims for EWRs in the first year of the transformational change program, and may not be applicable to other locations, or even the same location, at a future date. Furthermore, it should be noted that patient and staff stories of benefits from EWRs provide complementary sources of
information about EWRs, and are routinely elicited as part of the SWSLHD data collection plan.

Future Research

EWRs are often promoted as a standardised intervention leading to healthcare improvement, but the EWRs literature generally does not include reasons for measure selection. Eliciting formal EWRs knowledge from SMEs and organising it through driver diagrams shows logical links between local improvement aims, general and specific drivers and measures. The SME review produced a 36 per cent reduction in the number of potential measures included in the Delphi rounds. Developing driver diagrams before engaging the practical knowledge of healthcare executives in a modified Delphi technique resulted in no other measures being suggested.

This leads to propositions that SMEs completing a driver diagram may improve the measures presented in the Delphi questionnaire, and in doing so reduce time to complete the Delphi rounds and improve retention between rounds. These propositions need testing through quasi-experimental research designs in other healthcare settings and with other quality improvement interventions. This could include recruiting two executive panels from the same organisation, providing one panel with a Delphi questionnaire developed without, and the other questionnaire with driver diagram input, and then comparing the measures selected, time required, and ratings of the measures produced by the two panels. Research regarding measures selected as congruent with EWRs aims in varied healthcare settings would also make valuable additions to knowledge.

Within SWSLHD, the three selected EWRs measures support future research into the efficacy of EWRs across the district. They will also permit robust EWRs comparisons between facilities, inpatient and outpatient care settings and between clinical and corporate settings.

Conclusion

Understanding the aims for an improvement initiative in a particular context, identifying potential process, outcome and fidelity measures relevant to those aims, then reviewing and ratifying the measures with health services personnel making the improvement supports
robust evaluations of the initiative. The method described in this manuscript engages SMEs, executives and clinicians in clarifying, ratifying and selecting appropriate improvement metrics and can be employed in a variety of settings.
REFERENCES


Figure 1: Overview of three step method to develop improvement measures

STEP 1: IDENTIFY AIMS AND POTENTIAL MEASURES

STEP 2: DEVELOP DRIVER DIAGRAMS LINKING AIMS TO POTENTIAL MEASURES

STEP 3: CONDUCT THREE ROUND MODIFIED DELPHI TECHNIQUE
**Figure II:** Excerpt from driver diagram showing potential Executive WalkRounds measures

<table>
<thead>
<tr>
<th>Transformational Aim</th>
<th>Primary Drivers</th>
<th>Secondary Drivers</th>
<th>Potential Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two way communication, empowered staff</td>
<td>Standardised discussion (NOT ad-hoc)</td>
<td>Follow standard template</td>
<td>Opportunities identified during rounding</td>
</tr>
</tbody>
</table>

Follow standard template: What is working well? What can we do better? What tools, information or equipment do you need to do your job?

Oppotunities identified during rounding:

- Number of opportunities identified by staff in last month
- Number of opportunities identified by patient/carer/family in last month
- Solutions undertaken by frontline staff
- Number of solutions implemented by staff (proposer or team) per month
- Number of solutions implemented by department manager per month
- Number referred to fix team (via meeting) in last month, trending for 90 days
- To provide the opportunity for communication about key organisational goals and direction
- Yes/No recorded against rounding
- NSW PM:Q6a “I believe senior managers provide clear direction for the future of the organisation”
- Staff Survey (MEM every 90d)
- NSW PM

**Note:** In the above figure “NSW PM” refers to the New South Wales Government People Matter survey and “opportunities” refers to opportunities for improvement identified
**Figure III**: Schematic of Step Three: modified Delphi Method used to review and ratify measures

In Executive Delphi Rounds 2 and 3 anonymous, structured communication was used to identify potential measures with high ratings of relevance and consensus.
Table I: Multidisciplinary Executive Delphi Panel ratings and consensus thresholds for Executive WalkRounds measures

<table>
<thead>
<tr>
<th>EWR Aims</th>
<th>Potential Measures (Abbreviated Description)</th>
<th>Median Round 2</th>
<th>Median Round 3</th>
<th>Consensus After Round 3</th>
<th>IQR After Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>More positive and open workplace</td>
<td>My organisation focuses on improving the work we do</td>
<td>5</td>
<td>5</td>
<td>69%</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td>My organisation is committed to developing its employees</td>
<td>6</td>
<td>6</td>
<td>69%</td>
<td>2.00</td>
</tr>
<tr>
<td>To improve patient and staff satisfaction</td>
<td>I would recommend my organisation as a great place to work</td>
<td>6</td>
<td>6</td>
<td>100%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Overall, how would you rate the care you received while in hospital?</td>
<td>6</td>
<td>6</td>
<td>83%</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>If asked about your hospital experience by friends and family how would you respond?</td>
<td>6</td>
<td>6</td>
<td>92%</td>
<td>0.25</td>
</tr>
<tr>
<td>Adherence to rounding routine</td>
<td>Number of patients/careers rounded with last week, reported for one month</td>
<td>5</td>
<td>5</td>
<td>38%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Number of staff rounded with last week, reported for one month</td>
<td>5</td>
<td>5</td>
<td>69%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Proportion of available executive staff undertaking rounding in last 30 (or 60) days</td>
<td>5</td>
<td>5</td>
<td>54%</td>
<td>1.00</td>
</tr>
<tr>
<td>Executive Leader Rounding</td>
<td>Number of acknowledgments for CORE values or 4 Focus areas in last week (reported as trend)</td>
<td>4</td>
<td>5</td>
<td>54%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Number opportunities identified by staff in last month</td>
<td>5</td>
<td>4</td>
<td>31%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Number opportunities identified by patients/careers/family in last month</td>
<td>4</td>
<td>4</td>
<td>31%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Number of solutions or improvements implemented per month by staff</td>
<td>4</td>
<td>4</td>
<td>46%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Number of solutions or improvements implemented per month by department manager</td>
<td>5</td>
<td>4</td>
<td>46%</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Number of issues or opportunities referred to Exec team (via meeting) in last month, trending for 90 days</td>
<td>5</td>
<td>5</td>
<td>77%</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>I believe senior managers provide clear direction for the future of the organisation</td>
<td>5</td>
<td>5</td>
<td>92%</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>I feel that senior managers listen to employees</td>
<td>5</td>
<td>5</td>
<td>92%</td>
<td>1.00</td>
</tr>
<tr>
<td>Coverage of entire facility</td>
<td>Proportion of hospital’s departments rounded with in last month</td>
<td>6</td>
<td>5</td>
<td>85%</td>
<td>1.00</td>
</tr>
<tr>
<td>Repeat visits in “difficult” areas</td>
<td>Number of repeat visits (rounds) to same wards in last 6 months</td>
<td>6</td>
<td>5</td>
<td>69%</td>
<td>2.00</td>
</tr>
</tbody>
</table>

**KEY:** % consensus: measures rated 5 "good", 6 "very good" or 7 "excellent" N= 13. Strong consensus cut-off set at 75%

Strong consensus = > 75% or IQR <=1  
Moderate consensus >60% & <75%  
Poor consensus <60%