Shaping quality: the use of performance polygons for multidimensional presentation and interpretation of qualitative performance data

T. M. Cook¹,²*, M. Coupe¹,² and T. Ku¹
¹ Department of Anaesthesia, Royal United Hospital, Combe Park, Bath BA1 3NG, UK
² Department of Anaesthesia, CircleBath Hospital, Bath, UK
* Corresponding author. E-mail: timcook007@googlemail.com

Editor’s key points
- Measuring quality of anaesthesia practice is a challenge.
- The authors have proposed ‘performance polygons’ as a tool to represent multidimensional assessments.
- Importantly, the polygons encourage a balanced approach to measure quality.
- This tool will be useful in presenting individual, team, or organizational performance quality in a meaningful comparative manner.

Background. Measuring outcomes and quality in anaesthesia is challenging. In the UK, there is increased focus on these as a result of changes in Department of Health strategy and the imminent introduction of mandatory revalidation for all doctors. A definition of quality may differ according to the observer’s standpoint and numerous performance measures may contribute to overall quality. Patients, surgeons, anaesthetic assistants, recovery nurses, managers, and anaesthetic peers are each likely to have their own perspective on ‘anaesthetic quality’ and would perhaps suggest different metrics to measure it. Speed, efficiency, cost, interpersonal skills, complication rates, patient recorded outcome measures, and satisfaction are all valid as quality measures, but none alone captures anaesthetic quality. Performance data are frequently presented as single-dimension measurements (e.g. pain, postoperative nausea and vomiting, patient satisfaction), but this does not address the fact that two or more domains may be closely related (e.g. use of regional anaesthesia and quality of analgesia) or in opposition (e.g. use of regional anaesthesia and speed).

Methods. We introduce the concept of a ‘performance polygon’ as a tool to represent multidimensional performance assessment. This method of data presentation encourages balanced appraisal of anaesthetic quality.

Results. Performance polygons may be used to compare individual performance with peers, published outcome norms, trends in performance over time, to explore aspects of team performance and potentially capture data that are required for medical revalidation.

Conclusions. Performance polygons enable easy comparison with any relevant data set and are a visual tool that potentially has wider applications in healthcare quality improvement.

Keywords: anaesthesia, audit; anaesthesia, general; performance indicators; quality indicators
Accepted for publication: 10 January 2012

Measurement of outcome in healthcare is central to determining quality and driving efforts to improve it. In the UK, the importance of outcomes measurement, measurement of quality, and quality improvement (QI) has been increasingly emphasized as it is now both a strategic focus of the Department of Health and a central component of plans for mandatory revalidation of all doctors.¹–³

Quality of performance in anaesthesia, and in healthcare in general, is often assessed by a single measure or from a single perspective, for example, a ‘postoperative pain audit’. However, healthcare quality is more complex and using a surgical-anaesthetic example, performance of tonsillectomy might be measured by time taken for anaesthesia or surgery, overall time in theatre, operative blood loss, postoperative nausea and vomiting (PONV), pain on waking, the successful day-case rate, cost of the hospital episode, number of readmissions for bleeding, number of re-operations for bleeding, duration of postoperative pain, or time until the patient is able to return to school or work. The focus of the surgeon (operative time, bleeding, readmission rate) may differ from the anaesthetist (nausea, pain, day-case rate), theatre manager (total theatre time, cost), hospital management (cost, day-case rate), and the patient (pain and nausea, readmission rate, time off school or work).

The term ‘silo mentality’ describes the tendency for individuals or departments within an organization to lack communication and understanding of common goals. In medicine, a silo mentality may lead different groups to tend to observe
and manage only those performance measures of interest to themselves and this may adversely affect other measures: a surgeon wishing to reduce readmissions may be slow, treat all patients as in-patients, and ban use of non-steroidal anti-inflammatory drugs: this will improve surgical ‘performance measures’ but at the cost of increased theatre time and hospital cost, poorer patient experience, poorer anaesthetic quality measures, and perhaps delayed return to work. In anaesthetic practice, a desire for rapid turnover may discourage the use of regional anaesthetic techniques, leading to an increase in postoperative pain and PONV. Inappropriately focusing on a single outcome may have unintended effects. For example, a postoperative pain audit may result in consequences such as an increase in opioid use, increased rates of PONV, and longer recovery stays if individuals alter performance, either unconsciously or consciously, in an attempt to ‘play the system’. Many outcomes in anaesthesia cannot be judged in isolation but may be fundamentally interconnected and frequently in tension with one another.

With the imminent implementation of medical revalidation in the UK, a focus on quality, QI, and reporting outcomes will become increasingly important for all doctors. For anaesthetists, it is recognized that identifying appropriate outcome measures can be challenging and that different observers may judge anaesthetists by different criteria. This article introduces ‘performance polygons’ as a form of data representation that reflects the complexity of multiple outcome measures in a way that other visual forms of data representation do not. We use theoretical examples to indicate a range of practical uses of performance polygons. This article describes a method of representing data but does not seek to define which measures should be used to define anaesthesia (or other) quality. This is the subject of a project currently being undertaken by the National Institute of Academic Anaesthesia and Royal College of Anaesthetists Working Group on Quality Measures in Anaesthesia.

**Methods**

**Performance polygons**

Performance polygons are a qualitative method of representing multidimensional data in a two-dimensional form. They are designed to enhance appreciation of quality across a number of domains. They are derived from star charts, first proposed by Georg von Mayr, a leading German statistician in the late nineteenth and early twentieth centuries. Star charts have variously been called radar charts, web charts, spider charts, cobweb charts, star plots, irregular polygons, polar charts, and kiviat diagrams.

In essence, the performance polygon is a series of bar graphs rotated and spatially connected by their origins, radiating outwards much like bicycle wheel spokes. The joined data points create an inner polygon which is shaded to facilitate visual appreciation of the data set, although as with standard bar graphs, this is not intended to imply that the scales are equivalent.

A performance polygon is constructed as described below.

- An outcome measure is plotted on a single line from ‘lowest performance’ to ‘best performance’ (Fig. 1A).
- Additional measures are added as further lines spreading outwards from the same origin. It is important that all measures are plotted so that outer limit represents best performance. The polygon is constructed so that the angle between each ‘spoke’ is equal (four measures 90°, five measures 72°, six measures 60°, etc.) (Fig. 1B showing five outcomes).
- The performance data are then plotted and the lines joined to form a ‘performance polygon’. The points indicating best performance are joined creating the outer limits of the polygon (Fig. 1C).
- Markers may be super-imposed as benchmarks or comparators: in this example, the local ‘best performer’ and

![Fig 1](image)

**Fig 1** Stepwise construction of a performance polygon: (a) an outcome is plotted from worst to best, (b) further outcomes are added all radiating from the same origin, (c) the maximal limits of the polygon are completed by joining points indicating ‘best outcomes’ and the index polygon is plotted by joining the performance measures on each outcome spoke (this is usually shaded but is left unshaded for clarity here), and (d) comparator polygons are added.
Comparator polygons might be superimposed to highlight, for example, temporal changes in an individual anaesthetist’s multidimensional performance or comparison with the performance of peers or against predefined benchmarks. Performance polygons are not limited to the representation of an individual’s work as, using a data set containing the necessary data, they have the potential to be used to represent the behaviour or performance of groups (e.g. a theatre team, a hospital, or a healthcare organization as a whole).

**Results**

Examples of theoretical performance polygons are presented.

**Individual anaesthetist’s performance with 95th and 5th centiles as comparators**

Figure 2 shows the performance of an individual anaesthetist. The outcome measures are exemplar indicators of performance that might be recorded in recovery: pain, PONV, speed of patient recovery, successful use of regional anaesthesia, turnaround time, and correct prescription of rescue analgesia. These include measures of interest to patients, surgeons, recovery staff, managers, and, of course, anaesthetists. The measures are chosen to include measures of anaesthetic skill (regional block success), process variables (adherence with good prescription practice), efficiency measures (turnaround time), and patient-relevant outcomes (pain, PONV). All measure an aspect of anaesthetic performance. The comparator polygons used here are the 5th and 95th centiles of a reference group, in this case, the whole anaesthetic department in that hospital.

This anaesthetist performs with some above average and some very good outcomes (compared with the reference group) but achieves this at the cost of being slow. Criticism from surgical colleagues about the slow speed of the service may be deflected by the high quality of patient-relevant outcomes. Personal reflection by the anaesthetist should lead to a focus on improved turnover with maintenance of other outcomes.

**Individual anaesthetist’s performance compared with own previous performance**

In Figure 3, the same outcome measures are plotted using the same anaesthetist’s historical performance (based on the previous year’s best and worst months) as the comparator.

A large personal polygon suggests improved outcomes over the time period compared with historical performance. However, outcome improvement is at the cost of turnover speed. This anaesthetist might, for example, have increased their use of ultrasound for regional anaesthetic techniques (taking longer as they are on a learning curve, but generating better outcomes). If this anaesthetist’s performance has been scrutinized within a ‘pain audit’, they would be considered to be performing well, but an ‘efficiency audit’ might raise concerns. The multidimensional data allow a sensible discussion of whether the change in practice is justified and worth pursuing. Re-plotting the polygon after a suitable interval will reveal whether the anaesthetist has maintained their good outcomes while also returning to previous turnover times.
**Figure 3** Performance polygon: individual anaesthetist’s performance compared with own previous performance.

**Figure 4** Performance polygon: individual anaesthetist’s performance with absolute or ‘benchmark’ performance as the comparator.

**Individual anaesthetist’s performance with absolute or ‘benchmark’ performance as comparators**

In Figure 4, the same outcome measures are plotted again, but in this polygon, the comparators are an upper benchmark of ‘success’ and a lower benchmark of ‘unacceptable performance’. These may be generated from departmental data, published outcome data, or consensus opinion in the host department. Using external ‘benchmarks’ overcomes the limitation of using colleagues’ performance as a comparator (the first two examples described here) if a department performs to a very high, or very low, standard such comparison may lead to unfair judgement about the absolute performance of an individual. A ‘benchmark’ polygon might have a role in identifying poor performing trainees or as part of revalidation for trained anaesthetists.

This anaesthetist generates very comfortable patients but who spend a long time in recovery and who have high rates of nausea compared with the benchmarks. On the positive side, the anaesthetist is quick. Perhaps this anaesthetist uses excessive amounts of opioids and minimal adjuncts or is not good at, or avoids, regional anaesthesia. A pain audit would likely rank this anaesthetist highly, but
recovery staff are unlikely to rate his/her performance as good.

**Surgical team performance**

Performance polygons clearly need not be restricted to anaesthetic practice. In Figure 5, the performance polygon is used to represent multidisciplinary multidimensional outcomes after knee arthroplasty. The EQ5D is the EuroQual 5D test that measures global well-being in five domains and is used in healthcare outcome studies.⁹ The domains measured are mobility, self-care, ability to perform usual activities, pain/discomfort, anxiety/depression, and an overall health assessment as a 0–100 visual analogue scale. In this performance polygon, the comparator polygons are 95th and 5th centiles and the median performance of a reference group. The reference group might be historical data, performance of the team in the neighbouring theatre, a neighbouring trust, or nationally acquired data from a high performing trust.

All outcome measures are arguably of interest to all members of the team, but individuals may influence some outcomes more than others: anaesthetist (theatre time, time to mobilize, and EQ5D score), surgeon (theatre time, complication rate, and Oxford knee score), and nursing and physiotherapy care (time to mobilize, EQ5D score, and length of stay). Managers will be interested in the time in theatre and length of stay but might more interested than the rest of the team in the cost of the total hospital episode. Finally, and perhaps most importantly, the patient will likely be most interested in EQ5D, length of stay, and Oxford knee score.

The quality of performance of this surgical team is high. The 3 month outcome is excellent and perhaps QI efforts should focus on the EQ5D and complication rates which are the outcomes closest to median performance. This may turn length of stay from good to excellent.

If other outcomes are considered relevant to the team, they may be added or substituted for others, thus creating a polygon with a different focus.

This performance polygon might also be used to compare surgical or anaesthetic practice. If there is debate (as currently) about the best surgical or anaesthetic/analgesic method to use for knee arthroplasty, a polygon might provide a more rounded assessment of the utility of different techniques than a traditional pain audit approach. Introduction of an enhanced recovery programme could lead to comparison of performance polygons before and after its introduction. Rather than simply focusing on one aspect, such as length of stay (as is currently often the case), a performance polygon might be used to see how such a change in practice has affected a number of measures of quality, either positively or negatively, and thus allow a balanced assessment of the change in practice to emerge.

**Real data**

One of the authors has collected outcome data in the early postoperative period for the last 2300 patients anaesthetized. For each patient, the data collected include:

- first-time success rate at i.v. cannulation,
- percentage of patients arriving in recovery who were awake before completing handover to recovery staff (a surrogate measure of speed of recovery),
- percentage of patients without pain in recovery,
- percentage of patients without nausea or vomiting in recovery,
- percentage of regional blocks performed when indicated,
- percentage success of performed regional blocks.
These outcome measures were chosen because they are available to measure in recovery and are consistent with the literature on outcomes patients and anaesthetists believe are important in defining performance quality.\textsuperscript{10–12} These data are presented in Figure 6A as absolute data with the individual axes each ranging from 0 to 100. Percentages indicate the anaesthetist’s performance. In (a), the authors have agreed benchmarks with limits representing poor performance (small comparator polygon) and high-quality performance (large comparator polygon). The change in the shape and area of the polygons in (A) and (B) illustrates the importance of choice of comparator data and limits.

**Discussion**

Performance polygons are a simple but powerful way to represent data over several domains; they provide a visual representation of data which is easily understood by observers. The use of comparator polygons can enhance their
value and transform the polygons from simple graphical displays to a potential driver of change and QI. This is intuitively less easy to achieve with uni-dimensional outcome data.

Performance polygons might have numerous roles in a department. Individual performance polygons, as illustrated above, might prove a useful starting point for an appraisal discussion on all round performance. Where concerns are raised about an individual practitioner (which often focus on one domain of performance: a surgeon considering the individual slow, a recovery nurse considering analgesia is poor, etc.), a performance polygon enables examination of performance ‘in the round’. If a database is large enough, performance polygons might be used to examine individuals’ performance for specific operations to determine perhaps who performs best (so they may educate others) or whether any individual is a lower outlier (so they may learn from others).

Performance polygons are able to present a ‘balanced’ view of procedural performance, traditional outcome measures, and the increasingly important patient-reported outcome measures: all important components of quality measurement. If developed robustly, performance polygons might play a role in medical revalidation. Recent articles have commented on the importance of routine perioperative data collection and on the importance and challenge of collecting multi-source feedback from patients. Of the 16 domains described in the General Medical Council’s revalidation ‘matrix’, it is entirely credible that an appropriate polygon might include data to support (i) colleague multi-source feedback, (ii) patient multi-source feedback, (iii) clinical outcomes data, and (iv) evidence of clinical audit and QI. If used as a continuous tool, capturing data from all anaesthetics administered within a department, this would become increasingly valid and valuable. Widespread collection of similarly defined data could usefully contribute to national benchmarking: a process that is being developed in the USA.

Performance polygons may be constructed with as few (to a lower limit of three) or as many measures as are desired. Our experience is that polygons that have more than eight data points become complicated and perhaps lose some of their visual impact.

One problem with the measurement of clinical outcome is that clinicians are able to ‘play the system’, adjusting their performance according to the measure under scrutiny at one particular time. We believe that multidimensional assessment is less prone to manipulation by individuals who consciously or subconsciously ‘play the system’ and performance polygons enhance such assessment. As well-designed performance polygons examine many aspects of performance, these ‘players’ will have less to gain and should also be more easy to detect.

Clearly, the applicability of performance polygons need not be limited to anaesthesia: by using different measures, polygons might be constructed that examine different aspects of performance and it is quite possible to imagine using routinely collected data to construct individual polygons examining anaesthesia performance, surgical performance, theatre efficiency, ward staff performance, or patient outcomes (e.g. in knee surgery, as in the earlier example). Performance polygons could also be used to compare changes in practice. An example might be multidisciplinary outcomes before and after introduction of an enhanced recovery programme. In research, it might be possible to use a performance polygon as a method of presenting the impact of an intervention not only on a primary outcome but also on several secondary outcomes.

Polygonal representation of data in medicine is not new. In recent years, Intensive Care National Audit and Research Centre (ICNARC) has used polygonal representation of some if its outcome data. These data are used to help identify areas of strengths and weaknesses in a particular intensive care unit (ICU). These data include diverse domains such as in-hospital cardiopulmonary resuscitation before admission to ICU, *Clostridium difficile* infection rate, methicillin-resistant staphylococcal infection rate, and non-clinical transfers. Notably, these domains appear fundamentally unconnected. Less familiar to anaesthetists will be the use of ‘spydergrams™’ to represent the domains of the Medical Outcomes Study Short Form 36 (SF36). The SF36 is a tool to measure self-reported health-related quality of life which comprises 36 questions in eight domains: physical function, physical role, bodily pain, general health perception, vitality, social functioning, emotional role, and mental health index. The data obtained from the SF36 are conventionally represented by a series of bar graphs, but spydergrams™ have been used to represent the data points before and after treatment with various disease-modifying agents and in various disease states. With regard to the SF-36, it is notable that several domains tend to improve together in response to a treatment, although the observed magnitude of improvement is variable and differs between different drugs and disease states. The ICNARC polygons bring together fundamentally unconnected pieces of data and the rheumatology spydergrams bring together data that tend to be related and ‘move in the same direction’. In contrast to these two examples, we judge anaesthetic performance polygons to be most powerful when they represent domains that are both highly interconnected but in opposition to one another (i.e. where improved performance in one domain has the potential to adversely affect performance in another) and when comparator polygons are included.

There are several limitations to performance polygons. The main limitation is that large databases of robust data are required to generate meaningful, reliable polygons: in this regard, they are identical to any other data representation method. Another criticism that might be levelled is that they are unable to represent adverse event rates, particularly rare events. For example, an anaesthetist may choose to avoid regional anaesthesia on the basis that the potential benefits (improved analgesia, reduced nausea and vomiting) rarely outweigh the small risk of a devastating complication (permanent nerve damage). Consequently, the anaesthetist who chooses to avoid regional anaesthesia may compare poorly with an anaesthetist who performs more regional anaesthesia
but is logically less likely to cause such a neuropathy. Similarly, a very quick anaesthetist may cause dental damage more often. These rare outcomes are not represented in performance polygons. While acknowledging the validity of this argument, these limitations also exist when unimodal data are represented. Routinely collected data are unlikely to be an effective method for detecting or monitoring rare events.

In developing performance polygons further, it would be useful to move from using performance polygons as a qualitative tool to using them to aid quantitative analysis of performance. It is superficially attractive to think of the area of the polygon as being mathematically related to quality, but this presupposes that each of the domains measured contributes equally to ‘overall quality’. This can be achieved with careful design of the axes’ scales, but a further problem arises, in that the order in which the domains are represented around the ‘clock face’ affects the total area of the polygon, thus we are not aware of any clear method of relating the polygon area to an overall measure of ‘quality’ in a quantitative manner.

Perhaps inevitably we (or others) often judge ourselves by our most recent success or failure. In an attempt to make a more robust assessment of performance, a number of measures of performance have been suggested. However, many of these domains are not independent. Focusing on single domains can be unhelpful when assessing multidimensional or multidisciplinary outcome, thus we suggest the use of ‘performance polygons’. This method of data representation encourages a balanced appraisal of (anaesthetic) quality, allows comparison with other comparable data sets, and is a visual tool that may have wider applications in healthcare QI. Such a tool might meet several of the requirements for medical revalidation.

Acknowledgement

We thank Dr Andrew Padkin for his intellectual involvement in early discussion of polygonal representation of data without which this project would not have proceeded.

Declaration of interest

None declared.

References


