

A guide to using data for health care quality improvement

June 2008



Published by the Rural and Regional Health and Aged Care Services Division,
Victorian Government Department of Human Services, Melbourne, Victoria.

June 2008

This booklet is available in pdf format and may be downloaded from the VQC
website at www.health.vic.gov.au/qualitycouncil

© Copyright State of Victoria, Department of Human Services, 2008

This publication is copyright. No part may be reproduced by any process except
in accordance with the provisions of the *Copyright Act 1968*

Authorised by the Victorian Government, 50 Lonsdale St., Melbourne 3000.

Printed by Big Print, 50 Lonsdale St, Melbourne VIC 3000.

rcc_080507

Victorian Quality Council Secretariat
Phone 1300 135 427
Email vqc@dhs.vic.gov.au
Website www.health.vic.gov.au/qualitycouncil

Acknowledgements

This resource has been developed for the Victorian Quality Council by:

Project Health

- Fiona Landgren, author
- Jessie Murray, research and documentation

Its development has been overseen by a reference group comprising:

Victorian Quality Council members

- Associate Professor Caroline Brand, Chair
(Clinical Epidemiology and Health Service Evaluation Unit, Melbourne Health)
(Centre for Research Excellence in Patient Safety)
- Ms Kerry Bradley
(Mary MacKillop Aged Care)
- Dr Peter McDougall
(Royal Children's Hospital)
- Associate Professor Les Reti
(The Royal Women's Hospital)

External experts

- Dr Chris Bain
(The Western and Central Melbourne Integrated Cancer Service)
- Ms Anna Donne
(Department of Human Services)
- Ms Mary Draper
(The Royal Women's Hospital)
- Ms Pollyanna Hardy
(Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute)
- Dr Joseph Ibrahim
(Coronial Liaison Services, Victorian Institute of Forensic Medicine)
- Mr Steven McConchie
(Department of Human Services)

Project support was provided by the Victorian Quality Council Management Group

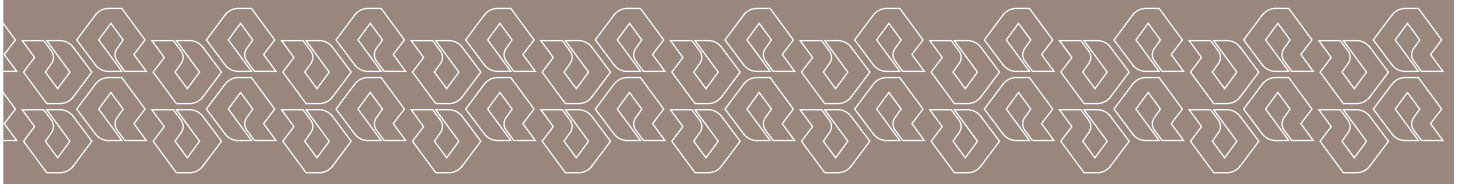
- Mr Oliver Furness
- Ms Jo Heard
- Ms Anna Laird

The contributions of those who reviewed the draft resource are gratefully acknowledged.

Contents

1. INTRODUCTION	4
1.1 Data and quality improvement	5
1.2 Purpose and scope of the guide	5
1.3 About the Victorian Quality Council	5
2. MEASURING AND IMPROVING QUALITY IN HEALTH CARE	6
2.1 What is quality?	6
2.2 What is the role of data in quality improvement?	7
2.3 Using data to help define your quality improvement project	9
2.4 Using data to evaluate existing processes and identify opportunities for improvement	11
2.5 Using data to formulate and prioritise interventions	12
2.6 Using data to measure impact	13
2.7 Using data to guide sustained improvement	13
3. GETTING STARTED – PLANNING YOUR DATA-RELATED ACTIVITIES	14
3.1 Taking a systematic approach	15
3.2 Getting the right advice	15
3.3 Staying on track	17
4. COLLECTING AND STORING DATA	18
4.1 Where can you access the data you need?	19
4.1.1 Existing internal data	19
4.1.2 Existing external data	19
4.2 Data collection techniques and tools	20
4.2.1 Process mapping	20
4.2.2 Brainstorming	20
How to brainstorm	23
How to create an affinity diagram	23
4.2.3 Cause and effect techniques	23
How to create a cause and affect diagram	23
How to use the ‘five why technique’	23
4.2.4 Audit (including clinical record reviews and observations)	25
4.2.5 Surveys and questionnaires	25
4.2.6 Focus groups and key informant interviews	36
How to use key informant interviews	26
How to use focus groups	26
4.3 ‘Good’ data – what is it and how to get it	27
4.3.1 Attributes of good data collection tools	27
4.3.2 Sampling	29
4.3.3 Data entry, checking and cleaning	29
4.4 Storing and managing your data	30

5. ANALYSING AND PRESENTING DATA	31
5.1 Analysing numerical (quantitative) data	32
5.1.1 Counts and sums	32
5.1.2 Ratios, rates and percentages	32
Using ratios, rates and percentages to make comparisons	32
5.1.3 Measures of centre	34
Mean or average	34
Median	34
Mode	34
5.1.4 Measures of variability and spread	36
Range	36
Interquartile range	36
Standard deviation	36
5.1.5 Using statistics to make comparisons	37
How to use and interpret confidence intervals	37
How to use and interpret p values	38
5.1.6 Other measures of causation	39
How to use correlation coefficients	39
5.2 Presenting data	40
5.2.1 Tabulating data	40
5.2.2 Graphing and charting data	42
How to use pie charts	43
How to use bar graphs	43
How to use bar charts to make comparisons	45
How to use box plots	46
How to use histograms and histograms	47
How to use a scatter diagram	48
How to use line graphs and time charts	49
5.3 Analysing qualitative data	50
6. INTERPRETING AND USING THE DATA	51
7. APPENDICES	53
Appendix 1. Data planning template	54
Appendix 2. National/state databases and registries	55
Appendix 3. Useful resources	57
Appendix 4. Glossary	59
Appendix 5. References	63



01

Introduction

Section 01

Collecting and analysing data are central to the function of quality improvement in any health service.

Introduction

1.1 Data and quality improvement

Quality improvement is now a driving force in health care and is an essential aspect of service delivery at all levels. Put simply, quality is everyone's business.

But, unless we measure, it's difficult to know exactly what to improve and whether we have in fact achieved improvement, so efforts to improve systems or processes must be driven by reliable data. Data not only enables us to accurately identify problems, it also assists to prioritise quality improvement initiatives and enables objective assessment of whether change and improvement have indeed occurred. Collecting and analysing data are therefore central to the function of quality improvement in any health service.

The good news is that you don't have to be a statistician to be successful in quality improvement. As this guide demonstrates, the fundamentals of data are accessible and understandable concepts that all health professionals can and should apply to their routine practice.

1.2 Purpose and scope of the guide

The purpose of this guide is to assist all members of the health care team to understand the role of data in quality improvement and how to apply some basic techniques for using data to support their quality improvement efforts.

The guide describes the fundamental concepts associated with data collection, analysis, interpretation and reporting, and how these relate to the various stages of the quality improvement cycle. It also describes how data informs and integrates with the

other key aspects of quality improvement, including communication, people and systems. It assumes a basic understanding of quality improvement principles and provides links for detailed information for those who wish to explore the topic in more detail.

Underpinning the content of this guide is the recognition that careful planning and effective teamwork are also essential elements of any quality improvement initiative.

The focus of the guide is on using data in quality improvement rather than research; however, the data management principles are also largely applicable to research.

1.3 About the Victorian Quality Council

The Victorian Quality Council (VQC) was established in 2001 as an expert strategic advisory group to lead the safety and quality agenda for Victorian health care services. The council is responsible for fostering better quality health services in Victoria by working with stakeholders to develop useful tools and strategies to improve health service safety and quality.

This project stems from the VQC objective to 'Promote access to and use of meaningful, targeted information, relevant to clinicians and patients, to improve practice', and further, to 'assist health services to measure and monitor safety and quality'.

More information:

See the VQC website at:
www.health.vic.gov.au/qualitycouncil

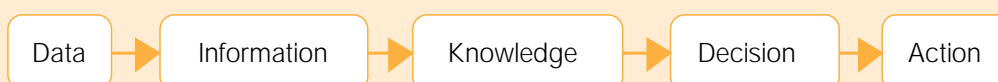


DATA TIP:

The more effort you put into understanding and utilising data, the more you will be rewarded in terms of solving the right problem in the right way.

Figure 1.1 From data to action

Data is the raw material from which information is constructed via processing or interpretation. This information in turn provides knowledge on which decisions and actions are based.



02

Measuring and improving quality in health care

Section 02

This section of the guide aims to equip readers to:

- ✓ recognise the key phases in the quality improvement cycle
- ✓ understand how data supports each stage of the quality improvement cycle
- ✓ understand the role of people and systems in data management.

Measuring and improving quality in health care

2.1 What is quality?

Quality in the health care setting may be defined as the 'extent to which a health care service or product produces a desired outcome'.¹

Quality improvement is a system by which better health outcomes are achieved through analysing and improving service delivery processes.

With this in mind, it is useful to consider quality and quality improvement as spanning a number of key domains. These provide a practical framework for quality measurement and improvement

The National Health Performance Committee² identified nine domains of health system performance:

- Effective
- Appropriate
- Safe
- Efficient
- Responsive
- Accessible
- Continuous
- Capable
- Sustainable

2.2 What is the role of data in quality improvement?

Health service organisations are complex adaptive systems. Making changes to improve quality of care can therefore be a complex business. Fundamentally it requires us to understand what is happening in the delivery of our health services, what factors affect delivery and how we can influence them to achieve improvement. In such a complex system, solid evidence is what we need to support decision making, rather than information based on isolated occurrences, assumptions, emotion or politics.

With this in mind it is useful to consider that quality improvement can be both **reactive** and **proactive**.

For example, most health services collect and analyse data routinely across various quality domains, thus **problems are often clear and self-evident** and the health service reacts to introduce appropriate improvements. Examples include adverse events, infection rates and a range of other clinical indicators.

In addition, health services are likely to proactively look for **opportunities for improvement**. These opportunities become evident when the processes and outcomes are examined more closely.

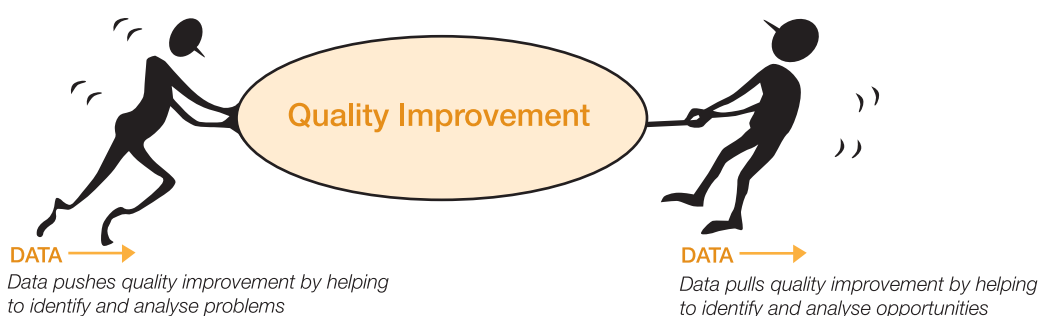
Data therefore helps to 'push' improvement by identifying problems, and to 'pull' improvement by identifying opportunities. Data helps us to understand and improve our service by giving us the tools to **describe** what's going on and to **compare** our performance, either against known standards or against previous performance.

To understand the role of data in quality improvement more clearly it is useful to consider the five phases of the quality improvement cycle. These phases are described briefly overleaf and are best viewed as a continuous cycle of activity as illustrated in Figure 2.3. As this figure shows, the phases take place within, and rely upon, an overall organisational context.

Data plays an important role in each of these phases, and therefore helps us to:

- tackle the right problem (phases 1 and 2)
- implement the right strategies/solutions (phase 3)
- demonstrate the required outcome and monitor for continued improvement (phases 4 and 5).

Figure 2.2 The 'push' and 'pull' of quality improvement

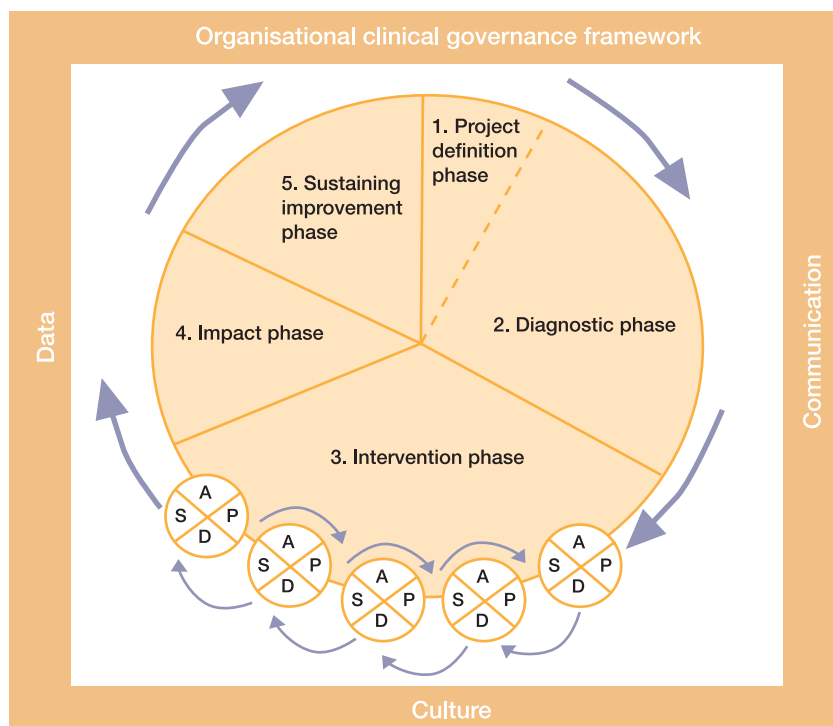


The following sections describe in more detail how data can be used at each of the phases of the quality improvement cycle.

The phases of the quality improvement cycle

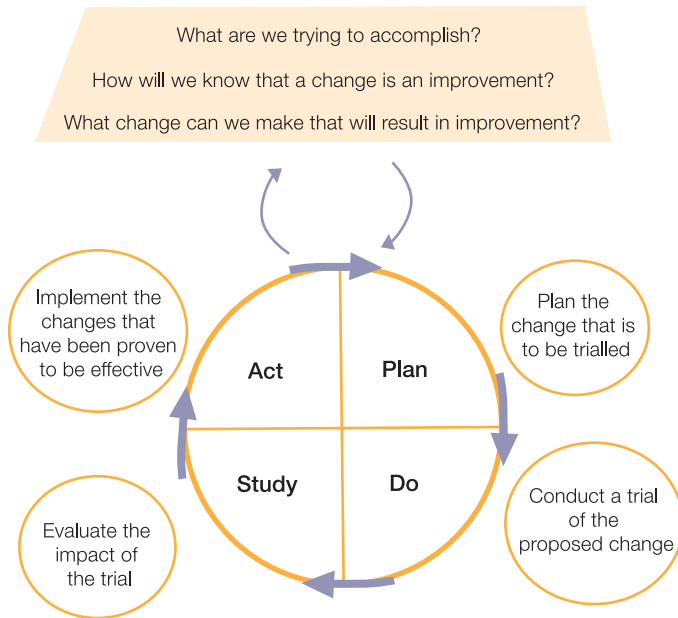
1. **Project definition phase – What is the question or problem?**
This involves identifying the ‘area of interest’ or potential problem area.
2. **Diagnosis phase – What can we improve?**
This involves evaluating existing processes within that area of interest to diagnose potential quality problems and/or opportunities for improvement.
3. **Intervention phase – How can we achieve improvement?**
This involves:
 - determining potential interventions for the processes that require improvement
 - defining performance measures
 - implementing interventions
 - monitoring the progress of improvement.
4. **Impact measurement phase – Have we achieved improvement?**
This involves evaluating the impact of interventions on the predetermined performance measures.
5. **Sustainability phase – Have we sustained improvement?**
This involves monitoring and refining interventions as well as providing feedback in order to sustain the improvement process and ensure the improved processes are integrated into health care delivery as appropriate.

Figure 2.3 The quality improvement cycle



Source: NSW Health 2001, *The clinician's toolkit for improving patient care*
Note: P = Plan, D = Do, S = Study, A = Act

Figure 2.4 The Plan-Do-Study-Act cycle



Like ‘wheels within a wheel’, the intervention phase of the improvement cycle is characterised by a sequence of repeating smaller quality cycles in which interventions are introduced and tested in the real work setting. These repeating cycles represent the Plan-Do-Study-Act model (see Figure 2.4), which provides for an organised and systematic approach to implementation of interventions.

Source: Langley G, Nolan K, Nolan T, Norman C, Provost L 1996, *The improvement guide: a practical approach to enhancing organisational performance*, Jossey Bass Publishers, San Francisco.

2.3 What is the problem or question? Using data to help define your quality improvement project

With (good) data you can:

- ✓ assess current performance and identify performance gaps
- ✓ understand the needs and opinions of stakeholders
- ✓ prioritise problems and improvement projects
- ✓ establish overall aims and targets for improvement
- ✓ establish a clear case for the need for improvement.

- Are clients satisfied with current waiting times at our clinic?
- Is the use of sedatives on discharge appropriate at our hospital and how can it be improved?

As mentioned earlier, data helps to ‘push’ improvement by identifying problems, and to ‘pull’ improvement by identifying opportunities. Without data we can only guess what issues we should address to benefit patients/clients/residents, the health service and other stakeholders.

The domains of quality provide a framework for considering what problems might exist in your health service. Table 2.1 provides some examples, including examples of the types of data that may guide decision making. These can be applied at both a clinical team level as well as an organisational level. Organising your quality issues and the relevant supporting data will help direct the focus of your quality improvement initiatives.

All quality improvement activities start with a problem or a question, for example:

- How can we improve the use of blood products within theatre in line with current clinical practice guidelines?
- Why is the post-surgery infection rate higher in our hospital/department than other comparable hospitals/departments?

**Table 2.1 Identifying areas of concern for quality improvement
(based on the domains of quality - National Quality Committee²)**

Quality domain/criteria	What data/types of measures might help you identify and prioritise quality improvement projects?
<p>Effective Care, intervention or action achieves desired outcome.</p>	<ul style="list-style-type: none"> ○ Clinical indicators ○ Benchmarking against other services/departments ○ Morbidity and mortality meetings/reports
<p>Appropriate Care/intervention/action provided is relevant to the client's needs and is based on established standards.</p>	<ul style="list-style-type: none"> ○ Clinical indicators ○ Audits against international standards/evidence-based guidelines ○ Benchmarking against other services/departments ○ Service utilisation data
<p>Safe The avoidance or reduction to acceptable limits of actual or potential harm from health care management or the environment in which health care is delivered.</p>	<ul style="list-style-type: none"> ○ Adverse events and incidents ○ Sentinel events ○ Clinical indicators ○ Benchmarking against other services/departments ○ Morbidity and mortality meetings/reports ○ Accreditation reports
<p>Efficient Achieving desired results with the most cost-effective use of resources.</p>	<ul style="list-style-type: none"> ○ Service utilisation data ○ Expenditure data ○ Audits of equipment/resource usage ○ Customer complaints ○ Waiting times ○ Failure-to-attend rates
<p>Responsive Service provides respect for all and is client orientated. It includes respect for dignity, confidentiality, participation in choices, promptness, quality of amenities, access to social support networks and choice of provider.</p>	<ul style="list-style-type: none"> ○ Service utilisation data ○ Customer complaints ○ Waiting times ○ Failure-to-attend rates ○ Accreditation reports
<p>Accessible Ability of people to obtain health care at the right place and right time irrespective of income, physical location and cultural background.</p>	<ul style="list-style-type: none"> ○ Service utilisation ○ Customer complaints ○ Waiting times ○ Failure-to-attend rates
<p>Continuous Ability to provide uninterrupted, coordinated care or service across programs, practitioners, organisations and levels over time.</p>	<ul style="list-style-type: none"> ○ Service mapping ○ Clinician feedback ○ Adverse events
<p>Capable An individual's or service's capacity to provide a health service based on skills and knowledge.</p>	<ul style="list-style-type: none"> ○ Waiting times ○ Adverse events ○ Accreditation reports
<p>Sustainable System or organisation's capacity to provide infrastructure such as workforce, facilities and equipment, and be innovative and respond to emerging needs (research, monitoring).</p>	<ul style="list-style-type: none"> ○ Accreditation reports ○ Organisational score boards ○ Integration with data systems ○ Business plans/resource allocation

So how do you actually decide what quality initiative should be addressed next?

A formal exercise called quality impact analysis can be useful. Best conducted as a group, **quality impact analysis** is a brainstorming-type activity that enables a structured consideration of the potential problems and opportunities for improvement within your health service. Based on the list of problems or potential quality initiatives and supporting data, the group may be asked to identify:

- five things that are done frequently
- five things that involve risk
- five things that are of concern to staff or clients.

Participants are then asked to score each item based on the frequency, risk level and general level of concern. Scoring may be, for example, from 1 to 3, where 1 is low frequency, risk or concern and 3 is high frequency of risk or concern. The highest scoring topics should indicate the priority of topics for attention. The activity may also be adapted to address a range of other criteria such as cost or clinical effectiveness.

As with all brainstorming activities, it is important that you involve all relevant stakeholders in this activity to avoid bias. See section 4.2.2 for more information.



DATA TIP: Setting targets for improvement

Use data to help set clear and measurable targets for your improvement initiative.

- *Make sure your targets are linked to your aims and objectives.*
- *Be realistic in your expectations – you won't be able to eliminate all adverse events or all inappropriate admissions.*
- *Express the target as a value, not as a percentage improvement. For example, if baseline throughput in a clinic is five patients per hour and you want to improve by 10%, then state your target as 5.5 patients per hour.*
- *Reassess targets throughout the project and be prepared to modify them in light of experience and in consultation with stakeholders.*

2.4 What can we improve? Using data to evaluate existing processes and identify opportunities for improvement

With (good) data you can:

- ✓ define the processes and people involved in the processes
- ✓ identify problem steps in the process
- ✓ identify and prioritise opportunities for improvement
- ✓ establish clear objectives for improvement of process steps
- ✓ identify barriers and enablers to change.

Once you have identified your area of concern, and you have agreement that the problem is worthy of attention, the next step is to analyse the issue and associated processes in sufficient detail to meet the needs of your project.

This diagnosis phase will enable you to establish the precise nature and cause of the problem and to identify where in the process improvements might need to be made.

Tools likely to be useful in gaining a better understanding of your processes and improvement opportunities include:

- **brainstorming:** members of the improvement team present and record ideas with respect to the 'what, where and how' of the problem and potential solutions
- **process mapping:** the steps in the process are identified and analysed
- **surveys, key informant interviews and focus groups:** stakeholder input is sought regarding the problem and potential solutions, as well as potential barriers to change
- **audits:** specific data is gained regarding existing practice and processes, in comparison to set performance standards
- **control charts:** process performance, based on existing data, is monitored over time and between populations to identify problems and patterns
- **benchmarking:** to compare existing processes/ outcomes with similar services.

These techniques will also help you to refine the overall targets for your improvement project and to develop specific objectives in relation to various process steps. For example, within a broad target of improving access to your service, process analysis may identify specific opportunities such as:

- allocating appointment times for urgent cases and establishing triage procedures to systematically prioritise urgent cases
- expanding the role of the practice nurse to support more effective use of medical personnel time
- reminding customers of their appointment times to reduce failure-to-attend rates.

For more information on these techniques, see section 4 of this guide.

2.5 How can we improve? Using data to formulate and prioritise interventions

With (good) data you can:

- ✓ determine the most appropriate interventions to address your particular problem and to suit your situation
- ✓ prioritise interventions and implementation strategies
- ✓ compare the benefits of alternative interventions and implementation strategies.

When you have a good understanding of your service processes and the opportunities for improvement, you are well placed to select the most appropriate interventions and to establish implementation strategies to suit your particular situation. Your ability to select the appropriate interventions and implementation strategies relies heavily on the data and information collected in the diagnosis phase, as described above.

For example, through audit and process mapping during your diagnosis phase, you may identify that inadequate information received on referrals for outpatient appointments presents a barrier to providing timely and appropriate service. Your interventions may therefore include standardised referral templates, educating referring health professionals and initiating central triage systems. Without data, you may well have approached this scenario in a different way.

The data techniques described in section 2.4 also help you to trial proposed interventions on a small scale to determine their impact and guide further refinement before applying them on a wider scale. 'Good' data is particularly important during this phase to ensure you are able to draw the right conclusions about what is happening as a result of your intervention and whether your findings can be applied to your broader service population.

What is the difference between 'interventions' and 'implementation strategies'?

We talk about the desired **clinical interventions**, that is, the clinical strategies supported by or derived from practice standards, evidence-based guideline recommendations or literature more generally. For example, interventions to reduce pressure ulcers may include using a skin integrity assessment, pressure-relieving underlays or promoting incidental activity.

We also talk about **implementation strategies**, that is, strategies to achieve the desired change in clinical practice. These strategies might include clinician awareness and education strategies, decision support tools, monitoring or reporting processes.

There is often a blurring or overlap between the terms but it is important to recognise the difference. In particular it is important to recognise that successful implementation of clinical interventions will depend on selecting or designing appropriate implementation strategies to suit your particular circumstances.

2.6 Have we achieved improvement? Using data to measure impact

With (good) data you can:

- ✓ assess the impacts of interventions and implementation strategies
- ✓ demonstrate the success of the improvement project to stakeholders.

'In God we trust – all others bring data.'

This is a popular saying among those involved in quality improvement and is particularly relevant to this phase. Without data, you simply cannot determine the impact of your change initiatives or demonstrate your success.

But it is important to understand what sort of data you will need to demonstrate the success of your initiative. For example, in many quality improvement projects you will be looking to introduce a clinical intervention that has already been demonstrated to improve patient outcomes – such as those featured in clinical practice guidelines. An example might be preventing deep vein thrombosis according to measures outlined in the *Best practice guidelines for the prevention of venous thromboembolism*.³ Thus the focus of your data collection and analysis will be on determining whether the change in clinical practice has occurred according to the guidelines, with less emphasis on the clinical impact. In other circumstances, where the expected benefits in terms of patient outcomes may be less defined, you may need to define your data requirements differently.

In this phase, repeated **audits** may be a feature, together with further **surveys**, **key informant interviews** and **focus groups**, to assess the impact of the interventions. Repeated analysis of existing data (such as through **control charts** and other presentations) is also likely.

In this phase you will need to come to grips with some of the **analytical tools and basic statistics** and presentation techniques used for describing and comparing what is happening in your service as a result of your quality improvement interventions. You will need to analyse and present your data effectively in order to draw appropriate conclusions and take appropriate action. You will also need to be able to articulate the benefits and outcomes and identify factors affecting the success of your intervention and its application within your organisation.

2.7 Have we sustained improvement? Using data to guide sustained improvement

With (good) data you can:

- ✓ provide feedback to reinforce change and demonstrate benefits for clinicians and clients/patients
- ✓ identify slippage in practice and the need for repeated intervention or change of intervention approach.

Data has an ongoing role in helping to monitor performance, sustain improvement and assist planning.

Not all improvement projects will use data in the same way. For example, **ongoing** measurement and feedback of performance may be necessary for particular problems or issues, particularly those involving conditions or treatments associated with high risk, high utilisation or high cost, such as incidence of caesarean section, use of particular diagnostic procedures or use of broad-spectrum antibiotics.

For other issues, **periodic** measurement may be appropriate to assess sustained improvement of clinical practice or to assess continuing appropriateness of the approaches.

03

Getting started — planning your data-related activities

Section 03

This section of the guide aims to equip readers to:

- ✓ understand the importance of planning in data-related activities
- ✓ understand the key considerations in data planning
- ✓ understand how to source the necessary expertise.

Getting started — planning your data-related activities

Good planning is the key to successful quality improvement activities. This is particularly so for the data management aspects of your project, which can be complex and resource intensive. So, you need to start with the end in mind.

Assuming you have identified your overall objectives and targets, the main planning issues you will need to consider in relation to data management are:

- the population of people relevant to your question
- what data you will need to accurately assess the problem and be able to judge improvement
- the appropriate methods for collecting and storing the data
- when/how often you will need to collect the data
- how the data should be organised, analysed and presented to demonstrate the extent of the problem and the impact of your improvement initiatives
- ethical considerations, including whether you will need to secure ethics approval for your project.

3.1 Taking a systematic approach

Planning is always easier if it is approached in a systematic and structured way. To assist you in this regard, a planning template is included in Appendix 1 of the guide. The template is presented in a simple table format that you can modify according to the needs of your particular project. The steps include:

1. formulating your overall goal for improvement
2. identifying specific objectives and improvement criteria
3. identifying your target populations both for the clinical interventions and for implementation strategies
4. based on your objectives and target populations, identifying the data that you need to collect to measure the problem and measure improvement
5. determining your data sources and collection methods
6. determining your data analysis techniques and storage requirements
7. determining the resources required accordingly, including skills, computer/software resources and financial resources.

A worked example is also shown in Table 3.1.

The table provides an overview of the data requirements for a project aimed at increasing the appropriate use of prophylactic antibiotics for surgery. Note this is an example only – your data plan may be more or less complex depending on the nature of your quality improvement project. There are no universal rules about what data you need to collect and how you need to go about it – it simply must meet the needs of your project, so the overall goal and objectives are key components to get right from the outset.

Subsequent to your initial overall planning, it is likely that you would want to undertake further detailed planning using a similar template for each project objective or measurement area.

Importantly, your data planning will integrate with the overall planning of your quality improvement initiative, including implementation of clinical interventions and change strategies. You may therefore want to expand the template to summarise findings and identify planned interventions. Other dimensions to address include stakeholder consultation and communication.

3.2 Getting the right advice

This guide provides information about commonly used data collection and analysis techniques and, as such, is a useful starting point for your project planning. Depending on the complexity of your project, it might also be a good idea to secure some data management expertise at the planning stage to save you heartache later on. It is likely that you will benefit from specific guidance in relation to issues such as:

- a literature review
- choosing the appropriate number of observations or sampling the appropriate number of people to make your results meaningful
- questionnaire design
- clinical audit design
- data collection and storage tools such as Excel spreadsheets and databases
- accessing existing data through your health service
- ethical and privacy issues
- more sophisticated statistical analysis methods.

Table 3.1 Example of a planning framework for data management

Improvement project: Prophylactic antibiotics for surgery

Overall goal	Objectives/ improvement criteria	Target populations	Data requirements	Data collection methods	Data analysis	Data resource needs
Improve appropriate use of antibiotics in surgical prophylaxis across all surgical areas	<ul style="list-style-type: none"> Increase percentage of patients receiving the correct antibiotic for prophylaxis Increase percentage of patients receiving prophylaxis at or before commencing surgery Increase percentage of patients receiving appropriate duration of prophylaxis Improvement/no negative impact on postoperative infection rates Reduce percentage of patients receiving antibiotic prophylaxis inappropriately 	<ul style="list-style-type: none"> Clinical intervention – all patients undergoing surgery Implementation strategies – all surgical and nursing staff 	<p>Current procedures</p> <ul style="list-style-type: none"> Nature of current procedures and how they compare with best practice (therapeutic guidelines) <p>Actual clinical practice</p> <ul style="list-style-type: none"> Type of surgery Duration of operation Patient characteristics (age, gender, comorbidities and other factors likely to influence the need for antibiotics) Whether antibiotics prescribed or not If prescribed – time of administration of the first dose of antibiotic; choice of antibiotic; dose of antibiotic; frequency and duration of administration, including discharge Post-operative infection <p>Current attitudes of surgical staff</p> <p>Human factors analysis</p>	<ul style="list-style-type: none"> Initial meeting with surgeons and other staff to identify area of concern Prospective audit over two weeks, prior to intervention, three months following and nine months following Process mapping to identify areas for improvement Staff forums to present initial data, gauge attitudes and barriers to change Nature of current antibiotic restrictions Financial reports of antibiotic expenditure 	<ul style="list-style-type: none"> Qualitative comparison of current documented protocols with recent evidence Quantitative analysis of pre-intervention and post-intervention prescribing data, with breakdown by type of surgery; comorbidities, post op infection. Thematic analysis of feedback from staff forums in relation to attitudes and barriers 	<ul style="list-style-type: none"> 0.5 EFT project manager for 12 months (including involvement in other aspects of implementation) Database to record and analyse prescribing data Independent facilitator to conduct staff forum

Sources of this expertise might include:

- the librarian at your health service or university
- health information services (medical records)
- your quality department or quality manager
- your information technology department
- data managers or research personnel
- experts in designing and using qualitative data
- an ethics committee
- existing staff with skills or interest in epidemiology or statistics
- a statistician at your university.

3.3 Staying on track

In addition to careful planning, there are a number of other 'data essentials' that you should bear in mind throughout your quality improvement project. These are reiterated throughout this guide and are summarised below.



DATA TIP: *If you can, include someone with data management experience on your project team. Start by surveying staff to determine their skills and interests to establish your existing skills bank.*

Data management essentials – tips to keep you on track

1. Plan carefully

Plan your data collection and analysis activities carefully before you start, considering:

- the scope and purpose of your project and the specific questions to be answered
- the availability of resources, including personnel and IT resources
- your target audience and stakeholders, including consumers.

2. Learn from others

Don't reinvent the wheel.

- Search the literature for projects that have tackled similar issues.
- Investigate existing sources of data before you initiate a new collection activity.
- If you need to collect new data, investigate existing and validated data collection methods – there are a wide range of audit tools and toolkits, which are likely to be relevant to your project (see Appendix 3, *Useful resources*).

3. Try things out to avoid mistakes

Test your data collection and analysis techniques on a small scale to identify and correct any problems. This might include:

- conducting a pilot survey
- trialling an audit form
- establishing whether you can access data as planned through your organisation.

4. Involve the team and consider the impact on normal work flow

- Integrate with existing work where possible and ensure everyone is knowledgeable about the requirements and reasons for data collection and the resulting benefits.
- As far as possible, minimise the impact of the data collection exercise on normal work – measurement should be used to speed improvement up, not to slow things down. Remember, the goal is improvement, not to develop a measurement system.

5. Don't be afraid to ask for help

- Get advice when you need it, particularly in the planning and design phase of your data management process – useful contacts include your IT department, librarian, quality department or contacts at a local university.
- Seek feedback at all steps in your data management and quality improvement project.

04

Collecting and storing data

Section 04

This section of the guide aims to equip readers to:

- ✓ understand potential sources of existing data
- ✓ identify how various data collection tools may be used at various stages of the quality improvement cycle
- ✓ understand the importance of data quality and how it can be achieved
- ✓ understand the options for storing and managing data.

Further information is available via the links at the bottom of each section and in Appendix 3, *Useful resources*.

Collecting and storing data

4.1 Where can you access the data you need?

Collecting data from scratch can take a great deal of time and effort, so investigate all possible sources of existing information before you initiate any new data collection processes.

4.1.1 Existing internal data

Useful internal data may already be collected and **reported routinely** by your health service. This data may be accessed from mandatory hospital reporting databases or from systems within individual departments, and may include:

- adverse events, incident reports and sentinel events
- infection rates, isolates
- length of patient stay
- a range of clinical indicators
- service utilisation data including diagnostics, pharmacy, specific procedures
- clinical outcomes from clinical registries
- waiting times for surgery
- waiting times for the emergency department
- customer complaints
- expenditure reports
- use of high-cost medication items.

Additional data that is not collated or formally reported on a routine basis may also be available, but may be accessible through your service computer system or through systems in particular departments. This may include **clinical registries** or **databases** providing access to a range of clinical information.

To find out what data is already being collected within your Organisation, speak to your IT and finance departments as well as data managers and health information services.

Information available in paper-based client medical records may also be considered as 'existing' data, though considerable effort will be required to access it through retrospective or prospective review (see section 4.2.4).

4.1.2 Existing external data

Existing national and state data may also be useful for your quality improvement project, particularly in terms of defining the problem and comparing performance to national averages. Appendix 2 includes many of the databases and registries currently maintained locally and nationally.

Remember that each data source is different in terms of the data elements collected, the time period over which these are collected and how reliable or accurate the data is. Read the small print, including the definitions of the data and the population to which it applies.

Administrative or clinical data?

Data may be administrative or clinical in nature – both are important and useful for quality improvement projects.

Administrative data helps to define your population and the services you deliver. It is often available through service databases and includes items such as:

- client demographics including age, gender, location, cultural background
- service delivery data such as frequency and duration of client contacts; dates and times of services; waiting times
- financial data
- readmission data, length of stay.

Clinical data relates to health needs of your population and the health-related impact of your services. It includes items such as:

- mortality and morbidity
- risk factors
- adverse events
- treatment practices, including drug usage data and diagnostic tests
- infection rates.



DATA TIP:

Talk to data managers or health information services before you start to establish what data already exists within your service.



**DATA
TIP:**

Where possible, use tools that are validated for consistency, and seek out examples of tools used in similar projects.

4.2 Data collection techniques and tools

If you find you need to collect new data, there are a wide range of methods and tools available. The simplest data collection tool is a **simple check sheet**, which can be used to record counts or observations.

A number of other common methods are described in this section.

Some methods, such as brainstorming, process mapping and focus groups, collect qualitative data about the nature of and reasons for your quality problem. Many of these techniques serve both the collection and analysis functions. Others, like clinical audit, collect quantitative data about patients, clinical practices and outcomes.

Most quality improvement projects will require a number of methods to be applied in order to adequately define the quality issue and measure the impacts of your improvement initiatives.

For further information see Appendix 3, *Useful resources*.

4.2.1 Process mapping

Process mapping is a way of collecting data and information about your current processes. It is an activity that should be conducted by the team involved in service delivery, and usually takes place in the **diagnostic phase** of the quality improvement cycle.

Process mapping involves outlining and analysing the steps in your existing processes in order to:

- confirm what is currently happening (sometimes we make incorrect assumptions about what is happening)
- identify problems within the process
- identify how people interact with existing systems and processes (human factors analysis)
- establish the causes of these problems and thus identify opportunities for improvement.

Process mapping helps you to hone in on where an improvement is required. It may involve developing flow diagrams or using matrices.

The simplest process mapping tool is the **basic flow chart**, which outlines the steps of a given process and identifies decision points that affect the outcome at the end of the process.

More complex process mapping techniques incorporate additional dimensions such as:

- **who** in the team is responsible for each step
- **how long** each step takes (program evaluation and review technique – PERT chart)
- the **most important steps** and sub-procedures in the process
- the **costs** associated with various steps and potential interventions.

Quantitative or qualitative?

Data may be quantitative, that is, numerical in nature. Examples include length of patient stay, number of patients treated and infection rates. Data may also be qualitative or descriptive, for example, attitudes and opinions of health care staff, or feedback from clients. Both types of data are valuable in informing your quality improvement initiatives.

More information:

See the Victorian Quality Council guide, *Process mapping: a guide for health service staff*, 2007 available at: www.health.vic.gov.au/qualitycouncil/downloads/process_mapping.pdf

4.2.2 Brainstorming

Brainstorming aims to generate a lot of ideas (a form of data) about a particular subject in a short period of time. It may also be seen as a data analysis technique as it can involve processing data to make decisions and draw conclusions.

Brainstorming may be used at various phases in the quality improvement process, for example, to:

- identify initial problems or areas of concern (phase 1)
- identify the potential causes of a problem (phase 2)
- identify potential interventions to address the problem (phases 2 and 3)
- examine reasons for the success or otherwise of the interventions (phases 3 and 4).

Brainstorming is not a normal meeting – it is deliberately structured to gain ideas from all participants and to avoid bias.

Example – process mapping

Process analysis in an orthopaedic outpatient clinic

In this example, the service delivery team at an outpatient orthopaedic clinic noted waiting times for specialist appointments for osteoarthritis patients as a quality issue. The issue was identified by comparing waiting times with other services as well as feedback from referring GPs and patients. Concern about possible decline in patients' conditions while they were waiting for an appointment was also a factor.

The team undertook a process mapping exercise to identify all the steps in the service provision and administration, from receipt of the GP referral, to the appointment with the specialist.

A simple flow chart was developed, which enabled the group to consider the various steps in the service delivery process. Other data considered included:

- average waiting times for appointments over the past 12 months – this showed a trend towards increasing waiting times (up to 20 weeks)
- rates for failure to attend over the past 12 months (between 15% and 20% per month).

The group identified a number of opportunities for improvement including:

- identifying and managing urgent cases
- communicating with GPs
- reducing failure to attend through improved communication with patients and GPs
- using limited specialist services more effectively through involving non-specialist staff in aspects of assessment as well as education
- identifying significant deterioration through monitoring patients waiting for an appointment.

A matrix (see Figure 4.1) representing activities undertaken by the members of the team was developed. It highlights proposed new activities or changes in existing process to help identify opportunities for improvement. This formed the basis for developing interventions for improvement.

Figure 4.1 Matrix representing activities undertaken by members of the team in the orthopaedic outpatient clinic

Activities	Clerk	Nurse	Doctor	Physio
Prior to appointment				
Receipt and record referral on database	✓			
New Triage referral to establish urgency		✓	✓ (if required)	
Make appointment	✓			
Improved Letter to GP acknowledging referral, indicating waiting time and requesting X-ray and any other information	✓	✓		
New Patient sent pre-appointment questionnaire and other information	✓			
New Patient phoned day before appointment to remind them	✓			
At appointment				
New Clinical assessment (weight, height, history of conservative management, review of questionnaire)		✓		
Clinical assessment (specialist)			✓	
Referrals for conservative management			✓	
New Education about self-management and accessing referred services		✓		
Follow-up appointment as appropriate	✓			
After appointment				
Improved Letter to GP regarding outcome, including referrals and scheduling of surgery if relevant	✓	✓		
New Patient periodically monitored by phone to establish deterioration and need for follow-up		✓		

How to brainstorm

- A group of people appropriate to the project is brought together to address a specific issue. Appropriate representation of stakeholders is an important consideration, so too is ensuring equal contribution from all participants.
- The group are asked to contribute ideas and these are recorded on a whiteboard or flip chart. The success of the session depends on participants being given the freedom to express ideas without restriction or judgement, thus skilled facilitation may be required. If some participants are likely to dominate, a variation may involve asking each participant to contribute one idea at a time moving around the table until all ideas are exhausted. Another approach, called **brainwriting**, requires team members to record a specific number of ideas (say, five ideas) and these are written up and discussed. This approach is beneficial in that it provides anonymity of contributions and therefore promotes open expression of ideas.
- All contributions are recorded, even if related or repetitive.
- When all ideas have been thoroughly exhausted the list is examined and discussed to determine which ideas are worth taking further.

How to create an affinity diagram

For complex quality improvement projects, the brainstorming session may also involve sorting ideas into categories to help focus thinking on key objectives. This results in the production of an **affinity diagram**, which simply means grouping related ideas together into categories. The ideas are usually recorded on sticky notes, categorised by the group and discussed to reach consensus about the nature of the categories and where to place the ideas within the categories. Similar activities are **concept mapping** and **mind mapping**, which describe organising concepts and ideas so that they may be better understood.

Where brainstorming addresses the causes of a particular problem, it may be used to produce a **cause and effect** diagram (see section 4.2.3).

4.2.3 Cause and effect techniques

Another valuable tool for collecting and processing data about the processes you are concerned with is the **cause and effect diagram**, also known as the fishbone diagram.

How to create a cause and affect diagram

The 'effect' is the problem you have identified from the process mapping exercise or other data collection technique. The causes may be identified in a brainstorming session or by more formal data collection techniques. The causes are matched to the effect diagrammatically to produce a visual representation of the problem, which can be considered by the group in terms of identifying possible solutions (see example overleaf).

How to use the 'five why technique'

The **five why technique** is another approach to establishing the causes of identified problems. It works by asking five 'why' questions in succession, each answer leading to another 'why' question. This is best explained by example:

Theatre phones remained out after an incident where power was lost to the theatre telephone and communication system during a planned test of circuits.

1. Why?

The programming software is backed by a battery on the motherboard that failed.

2. Why?

The battery had not been maintained properly.

3. Why?

The maintenance contract had effectively lapsed.

4. Why?

There was an absence of documentation in a past contract transfer.

5. Why?

There is a 'person dependant' culture, along with gaps in documentation.

More information:

NHS The Improvement Network, East Midlands.
www.tin.nhs.uk/tools--techniques/implementation/tools-to-use-here



DATA TIP:

There are now many easy-to-use and cheap (or free) software packages that help you to generate professional-looking affinity diagrams as well as other diagrams and flow charts. Have a look on the internet.

Example – cause and effect diagram

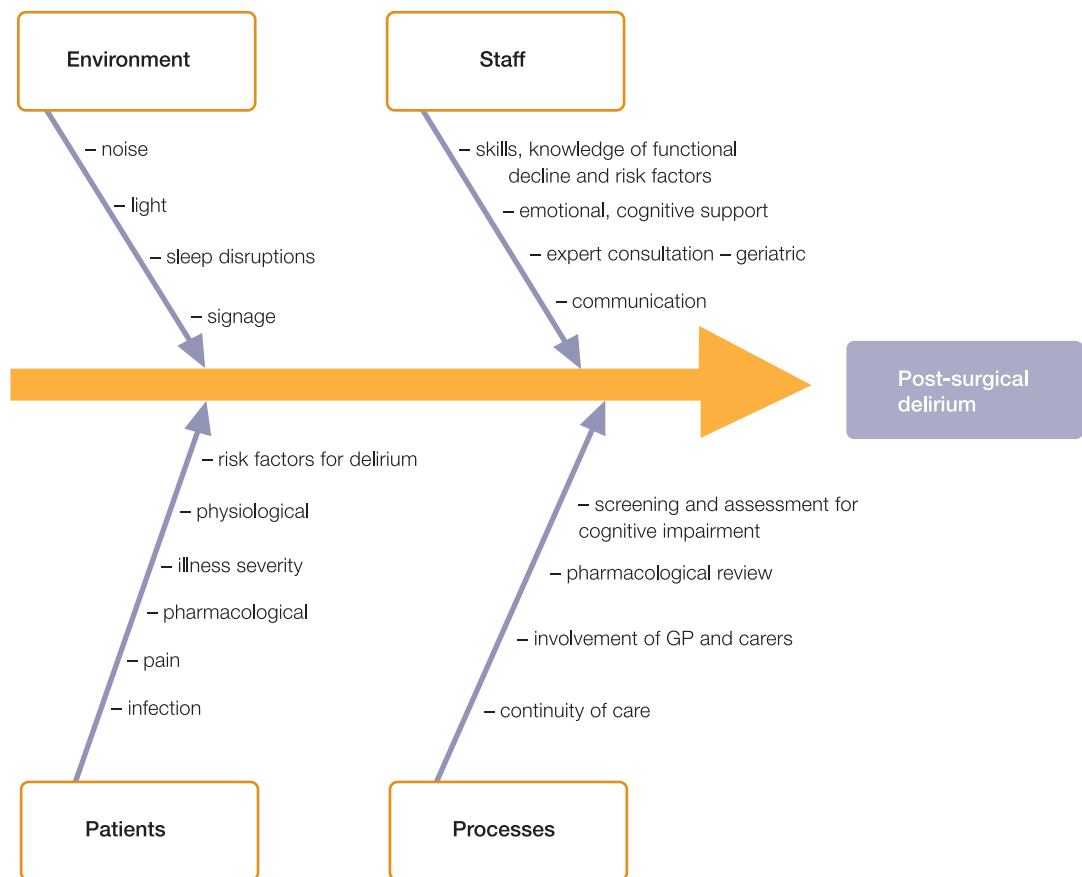
Post-surgical delirium in older patients

In this example, a service delivery team identified a problem with the incidence of post-surgical delirium among older patients in a general surgical ward.

The group conducted a brainstorming session to identify possible causes and opportunities for improvement. Participants in the session included nursing, surgical and pharmacy staff as well as non-clinical support staff. As preparation for the brainstorming session, participants were provided with a copy of recent clinical guidelines on delirium and data relating to the incidence of post-surgical delirium established through a recent audit.

Ideas freely generated during the session were collated under headings and a cause and effect diagram was generated. Further discussion helped to identify priorities for intervention.

Figure 4.2 Cause and effect diagram



4.2.4 Audit (including clinical record reviews and observations)

Audit usually describes a process by which current performance is measured against a known standard or benchmark.

Audits may be used:

- at the diagnosis phase to identify and quantify quality problems, including to establish baseline data prior to the intervention
- at the **impact phase** to establish the effect of the intervention
- at the **sustainability phase** to monitor for sustained improvement.

Record review is a common audit method whereby information is retrospectively or prospectively collected from existing client or system records. The limitation of record review is that some variables may not be recorded and some may be recorded inconsistently, thus a degree of interpretation is required. For this reason it is important to set up a data collection form so that the data is collected as consistently as possible and issues regarding quality and consistency can be flagged.

Audits may also be based on **observations of practice** (such as hand washing). As with record review, ensuring consistency of data collection is vital and should be guided by a standard collection form, as well as clear data definitions and data collection 'rules'. As for all data collection methods, it is also important to test your audit forms and procedures on a small scale first.

Increasingly in the future, record reviews will be based on asking questions of computer-based patient records, which will make them far more efficient and useful for quality improvement activities.

More information:

Clinical audit – how to guides. NHS United Bristol Healthcare. Available from www.ubht.nhs.uk/healthcare-professionals/clinical-audit/how-to-guides.html

4.2.5 Surveys and questionnaires

Surveys and questionnaires provide information about the characteristics, attitudes or behaviours of a group of individuals. They are useful for gathering data to inform the diagnosis phase of your quality improvement project and to assess changes following intervention. Surveys may be administered to staff, customers, suppliers or other stakeholders.

Surveys are not easy to develop, so wherever possible seek existing surveys that have been validated. You may need to adapt an existing survey to your particular needs, but bear in mind that if you make changes, the survey is no longer validated.



DATA TIP: Consider ethical issues... Data collection and analysis should comply with human research and ethical standards – collect only the information that is necessary to measure practice. For more information see *When does quality assurance in health care require independent ethical review?* National Health and Medical Research Council, 2003. Available from: www.nhmrc.gov.au/publications/synopses/e46syn.htm

If you do have to write your own survey, seek guidance from someone with experience and ask for comment. Also be sure to test the survey on a small group of your target audience to identify problems with interpretation and understanding. As with all data collection techniques, it is important to establish your objectives before you write your survey.

It is also important to:

- make sure you have used 'good' survey questions – see the tips overleaf and ask for advice from someone with survey experience
- optimise response rates through appropriate length, appropriate delivery mechanisms (online, hardcopy) and follow-up
- make sure it is clear to your audience why you are collecting the data and the benefits of their participation
- make sure you have addressed confidentiality in your design, collection and analysis processes.


**DATA
TIP:**

Free online survey hosting websites are now available, saving you time and effort in design and distribution.

Good survey questions:

- ✓ are specific
- ✓ are easy to read and understand
- ✓ ask for knowledge or opinion, not both
- ✓ are appropriate to your target audience's level of knowledge and understanding
- ✓ are not 'loaded' or 'leading'
- ✓ do not ask more than one question ('double-barrelled' questions)
- ✓ do not include jargon or acronyms
- ✓ allow for choice of only one option (unless deliberately seeking numerous options)
- ✓ provide reasonable ranges of variation in the response options
- ✓ are unlikely to elicit socially desirable answers
- ✓ are appropriate for age, culture and literacy
- ✓ provide adequate demographic information that will help you analyse the data in a more meaningful way.

More information:

See Appendix 3, *Useful resources*.

4.2.6 Focus groups and key informant interviews

Focus groups and key informant interviews are common methods of gaining qualitative data to guide quality improvement initiatives. They can be used to solicit views, insights and recommendations of staff, clients/patients, technical experts and other stakeholders, therefore providing valuable input into all phases of the quality improvement process.

How to use key informant interviews

Key informant interviews are generally used to gain in-depth information from a limited number of well-informed people. These experts, with their knowledge and understanding, can provide insight into the nature of the quality problem and give recommendations for solutions. They may also be helpful in alerting you to previous work in the area and sources of information, thus saving you time and avoiding duplication of effort.

Key informant interviews are used in preference to focus groups when frank and in-depth answers are required, and when a larger group dynamic might be expected to limit or bias discussion. They may be structured or semi-structured and are most commonly guided by a survey-like proforma or 'script' that lists open-ended questions.

Interviews may be conducted by telephone or face to face. Face to face is most frequently used as it is more conducive to a free exchange of ideas. Choice and training of the interviewer is another important consideration in order to assure an open and appropriate level of discussion.

How to use focus groups

Focus groups are also valuable for informing the various phases of your quality improvement project. They can provide input into diagnosing your quality issue and gaining an understanding of the perspectives of various stakeholders in a short amount of time. They are also useful for identifying and exploring barriers to implementation of your planned change and for resolving problems that arise.

Focus groups provide a flexible format to explore unanticipated issues and to allow interaction of participants, however a clearly stated objective for the group remains important. A disadvantage of focus groups is that discussion can become sidetracked and can be dominated by some participants. Findings are difficult to analyse and care must be taken not to generalise comments to the whole population.

More information:

The UCLA Center for Health Policy Research has some useful and brief guidance material relating to focus groups and key informant interviews as well as a number of other techniques, see: www.healthpolicy.ucla.edu/HealthData/links.html

4.3 'Good' data – what is it and how to get it

'Good' data is data that accurately, reliably and consistently reflects what is really happening in the population or service that you are studying.

To get good data you need to:

- ✓ use good data collection techniques, including appropriate measurement instruments or tools and appropriate sampling techniques
- ✓ ensure data is accurately and consistently entered and stored without duplications or other errors.

4.3.1 Attributes of good data collection tools

Good data is reliable

For a tool to be useful in measuring quality improvement, you need to be confident of its reliability – meaning its ability to get repeatable results with subsequent measurements of the same thing, that is, when you don't expect a change to have occurred. Only with reliable measurements can you be confident that what you are observing is the true situation, and that what you measure after your quality improvement interventions represents a true improvement.

Unreliable results come from unreliable tools, for example a poorly designed survey or audit form, which may be completed differently by different people (inter-observer reliability) or differently by the same person at different times (intra-observer reliability).

Similarly, you want the tool you use to be **responsive**, that is, to identify a change when it has occurred.

The reliability and responsiveness may be related to the nature of the question or data field. For example, a description of disease or clinical outcome may be ambiguous and may therefore be interpreted differently by different people. Clear rules and definitions are important for your data fields and should be established from the outset. A useful exercise is to set up a data dictionary that includes a definition and 'rules' for the various data fields (see Table 4.1 overleaf).

The layout of a form or survey is also important for ensuring reliable completion, so too is trialling your instruments to identify reliability problems.

Good data is valid

Validity is another important attribute of a good measurement tool. Validity simply means that the tool measures what it is supposed to measure.

Many things we measure in quality improvement are physical quantities such as height, temperature and medication use. These are readily observable either directly or indirectly using an appropriate instrument, thus we are confident that we are measuring what we are intending to measure. Testing validity becomes important however when we are looking at variables such as 'quality of life' or 'range of motion' or 'social support', since the measurements of these factors is dependent on their definitions, which may vary from person to person and in the way they are measured. Different instruments may therefore result in different answers, raising the question – what is the right answer?

Validating an instrument or scale is a process by which we determine the confidence with which we make inferences about people based on the scale. A discussion of how we go about validating an instrument or scale is beyond the scope of this guide – suffice it to say, you should, wherever possible, seek out validated instruments to conduct your quality improvement work. When choosing an instrument ask yourself whether it has been tested against a 'gold standard' or other accepted measurement tool.

Good data is unbiased

Bias is created when a tool over measures or under measures the true result, leading to an invalid result. Bias can occur as a result of survey questions that lead the participants in a certain way or when researchers know the treatment being used and their expectations influence their observations. Bias can also result from choosing sample populations that do not represent the broader population (see Table 4.1).

Table 4.1 Sample data dictionary

Variable name label	Data type	Definition/ purpose	Possible values	Restrictions/ checks
Patient information				
Gender	Categorical	Male or female	M or F	Must be M or F
Age	Numerical	Age in years to nearest year	1–110 years	Must be a whole number
Height	Numerical	Height in centimetres	0–300cm	Measure to one decimal point
Weight	Numerical	Weight in kilograms	0–250kg	Measure to one decimal point
Operation				
Operation	Description	Name of operation	Text (refer checklist)	Must be from checklist
Surgeon	Description	Surname and first initial	Text (refer checklist)	Must be from checklist
Duration	Numerical	Duration of operation from full anaesthesia in hours and minutes	Hours	To one decimal point
Antibiotics administered				
Antibiotic	Description	Full generic name of antibiotic	Text (refer checklist)	Must be from checklist
Dose	Numerical	Dose of antibiotic in milligrams	Any	Must be a whole number
Number of doses	Numerical	Number of doses of antibiotic given	0–10	Must be a whole number
Frequency	Numerical	Number of doses ordered per 24-hour period	0–6	Must be a whole number
Qualitative data				
Barriers to change	Description	Barriers identified in interviews	Text	Up to 100 characters

4.3.2 Sampling

Sampling is one of those terms that strikes fear into the hearts of the statistically uninitiated – and for good reason – it is a complex area that can become very technical.

The important thing to understand is that it is unlikely that it will be feasible to collect data on all relevant patients or services in your target population. You will therefore have to work with a ‘sample’ and you will have to give some consideration to whether the sample you select is reasonably representative of the population affected by your quality project.

If your sampling is not representative of your population, **selection bias** is likely to occur, which will affect the validity of your conclusions. Selection bias might occur if, for example:

- you choose to examine a sample group during a time period that is not representative of your usual practice, such as during the holiday period or on weekends
- the health professionals participating in your brainstorming activity do not represent all major stakeholders
- the survey respondents do not include non-English speaking clients.

Sampling refers to the number of observations or subjects as well as how they are chosen. You’ve probably heard of ‘random sampling’, which means the subjects are chosen at random from the population. This is the best sort of sampling, based on the assumption that all subjects in the population will have an equal chance of being selected.

In terms of sample size, there is no magic number as to exactly how many subjects should be included. It will depend on a range of factors including:

- the nature of the question being asked and the information being collected
- the degree of confidence you want to have in demonstrating that any change is a true change, and not just normal variation
- the amount of information being collected, how easy it will be to obtain that information and the resources available.

Sample size calculators (available on the internet) might be useful when trying to determine an appropriate sample size, but the best approach is to seek advice from a statistician, particularly if you are planning a before and after data collection to measure the effects of an intervention. In this case, it will be important that the samples you select for the before and after studies are not only representative of the overall population but are comparable to each other.

More information:

See the various statistical references in Appendix 3, *Useful resources*.

4.3.3 Data entry, checking and cleaning

Data entry is an important step in your data management process and can be a source of considerable error if not undertaken carefully. Appropriate training for the data entry personnel is important, including ensuring a basic understanding of the data they are entering. The data dictionary mentioned earlier is important in this regard.

Automatic edit checks are a further means of ensuring data accuracy. Programmers can add rules to database programs to:

- prevent certain fields from being filled out when they shouldn’t be
- restrict the format or range of data values that can be entered
- ensure the completeness of the record.

Manually edit checking the data is also advisable. This can be done by scanning the data in a spreadsheet (such as to detect gaps or unusual values) or can be done by creating tables. These tables might identify, for example, where there is an unacceptable level of missing information or it might identify incorrect data values; for example, an age value of 125 is likely to be incorrect (refer also to section 5.2.1 *Tabulating data*).

If problems are identified you will then need to:

- investigate the reasons for the bad data
- determine whether and how it can be corrected
- perform accuracy checks to compare the original data collection tool with the entered data.



EpiData (www.epidata.dk) is an easy-to-use public domain program that permits detailed data entry, quality control and datachecks.



DATA TIP:

Data tip – Whether you use computer- or paper-based storage methods, be sure to take a copy of your data. If you make a mistake in your calculations or accidentally erase part of your data, you will have the original to go back to. It's also a good idea to give your copy file a different name.

4.4 Storing and managing your data

Data storage is another important consideration and one that should be addressed early in your project planning. For straightforward projects, data may be stored in paper records. Computer-based storage of some kind is a practical and safe approach. There are three main choices:

- spreadsheet programs (such as Microsoft Excel, Lotus 1-2-3, Quattro Pro)
- database programs (such as Microsoft Access, Lotus Approach, Filemaker, Ability Database)
- statistical programs (such as Statistical Package for the Social Sciences – SPSS, SAS).

The best option for you will depend on the nature of your data collection project and the skills and resources available to you. Before you make a decision, speak to someone within your organisation about the options already available and what options might best suit your project.

Whether you are collecting by paper or electronic means, structure your data in a logical, relatable way. For example, keep all data related to patients together and keep all data related to a ward together. If in doubt, seek assistance – it is better to set up your data properly in the first place.

Table 4.2 Data storage/analysis options

Storage method	Features/ advantages	Disadvantages	Typical usage	Skills required
Spreadsheet programs	<ul style="list-style-type: none"> ○ Easy to set up and enter data ○ Can produce simple graphs ○ Readily available in most organisations ○ A range of 'add ons' available via the internet to enhance applications 	<ul style="list-style-type: none"> ○ Information is not stored as a complete record 	<ul style="list-style-type: none"> ○ Straight forward projects with limited variables and limited requirements for analysis and comparison 	<ul style="list-style-type: none"> ○ Basic to intermediate computer skills required for setup and analysis
Database programs	<ul style="list-style-type: none"> ○ Information stored as complete record ○ Can easily query data ○ Readily available in many organisations 	<ul style="list-style-type: none"> ○ Specific skills to set up ○ More difficult to undertake analysis as user requires knowledge of how to run a query 	<ul style="list-style-type: none"> ○ More complex projects especially data collected over multiple time points 	<ul style="list-style-type: none"> ○ Specific training required for set up
Statistical packages	<ul style="list-style-type: none"> ○ Enables data analysis as well as data entry storage ○ Enables manipulation of data, comparisons and statistical analysis 	<ul style="list-style-type: none"> ○ Needs specific training ○ Often needs to be explicitly purchased and can be expensive 	<ul style="list-style-type: none"> ○ Projects requiring more complex analysis, for example, to find associations between variables and outcomes 	<ul style="list-style-type: none"> ○ Training in use of system and in statistical processes and approaches



05

Analysing and presenting data

Section 05

This section of the guide aims to equip readers to:

- ✓ recognise and understand the common terms used in data analysis
- ✓ understand and apply some of the basic methods for describing, comparing and presenting data
- ✓ understand the purpose of more complex data analysis techniques.

Analysing and presenting data

There are a number of basic methods that help you to organise, analyse and present your data in order to support your quality improvement activities. These methods help you:

- ✓ describe what is happening in your study population
- ✓ identify relationships between variables
- ✓ identify whether improvements have occurred
- ✓ monitor improvements over time
- ✓ determine the significance of your results
- ✓ communicate your conclusions effectively.

Some of these methods are described in this section. Further information is available via the links included in Appendix 3, *Useful resources*.

Understanding your variables

Variable is another term used to describe a value or type of data. Height, age, gender, amount of income, country of birth, language, diagnosis, infection rate are all examples of variables. Variables may be classified in a number of ways.

Numerical and categorical data/variables

Numerical variables are expressed as numbers such as height, weight, dosage and time. Categorical variables are those that can be sorted according to categories, such as type of operation and gender.

Continuous and discrete variables

A continuous variable is a numeric variable that can assume an infinite number of values. For example, height, weight, age, distance, time and temperature. A discrete variable is one that can only assume certain values. For example, responses to a five-point rating scale can only take on the values 1, 2, 3, 4 or 5. The variable cannot have the value 1.7.

5.1 Analysing numerical (quantitative) data

Raw numerical data is hard to absorb, thus basic statistics are used to **organise and summarise** information about a dataset. This helps describe what is happening in your sample population and can help guide the need for further analysis. The basic summarising statistics and techniques that you are likely to use when you first look at your data are described below. Also described are techniques for comparing data, being one of the basic requirements for quality improvement.

5.1.1 Counts and sums

Counts are simply a count of how many items or observations you have in your sample, for example the number of people receiving a particular treatment or the number of people responding to a survey. In statistics they are sometimes referred to as 'n', indicated by a small letter n.

Sums involve adding up the numbers in each set of observations. For example, 20 people responding to the survey feel that current processes for counselling discharge patients are inadequate. Sums are usually expressed in relation to 'n', that is, 20 of the 100 people surveyed feel that current processes for counselling discharge patients are inadequate.

5.1.2 Ratios, rates and percentages

Simple counts and sums are just the beginning. Statistics such as rates, ratios and percentages help you to standardise your data so that it is expressed in a meaningful way that can be readily compared with other data.

A **ratio** is a fraction, expressed in its simplest terms, that describes two groups relative to one another. For example, the ratio of females to males in your study group may be 3 to 2, meaning that for every three females there are two males.

A **rate** is a ratio that describes one quantity in relation to a certain unit. For example the rate of hospital readmissions may be expressed as four per 100 people discharged, the rate of falls or other risks may be expressed per 1,000 bed days. With most types of data there are conventions relating to how rates are expressed. If in doubt, seek advice or do some research to find similar work done by others.

Ratios and rates may also be expressed as **percentages**, such as 60% females and 4% hospital readmissions in relation to the above examples.

How you choose to express your data will depend on the nature of the data and how you plan to use it. You will find that for certain types of data there are certain traditions or standards of expression.

Table 5.1 Describing your population with summary statistics

	Females versus males
Counts	228 females, 152 males
Ratio	3 to 2
Rate	60 females per 100 population
Percentage	60% females

Using ratios, rates and percentages to make comparisons

Ratios, rates and percentages are also useful when it comes to comparing datasets. For example, you might want to compare populations within your health service to see where problems lie, or you might want to make comparisons before and after a quality improvement initiative.

Table 5.2 compares prevalence of pressure ulcers across a number of specialty areas before and after an intervention to address risk factors. From this table you can make some general observations about how the specialty areas compare with each other and the impact of the intervention.

For more information about using statistical techniques to compare percentages and interpret the comparative results, see section 5.1.5.



DATA TIP: Be careful when interpreting and using percentages though. You can't assume that when a large percentage increase is reported that the result is a large or significant one. For example, a rare event or disease may be reported to increase by a large percentage but, given the small numbers involved, this impact might not be as meaningful as it seems by looking at the percentages alone. It is therefore important to consider both the actual numbers and the percentages when interpreting your data.

Table 5.2 Prevalence of pressure ulcers before and after intervention

Specialty area	% prevalence before intervention	% prevalence after intervention	Percentage change
Spinal	41.4 (n=35)	60.9 (n=23)	19.5
Medical	27.8 (n=1,460)	18.6 (n=1,645)	-9.2
Rehabilitation	29.9 (n=946)	24.0 (n=1,101)	-4.7
Surgical	22.4 (n=1,317)	14.1 (n=1,645)	-8.3

5.1.3 Measures of centre

The most common way of summarising and comparing numerical data is to describe where the centre is. This gives you an idea of what the 'most' common, normal or representative results might be. The centre can be measured in different ways so it is important to understand how these different approaches might affect the conclusions you draw about your data.

Mean or average

The mean is the most commonly used measure of centre. It is calculated by adding up all the numbers in the dataset then dividing by the number of numbers in the dataset.

For example, the average time waited for people currently on the waiting list to receive joint replacement surgery is calculated by adding up all the waiting times of people who are currently on the list and dividing by the number of people on the waiting list.

A limitation of the mean is that it can be affected by 'outliers', that is, extreme values at either end of the scale, thus it may not be truly representative of the sample population.

Median

The median is the value that lies exactly in the middle of the dataset, that is, 50% of values lie above it and 50% of values lie below it. To find the median:

1. arrange the numbers in your dataset from smallest to largest

2. if there are an odd number of numbers, choose the number that is in the middle – that is your median
3. if there are an even number of numbers, take the two middle numbers and average them to find the median.

Unlike the mean, the median is not affected by outliers. It is the preferred measure of centre for skewed distributions (see overleaf).

Mode

The mode is the most common value, that is, the value with the highest frequency. To find the mode, arrange the numbers in your dataset from smallest to largest. The number with highest frequency is the mode.

The figures below and overleaf show data from three different surgical waiting lists. The data in Figure 5.1 forms 'normal distribution', that is, it is symmetrical with a single bump in the middle. For this data, the mean, median and mode are all the same.

The data in Figure 5.2 has 'outliers' on the lower end of the horizontal axis. It is described as skewed left or negatively skewed. For this list, the mean or average will be smaller than the median.

The data from Figure 5.3 has 'outliers' on the higher end of the horizontal axis. It is described as skewed right, or positively skewed, and the mean or average is higher than the median.

If the data is 'skewed' you should use the median rather than the mean to describe the data.

Figure 5.1 Waiting list 1 – normal distribution

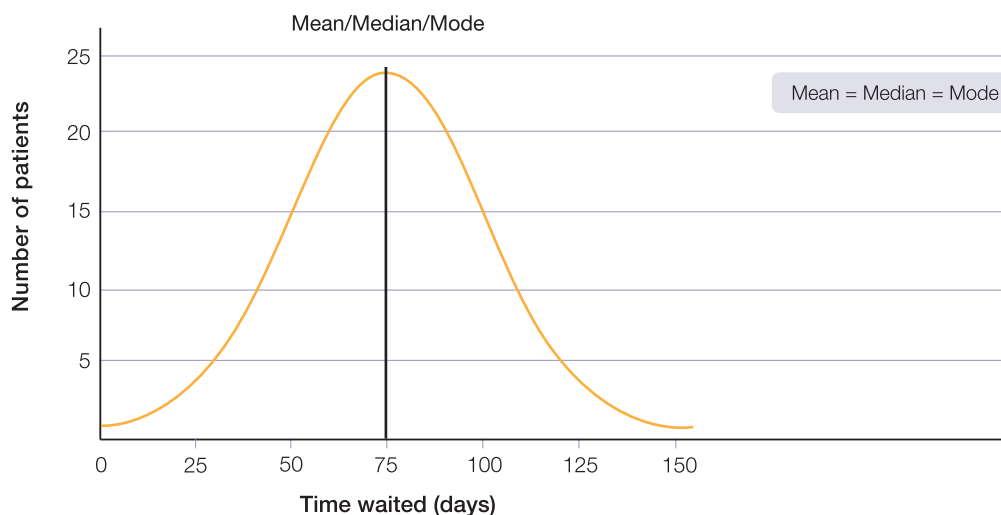


Figure 5.2 Waiting list 2 – left (negative) skew

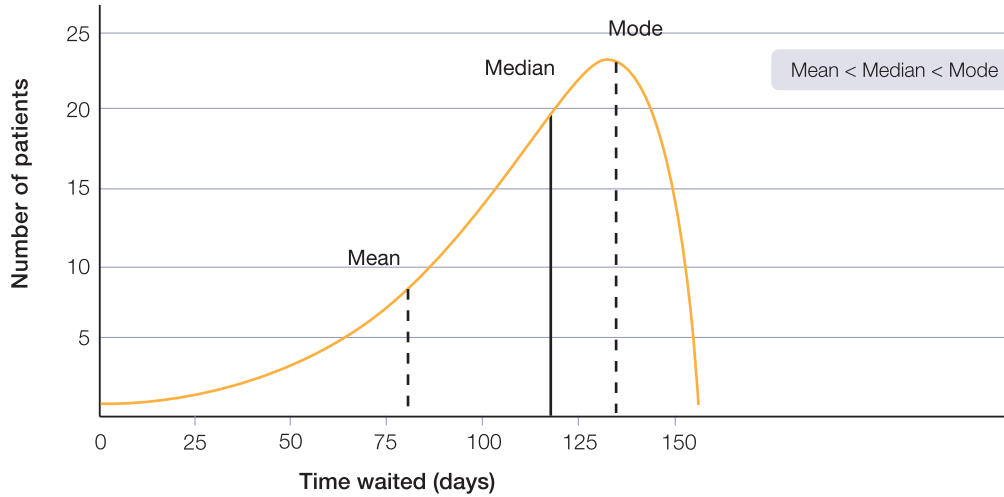
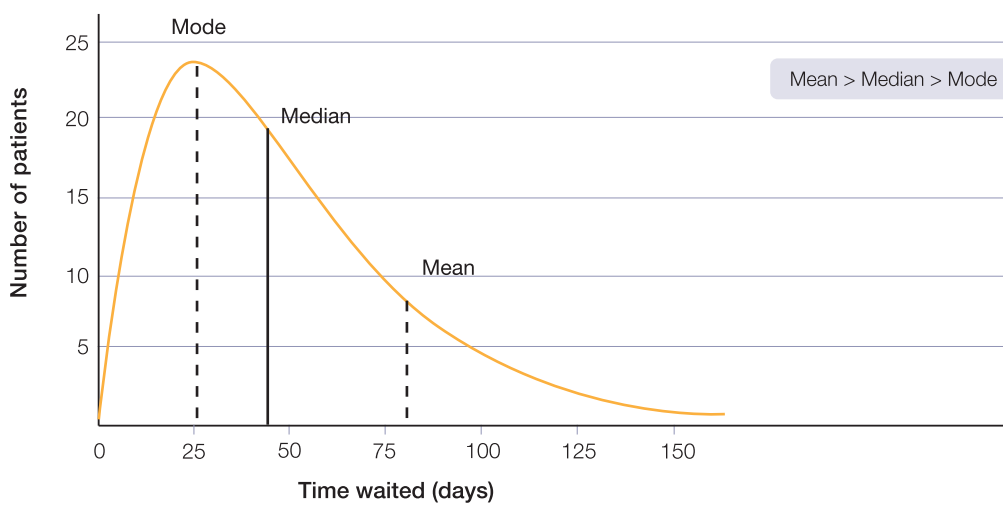


Figure 5.3 Waiting list 3 – right (positive) skew



DATA TIP: Confused about mean, median and mode?

As a general rule, place them in alphabetical order, that's mean, median, mode, then:

- If your data forms a normal distribution then **mean=median=mode**.
- If your data is skewed left then **mean<median<mode**.
- If your data is skewed right then **mean>median>mode**.

Measures of centre are a common basis for comparison in quality improvement activities. For example, you may compare means in your population before and after a quality intervention.

For more information about using statistical techniques to compare means and interpret the comparative results, see section 5.1.5.

5.1.4 Measures of variability and spread

Consistency of practice is a key aspect of quality in health care, thus an important characteristic of a dataset can be how variable it is. For example, in assessing the prescribing patterns of antibiotics, it will help you to know not only whether practice complies with guideline recommendations or not, but how the practice varies. Knowing the variability will help you develop appropriate intervention strategies and will also help you to demonstrate the success of your implementation strategies in improving practice.

Variability is best visualised using **box plots**, but **histograms** and **time charts** can also be useful (see section 5.2.2). However, sometimes it can be helpful to actually measure the variability – this is where statistics come to the fore.

Range (minimum and maximum)

The range is a commonly used measure of variability, mainly because it is easy to calculate. It is defined by the minimum or lowest observation and the maximum or highest observation. The range however is not very useful for describing a dataset and certainly not for comparing datasets, as it relies on only two numbers, the minimum and the maximum. It does not describe the overall variability. Thus the standard deviation is a preferred measure of variability.

The range however can also be used to identify incorrect data, for example, an age value of 125 is likely to be incorrect and should lead you to consider the accuracy of your data.

Interquartile range

The interquartile range is another way of expressing range. It is the difference between the upper and lower quartile values where each quartile represents the division of the data into four equal sized groups (see also box plots, section 5.2.2).

Range = highest value – lowest value

Interquartile range = upper quartile – lower quartile

Standard deviation

As a measure of variability, the standard deviation describes your sample population in terms of the average distance from the centre or mean. 's' is used to denote the standard deviation of a sample population. The smaller the standard deviation, the more closely the data clusters around the mean. The larger the standard deviation, the further away the dataset is from the middle, in other words, the more variable your data is.

Using standard deviation as a measure of variability relies on the assumption that your data resembles a 'normal distribution' (see Figure 5.4).

How to calculate the standard deviation

1. Find the average of the dataset (add up all the numbers and divide by the number of numbers in the dataset).
2. Take each number and subtract the average from it.
3. Square each of the differences.
4. Add up all the results from step 3.
5. Divide the sum of squares (step 4) by the number of numbers in the dataset minus one ($n-1$).
6. Take the square root of the number you get.

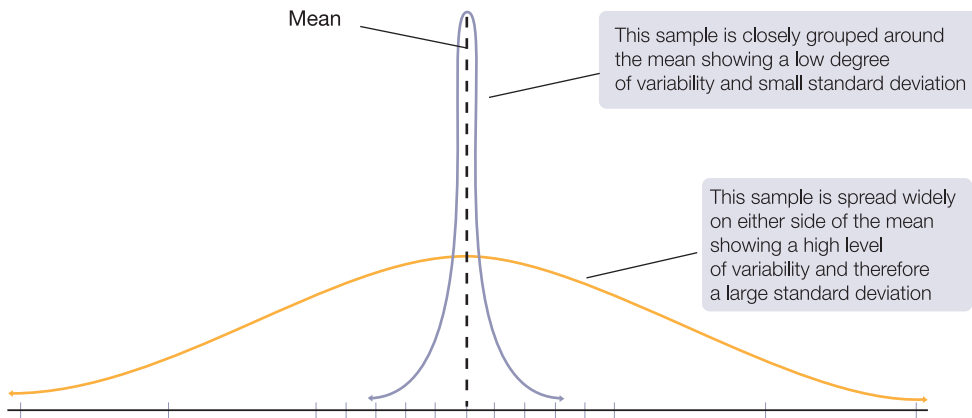
$$\sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$



DATA TIP:

Most statistical packages calculate the standard deviation automatically.

Figure 5.4 Standard deviation



5.1.5 Using statistics to make comparisons

In quality improvement, much of your data analysis is aimed at determining whether your initiatives have resulted in improvement and whether the size of that improvement is significant or beneficial. While straight comparisons of sample means and proportions are helpful, some further statistical techniques can add to the power of your conclusions. Here we talk about two of the most common statistical tools – confidence intervals and statistical significance (or ‘p’ values). This discussion is focussed on how you might interpret the application of these tools, rather than how you might calculate the statistics. For more detailed information see the general statistical references listed in Appendix 3, *Useful resources*.

A common process in making comparisons involves:

- estimating the size of the difference between sample observations (such as by comparing population means or proportions)
- deriving a confidence interval for the difference
- deriving a p value to test whether there is a true association between the differences observed.

How to use and interpret confidence intervals

As discussed in section 4, when you collect data for your quality improvement projects, it is not possible to collect data from your whole population – you have to take a sample and use that sample to estimate your population. This means that there will be a certain amount of error involved. The **margin of error** is a useful statistic to measure the amount that your sample might be expected to differ from the actual population. The margin of error is expressed as an amount above and below the sample value. The larger your sample size, the more likely it is to approximate your whole population, thus the larger your sample, the smaller the margin of error.

When you combine your sample value with the margin of error, you come up with a **confidence interval**, that is, a range of values that you would expect your population to fall within. The confidence interval expresses your level of confidence that the sample value represents the true value as seen in the overall population.

The calculation of the confidence interval depends on just how confident you want to be that your population value falls within the range. Commonly, statisticians are satisfied with a confidence level of 95%.



DATA TIP:

The applicability of these and other statistical techniques will depend on a range of factors including your sample size, so seek appropriate advice before applying them to your own data.

There are a number of common applications of confidence intervals that you are likely to come across in your quality improvement work. These are described briefly below. For more information, including formulas for calculating confidence intervals, see the statistics references in Appendix 3 or speak to a statistician.

○ *Confidence interval for the population mean*

This is used to describe a population when the characteristic being measured is numerical (such as height, weight, age, dose, physiologic measure or waiting time).

For example, the waiting time for appointments for clients referred to your clinic might be expressed as a mean of 13.5 weeks with a 95% confidence interval of 11.6 to 15.3 weeks (95% CI 11.6-15.3). This means that you expect your population on average would wait between 11.6 and 15.3 weeks for an appointment.

○ *Confidence interval for the difference of two means*

This is used to determine the precision of the estimate of the difference between the two means, such as when you observe differences before and after an intervention.

For example, comparing waiting times before (13.5 weeks as above) with after (11.6 weeks). The difference in the means is 1.9 weeks. Let's say, based on your sample sizes, the margin of error for the difference in the means is calculated as plus or minus 0.71, and thus your 95% confidence interval is 1.19 to 2.61 weeks. This means you can say with 95% confidence that your new systems, on average, improved the mean waiting time from between 1.19 and 2.61 weeks.

○ *Confidence interval for the population proportion*

This is used when the characteristic being measured is categorical and the data is expressed in terms of a percentage or proportion, for example, the proportion of people with pressure ulcers, the proportion of people with post-operative infection or the proportion of people maintaining a particular opinion.

For example, the prevalence of pressure ulcers in a sample population may be expressed as 9.6% with a 95% confidence interval of 6.0% to 14.2% or (95% CI 6.0%-14.2%). This means that you are 95% confident that your population value falls between 6.0% and 14.2%.

○ *Confidence interval for the difference of two proportions*

This is used to determine the precision of the estimate of the difference between the two proportions.

For example, in our example above the prevalence of pressure ulcer is measured as 8.5% following the intervention. The difference is 1.1% when compared with our pre-intervention sample. When you calculate your margin of error, you get a value of 1.4%, and thus your 95% confidence interval is - 0.3 – 2.5%. This means you can say with 95% confidence that your intervention may have caused a worsening of ulcer prevalence by 0.3% or up to 2.5% improvement.



DATA TIP: What is the difference between confidence intervals and p values when making comparisons?

- A confidence interval gives you the best range of feasible values for the difference.
- The p value addresses whether the observed difference in the sample could be due to chance or whether there is a significant difference in the populations.

How to use and interpret p values (significance test)

The p value is an expression of whether the difference you have observed in your study samples could be due to chance or whether there is indeed a 'significant' difference. The smaller the p value, the more likely the result is not due to chance and thus represents a significant result. The larger the p value, the greater the chance that the difference observed could be the result of sampling variation.

p values less than 0.05 are commonly interpreted as 'statistically significant'.

The size of the p value depends on the size of the sample, so be alert to possible mistakes that can occur in interpreting these values. For example, potentially important differences observed in small studies may be ignored if the interpretation is based on the p value alone. As such, you should always consider the confidence interval in association with the p value.

It is also important to remember that 'statistically significant' does not necessarily equal 'clinically significant' so the interpretation of your results should involve input from appropriate clinical experts.

More information:

For a more detailed discussion about the interpretation of confidence intervals and p-values, including examples, see the general statistics references listed in Appendix 3, *Useful resources*.

5.1.6 Other measures of causation

The technique of correlation is also used in statistics to determine whether a relationship exists between two variables. In cases where both variables are numerical (quantitative) the data can be organised in a scatter diagram or scatter plot to determine a possible relationship. Once plotted, if the points resemble an uphill line this indicates a positive linear relationship. If the data points resemble a downhill trend this indicates a negative (or inverse) relationship. If the data doesn't seem to resemble a line at all this means there is no linear relationship. It is important to remember that plotting the data only show a *possible* relationship and does not confirm cause and effect. Also remember that relationships may be non-linear – but that is a topic best left to the statisticians.

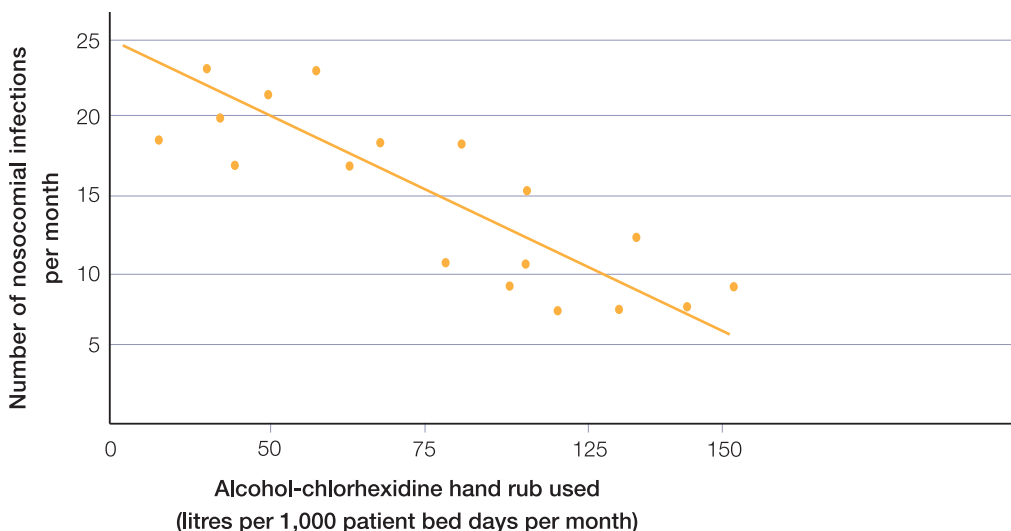
How to use correlation coefficients

You can also determine the **strength of the relationship** between the two numerical variables by calculating the correlation coefficient. The value of the correlation coefficient (r) is always between -1 and 1. Most statisticians like to see correlations of >0.6 or < -0.6 .

- A correlation of exactly +1 indicates a perfect positive relationship.
- A correlation close to +1 indicates a strong positive relationship.
- A correlation of exactly -1 indicates a perfect negative (inverse) relationship.
- A correlation coefficient close to -1 indicates a strong negative (inverse) relationship.
- A correlation close to 0 means there is not linear relationship.

In Figure 5.5, an improvement project relating to hand hygiene seeks to demonstrate a possible relationship between nosocomial infections and the use of alcohol-chlorhexidine hand rub. Monthly figures for the two variables are monitored over 18 months from before the hand hygiene intervention. The plot shows a negative relationship between the two variables, reflecting that increased use of the handrub (plotted on the horizontal axis) is associated with reduced infection rates (plotted on the vertical axis).

Figure 5.5 Nosocomial infection (per month) vs alcohol-chlorhexidine hand rub usage (June 2006 – December 2007)



5.2 Presenting data

There are various techniques for organising and presenting data. These are helpful in guiding your analysis and are also valuable for communicating your findings. A number of commonly used techniques are described in this section.

5.2.1 Tabulating data

Tabulating data can be a very useful way of coming to grips with your dataset. A table simply presents the data in row and column format, and can provide a useful overview to guide further analysis or further tabulations. Tables can also be used to make comparisons between datasets and lead to some initial conclusions.

Creating good tables takes a bit of practice, but the following tips might prove useful:

- Keep it simple – don't try to include too much information in one table.
- Watch your units – make sure they are consistent throughout.
- Don't get carried away with decimal points – round off to one decimal point or whole numbers as appropriate to make tables easier to read and be consistent throughout.
- Include both raw numbers and percentages: n (%).
- Always include 'n', being the number in your total population.
- Identify where there is missing data.
- If grouping data, be sure that the groups don't overlap and that the groupings are evenly spaced.
- When comparing data before and after or between groups, include **confidence intervals** where possible.

Overleaf are some examples of tables developed to present results of a client satisfaction survey for a number of outpatient services.

The first table (Table 5.3) shows response rates for the survey across different services and across the two different sites. This sort of data provides a basis for you to assess whether the information you have collected is representative of your overall population – for example, a low response rate may make you cautious about your conclusions.

This table includes responses from all services involved in the survey. It enables you to:

- identify the response rate for each service (and consider whether it is reasonably representative of the service)
- identify the response rate for each site
- identify the overall response rate across both sites
- compare the response rates between similar services at different sites (such as orthopaedics)
- compare overall response rates between sites.

The second table (Table 5.4) presents survey results for one of the sites. You may want to prepare an additional table to compare the two sites but it is best to start simple. The table shows basic frequencies (how often people gave a particular answer) for various aspects of the service. It shows 'dissatisfaction' data as this can point you more clearly to opportunities for improvement.

This table focuses on the survey results for one site so that you are not overwhelmed by data in the first instance. It presents information about the level of dissatisfaction (rather than satisfaction) with elements of the service to help you identify opportunities for improvement.

The table enables you to:

- identify areas where there may be a significant level of dissatisfaction
- identify whether dissatisfaction is limited to one service or is common across the site.

More information:

See *How to display data* by Freeman, Walters and Campbell, Blackwell Publishing, 2008.

Table 5.3 Client satisfaction survey response rates (two sites, five services)

Sites	Site 1			Site 2				Total
Services	Service 1 Ortho	Service 2 Maternity	Total	Service 3 Ortho	Service 4 Diabetes	Service 5 Oncology	Total	
No. of surveys sent	140	76	216	47	50	36	133	349
Number of responses by service (%)	75 (54)	50 (41)	125 (58)	35 (74)	32 (64)	31 (86)	98 (74)	223 (64)

Table 5.4 Client satisfaction survey results – site 1

Elements of service	Dissatisfied or very dissatisfied (combined responses 1 and 2 from 5 point scale) *, n (%)		
	Service 1 (n=75) Orthopaedics	Service 2 (n=50) Maternity	Total (n=125)
Availability of suitable appointment time	15 (20)	10 (20)	25 (20)
Waiting time to be seen	17 (23)	5 (10)	23 (18)
Courtesy of staff	4 (3)	0 (0)	4 (3)
Information provided	4 (3)	8 (16)	12 (10)
Opportunity to contribute own thoughts and opinions during consultation	8 (11)	4 (8)	12 (10)
Overall satisfaction with service	10 (13)	7 (14)	17 (14)



Most of the graphs shown in this section can be produced from a simple spreadsheet program like Excel.

5.2.2 Graphing and charting data

Graphs and charts are like visual tables and are a useful way of presenting data to identify patterns or trends. They are also useful for communicating your findings to others.

Graphs usually plot two types of data using a grid in which one set of data is plotted along the horizontal or X-axis, and the other is plotted along the vertical or Y-axis. Table 2.5 summarises some of the common graphing and charting methods, indicating when they might be useful.

Table 5.5 Summary of graphs and charts

What do you want to show?	What graphs/charts to use?
Basic population characteristics – such as age ranges, ethnicity	Pie chart Bar chart
Measures of magnitude including comparisons	Bar chart Box plot
How often something occurs (frequency) –such as a clinical practice; an adverse event, including comparisons	Pie chart Bar chart Pareto chart Box plot
Trends over time	Line graph (also known as a time chart) Control chart
Distribution of data	Histogram Histogram Scatter diagram
Whether there is relationship or association between two things (cause)	Scatter diagram Box plot

When creating graphs and charts, similar rules apply to those mentioned above in relation to tables. In particular you should:

- Keep it simple. Don't try to include too much information – use a series of graphs or charts rather than trying to communicate too much at once.
- Avoid complex colour schemes and three dimensional graphs – these are difficult to read.
- Choose a clear heading which describes the purpose of the graph and the population.
- Mark the names of the variables and the units clearly.
- Choose scales carefully so as not to under represent or over represent differences in the data.
- Include both raw numbers and percentages: n (%).
- Always include 'n', being the number in your total population.
- If grouping data, be sure that the groups don't overlap and that the groupings are evenly spaced.
- When comparing data before and after or between groups, include **confidence intervals** where possible.

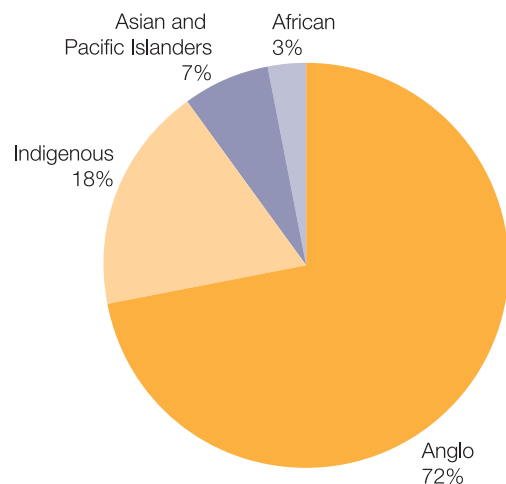
How to use pie charts

The pie chart is a popular and simple way of presenting data as it is easy to read and can quickly make a point. Pie charts are used to present categorical data and show how the percentage of individuals falls into various categories so that they may be compared.

Pie charts can only be used however when each individual in the group falls into one category and only one category as the sum of all the slices must be 100%. Thus pie charts are suitable for presenting data such as age ranges, cultural background, expenditure, types of surgery and cause of death, and are of limited value in quality improvement activities.

When presenting a pie chart always be sure to include the number in the total population, not just the percentages of the groupings. Also be sure that the percentages do not add up to more than 100%.

Figure 5.6 Pie chart showing breakdown of ethnic group in a sample population (n=210)



How to use bar graphs

Simple bar graphs are also used to present **categorical data**, where the groupings are discrete categories.

Bar graphs consist of a series of labeled horizontal or vertical bars with the bars representing the particular grouping or category. The height or length of the bar represents the number of units or observations in that category (also called the frequency).

As an example, Figure 5.7 shows how the percentage incidence of pressure ulcers varies between the various specialty areas. The categories are the specialty areas that are identified along the horizontal axis. The height of the columns corresponds to the prevalence of pressure ulcers reported on each of the areas.

Another way to look at the data is by using a Pareto chart. In a Pareto chart the purpose is to highlight the most important among a set of factors, thus:

- the values being plotted are arranged in descending order
- frequency of occurrence is presented on the left vertical axis (it can also represent cost or another unit of measure)
- the cumulative percentage of the total number of occurrences, total cost, or total of the particular unit of measure is shown in the right vertical axis.

In quality improvement, the Pareto chart is often used to show the most common sources or causes of quality problems. For our pressure ulcer example, the Pareto chart shows that while the prevalence of ulcers is highest in the spinal group, the numbers of patients affected in the general medical areas is highest, therefore intervention is required across all these areas.

In the process of further exploring the reasons for the pressure ulcer rates in these areas, a Pareto chart may also be used to identify and prioritise risk factors that need to be addressed through your intervention phase of the project.



DATA TIP:

Using the Pareto chart gives rise to the 80–20 rule – that 80% of the problems stem from 20% of the causes.

Table 5.5 Pressure ulcer prevalence across five specialty areas

	Spinal	Rehabilitation	Medical	Critical care	Surgical	Total
Number patients	23	1,101	3,053	483	1,645	6,305
Number patients with ulcers	14	264	569	72	232	1,151
Prevalence per specialty (%)	60.9%	23.9%	18.6%	14.9%	14.1%	18%
Total prevalence (%) n = 1,151	0.2%	23%	49.4%	6.3%	20.1%	100%

Figure 5.7 Bar graph showing prevalence of pressure ulcers per specialty (n=6,305)

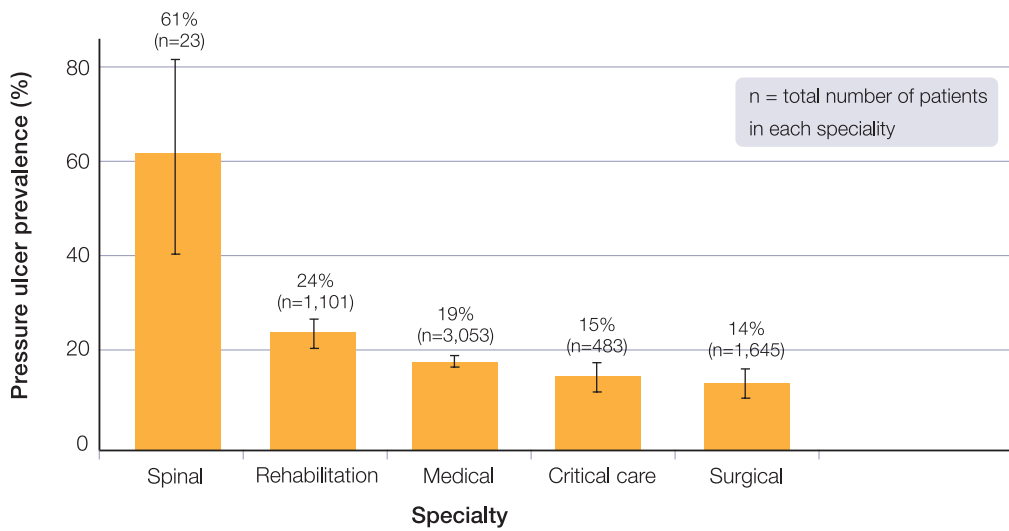
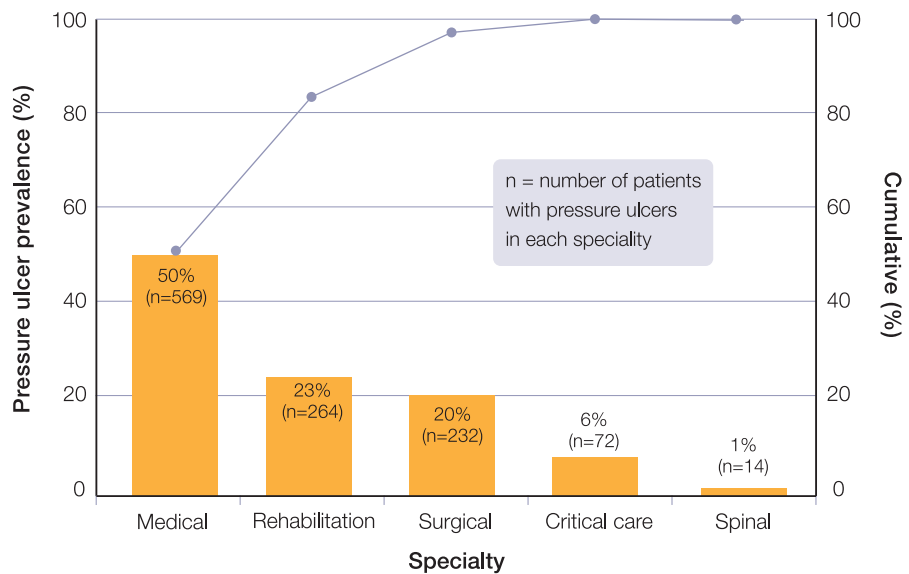


Figure 5.8 Pareto chart showing total prevalence of pressure ulcers (n=1,151)



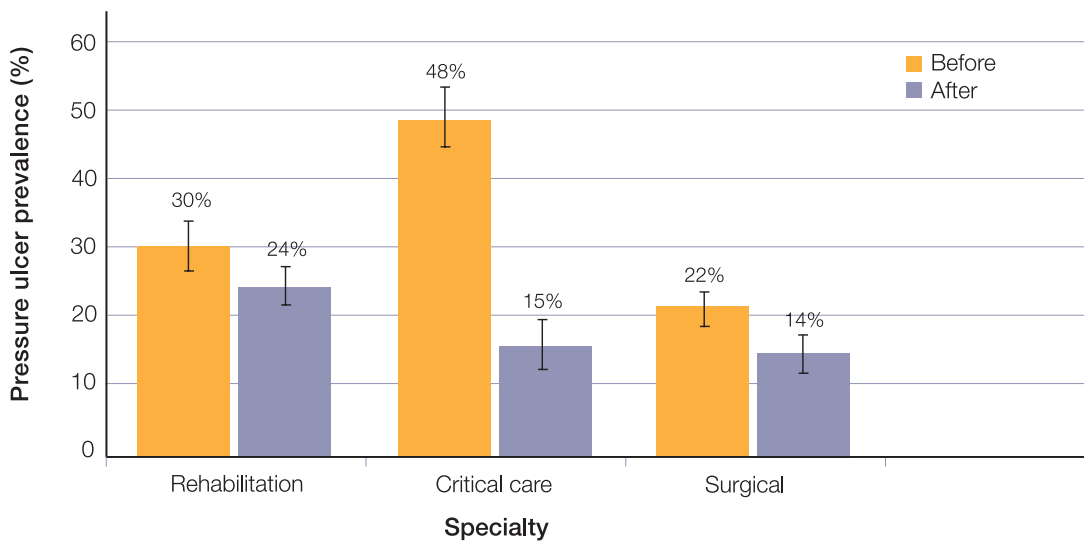
How to use bar charts to make comparisons

Bar charts can also be used to present comparative data, that is, to show changes that have occurred following your quality improvement intervention. One series of bars presents the situation before the intervention (baseline data) and the other, usually shown in a different colour, shows the situation after the intervention. The difference in the heights of the bars reflects the change that has (or has not occurred).

Figure 5.9 shows data before and after the introduction of the pressure ulcer prevention program in three specialty areas. The vertical lines are the confidence intervals.

Figure 5.9 Using bar charts to make comparisons before and after a quality intervention

Impact of a pressure ulcer prevention intervention



How to use box plots

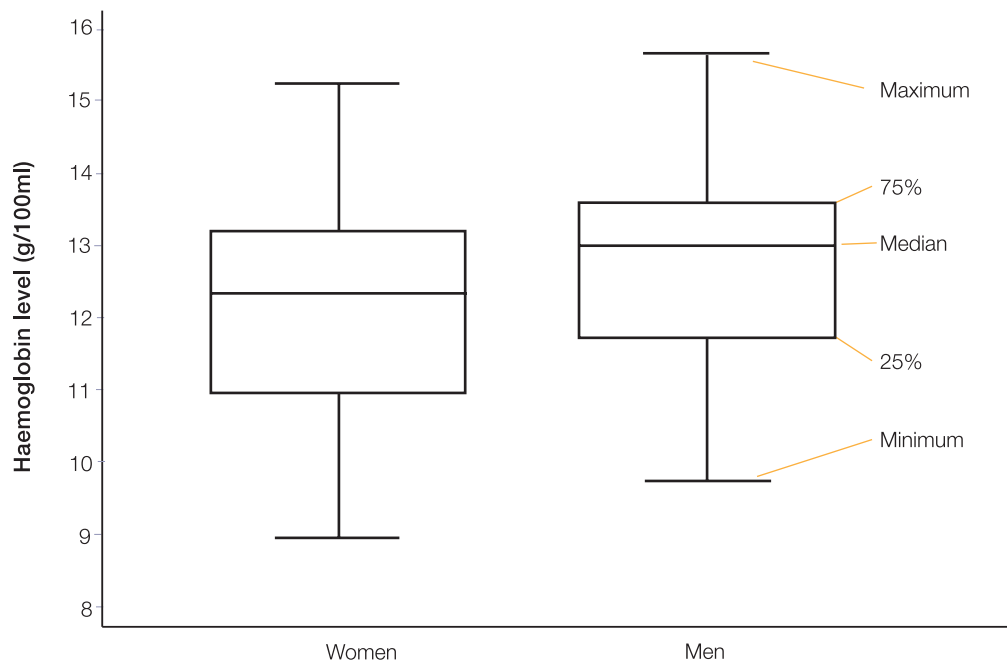
The box plot, also known as a box and whisker diagram, is a useful way of summarising and visualising your data to show:

- the median or 50th percentile (depicted by a line in the middle of the box)
- the lower quartile (that is, the value within which a quarter of values lie at the lower end of the scale)
- the upper quartile (that is, the value within which three quarters of values lie)
- the range of the data (minimum and maximum) depicted by the vertical lines extending from the box (the 'whiskers').

The box plot is particularly useful for comparing distributions between several groups of data (see Figure 5.10)

Figure 5.10 Sample box plot

Haemoglobin levels of women and men



How to use histograms and histograms

A **histogram** is a bar graph that is used to display **numerical data** as distinct from categorical data.

A **histograph**, or frequency polygon, is a graph formed by joining the midpoints of histogram column tops. These graphs are used only when depicting data for continuous variables shown on a histogram.

A histogram smoothes out the abrupt changes that may appear in a histogram, and is useful for demonstrating continuity of the data being studied.

You can use histograms/histographs to tell you three main features of your numerical data:

- how the data is distributed
- the amount of variability in the data
- where the centre of the data is (approximately).

Producing meaningful histograms takes a bit of practice. Be sure to:

- use an appropriate scale so that differences in frequencies are not played down or exaggerated
- make it clear whether you are using numbers or percentages to quantify frequency on the Y-axis
- choose appropriate ranges for your groupings along the X-axis, so as not to inappropriately represent the variability of the data.

Figure 5.11 shows a histogram and histograph representing waiting times for elective surgery. Note that the waiting time groupings are equally spaced (30 days) to adequately show the distribution of the data. These graphs might be used to help improve waiting list management.

Figure 5.11 Sample histogram and histograph

Histogram and histograph – Waiting time for elective surgery (n= 96)

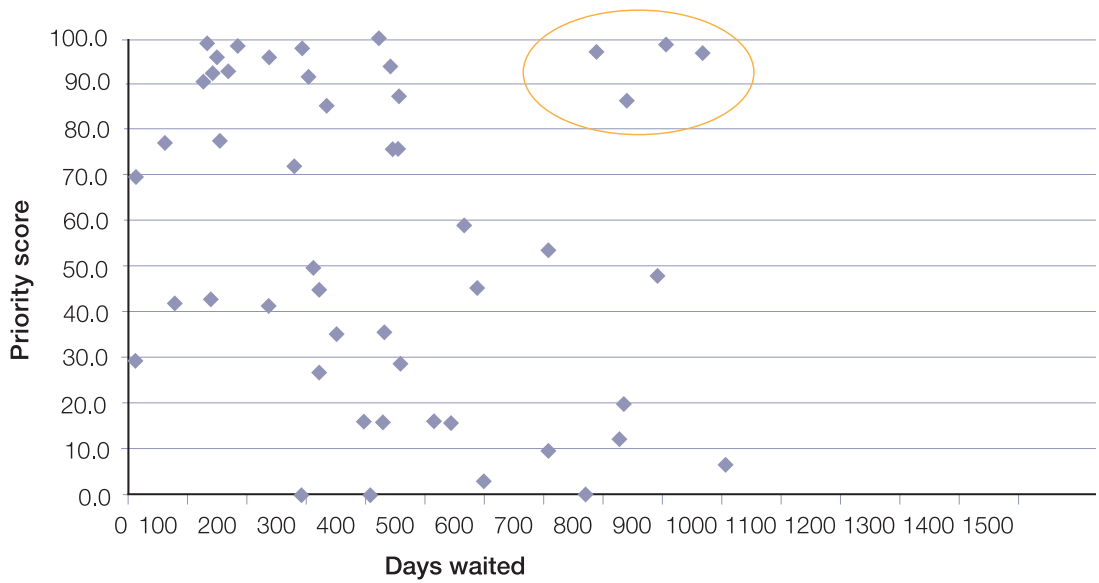


How to use a scatter diagram

In cases where both variables are numerical (quantitative), the data can be organised in a scatter diagram or scatter plot. This simple graph plots two characteristics of each observation. In the surgical waiting list example, a questionnaire might be used to calculate a score that reflects the need for surgery. For each patient, this score is plotted against the time that person has been waiting for surgery. Consideration of the scatter diagram helps to identify which patients should have surgery first, in the first instance those who have a high score and have waited a long time.

When interpreted in conjunction with Figure 5.11, this helps to make the most appropriate decisions about which patients should have surgery first, rather than just focusing on those who have waited the longest.

Figure 5.12 Scatter plot of patient priority score vs time waited for surgery (n=50)



Source: Orthopaedic Waiting List (OWL) Project 2007, *Phase II final report (unpublished)*.

How to use line graphs and time charts

A line graph is a visual depiction of how two variables are related to each other. It shows this information by drawing a continuous line between all the points on a grid. Line graphs compare two variables: one is plotted along the horizontal X-axis and the other along the vertical Y-axis.

The Y-axis in a line graph usually indicates quantity or percentage (such as infections, readmissions and adverse events), while the horizontal X-axis often measures units of time. As a result, the line graph is often viewed as a time series graph or time chart. At each time period, the amount is represented as a dot and the dots are connected to form the chart.

Line graphs can also depict multiple series, for example plotting results for different population subgroups or different services as shown in Figure 5.13.

Bar graphs and line graphs share a similar purpose. The bar graph, however, reveals a change in magnitude, whereas the line graph is used to show a change in direction. Line graphs are therefore used to reveal trends and relationships between data and to compare trends in different groups of a variable.

In quality improvement, time charts are useful for monitoring change over a period of time and for monitoring sustainability of change (see Figure 5.13).

Another version of a line graph is a **control chart** in which changes are demonstrated in relation to a specific target (see Figure 5.14). Control charts are time charts that track the consistency of data through time. They are often used to evaluate processes and to monitor quality performance, for example to monitor adverse events such as nosocomial infections.

Figure 5.14 shows a control chart for nosocomial infections over a four-year period. The chart features a centreline of the overall average of the data points and two lines representing the upper control limit and the lower control limit. The control limits help identify when significant variation in the process is occurring and thus whether intervention is required.

More information:

See the references in Appendix 3, *Useful resources*.

Figure 5.13 Line graph representing waiting times in the emergency department of two hospitals

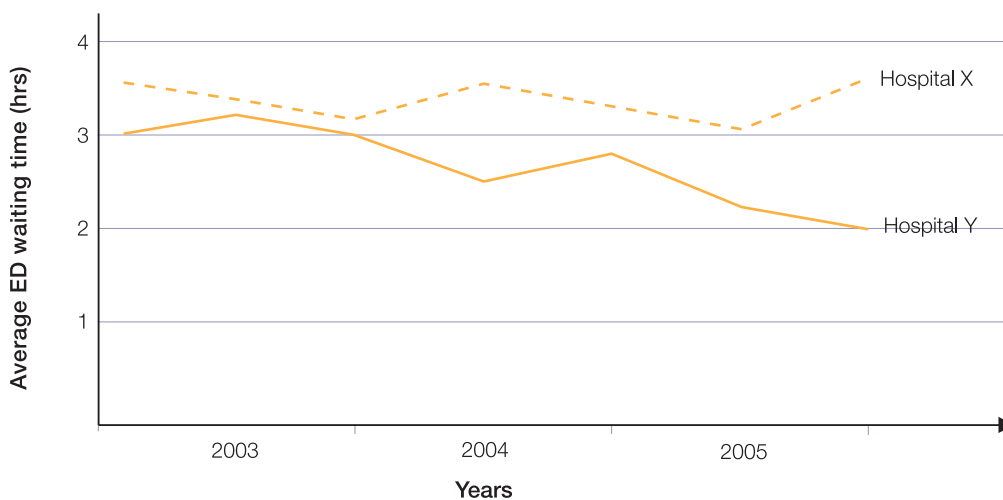
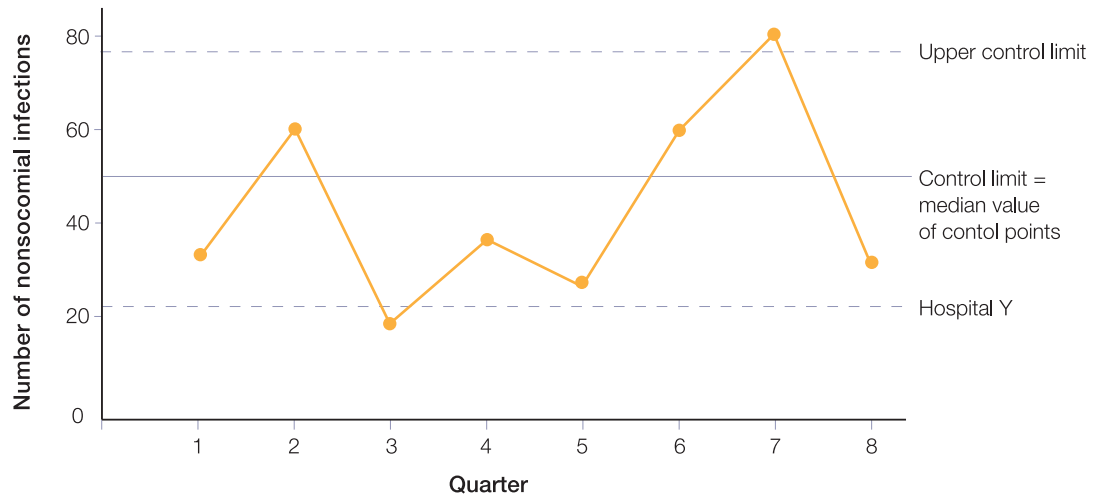



Figure 5.14 Control chart representing nosocomial infections in the emergency department



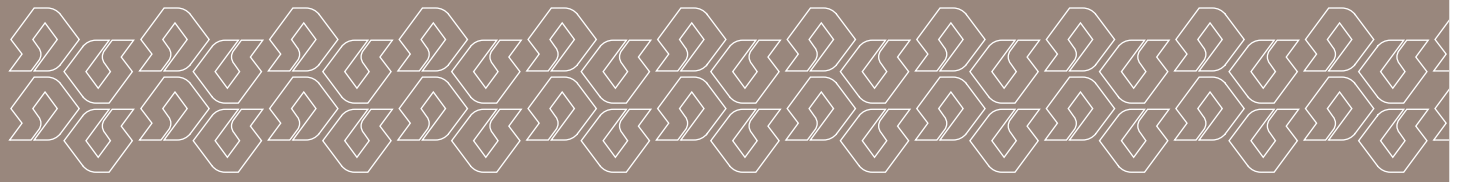
5.3 Analysing qualitative data

Qualitative data can also be analysed and provides an important input into your quality improvement initiatives. Analysis generally involves identifying themes and categories of data. More complex analysis may involve trying to identify relationships between themes. The process can be lengthy (and sometimes tedious); however, a number of software programs have been developed to assist in analysing quantitative data.

The more rigorously you design the project and analyse the data the more credible it will be. If you plan to submit the results of your project for publication using qualitative data it is advised you seek expert input before starting the project.

 **More information:**

See *Qualitative research in health care* by Catherine Pope, Nicholas Mays, Blackwell Publishing, 2006.



06

Interpreting and using data

Section 06

Interpreting and using data

So you've collected and analysed your data – now what? How do you come to conclusions about what is going on in your health service and how to link these conclusions to action?

Depending on your project, this can be a complex process, so this is when you should involve your stakeholders and seek input regarding the meaning of your data in the context of your organisation, including what your data means for design, implementation or refinement of your improvement initiative.

This should be undertaken at each phase of your quality improvement initiative, that is, at the project definition/diagnosis, intervention, impact and sustainability phases, so that your next steps are well informed and agreed upon.

Clear communication of your findings at each phase is therefore important, which starts with organising and presenting your data using some of the presentation

techniques described in the previous section. Your presentation should outline clearly:

- your original objectives
- a brief description of your data collection strategy, including consideration of your sample population, existing data sources, supporting literature and data collection tools
- a brief description of your analysis strategy
- tabulations, graphs and statistics that describe your findings.

This information can then be considered in light of what else is happening in the organisation as well as what is happening externally.

When looking to interpret your data and decide on appropriate actions, we return to the quality improvement cycle and the questions posed in section 2.

1. What is the problem or question?

What is your current performance and what are the performance gaps?

What are the needs and opinions of stakeholders?

What are your priorities in terms of problems and improvement projects?

What are your overall aims and targets for improvement?

2. What can you improve?

What are the processes and people involved in the processes?

What are the problem steps in the process?

What are the opportunities for improvement and which opportunities should you pursue?

What are your objectives for improvement?

What are the possible barriers and enablers to change?

3. How can you improve?

What are the most appropriate interventions to address your particular problem and to suit your situation/organisation?

What interventions and implementation strategies should you implement first?

4. Have you achieved improvement?

Have your interventions been implemented as planned?

What were the barriers to implementation?

Has the intervention achieved improvement in patient outcomes?

Is the improvement attributable to the intervention?

What else is happening within the organisation that has affected the intervention?

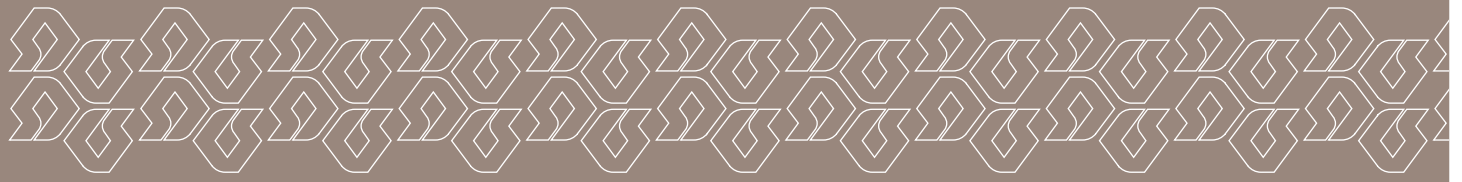
5. Have we sustained improvement?

What trends are occurring over time?

Is there a need for repeated intervention?

Have improved processes been integrated into routine practice?

What else do you need to do achieve quality improvement in this area?



07

Appendices

Section 07

Appendix 1. Data planning template

Planning template

Overall goal	Objectives/ improvement criteria	Population	Data requirements	Data collection methods	Data analysis	Data resource needs

Appendix 2. National/state databases and registries

General national databases

Australian Bureau of Statistics (ABS)
www.abs.gov.au

Australian Institute of Health and Wellbeing (AIHW)
www.aihw.gov.au

AIHW National Mortality Database
www.aihw.gov.au/mortality/mortality_database.cfm

Centre for Burden of Disease and Cost Effectiveness (BODCE)
www.uq.edu.au/bodce/

Centre for Research Excellence in Patient Safety (CRE-PS)
www.crepatientsafety.org.au

Chronic Disease Indicators Database
www.aihw.gov.au/cdi

Metadata Online Registry (AIHW)
<http://meteor.aihw.gov.au>

National Centre for Classification in Health (NCCH)
www3.fhs.usyd.edu.au/ncch

Public Health Information Development Unit (PHIDU)
www.publichealth.gov.au

General Victorian databases

Victorian Admitted Episodes Data Set (VAED)
www.health.vic.gov.au/hdss/vaed

Victorian Emergency Minimum Dataset (VEMD)
www.health.vic.gov.au/hdss/vemd

The Victorian Integrated Non-Admitted Health (VINAH) Minimum Dataset has been developed to incorporate the Community Rehabilitation Centre Minimum Dataset, the Specialist Clinics Minimum Dataset and HARP into a single dataset (VINAH MDS).
www.health.vic.gov.au/hdss/vinah/index.htm

VICNISS – Hospital Acquired Infection Surveillance System
www.vicniss.org.au

Bleeding disorders

Australian Bleeding Disorder Register
www.ahcdo.org.au/abdr

Haemostasis Registry
www.med.monash.edu.au/epidemiology/traumaepi/haemostasis.html

Cancer

Biogrid
www.biogrid.org.au

BreastScreen Australia Program
www.cancerscreening.gov.au/internet/screening/publishing.nsf/Content/breastscreen-1lp

Cancer Institute NSW
www.cancerinstitute.org.au/cancer_inst/statistics

Queensland Cancer Registry
www.qldcancer.com.au/research/qcr/qld_cancerReg.asp

South Australian Cancer Registry
www.health.sa.gov.au/pehs/branches/branch-cancer-registry.htm

Tasmanian Cancer Registry
http://menzies.utas.edu.au/re_cancer_reg.html

Victorian Cancer Registry
www.cancervic.org.au/cancer1/research/cec/vcr.htm

Victorian Cervical Cytology Register
www.vccr.org

Western Australian Cancer Registry
www.health.wa.gov.au/wacr

Western Australian Cervical Cytology Registry
www.kemh.health.wa.gov.au/services/cervical_cancer/index.htm

Cardio/respiratory

Australian Society of Cardiac and Thoracic Surgeons Database Project
www.ascts.org/sections/outcomes/index.html

Australian Centre for Asthma Monitoring (ACAM)
www.asthmonitoring.org

Cardiac Surgery in Victorian Public Hospitals
www.health.vic.gov.au/specialtysurgery/cardiac.htm

Melbourne Interventional Group Interventional
 Cardiology Register
www.ccretherapeutics.org.au/research/mig/index.html

Melbourne Vascular Surgery Quality Initiative
www.health.vic.gov.au/specialtysurgery/vascular.htm

Dental

Dental Statistics and Research Unit (DSRU)
www.arcpho.adelaide.edu.au

Diabetes

National Diabetes Register
www.aihw.gov.au/diabetes/ndr.cfm

Elective surgery

Elective Surgery Information System
www.health.vic.gov.au/hdss/esis

General practice

Australian General Practice Statistics and
 Classification Centre (AGPSCC)
www.fmrc.org.au/agpscc

Injury and trauma

CONROD Trauma Registry (QLD)
www.uq.edu.au/conrod/index.html?page=11386

Monash University Accident Research Centre
 (MUARC)
www.monash.edu.au/muarc

National Injury Surveillance Unit (NISU)
www.nisu.flinders.edu.au

Trauma Registry, Royal Perth Hospital
www.rph.wa.gov.au/traumaservices.html

Victorian Orthopaedic Trauma Registry
www.med.monash.edu.au/epidemiology/traumaepi/orthoreg.html

Victorian State Trauma Outcomes Registry (VSTORM)
www.med.monash.edu.au/epidemiology/traumaepi/traumareg.html

Intensive care

Australian and New Zealand Intensive Care Unit
 Society Adult Patient Database
www.anzics.com.au/section.asp?Section=adult

Mental health

Mental Health Client Management Interface and
 Operational Data Store (VIC)
www.health.vic.gov.au/healthsmart/mental.htm

Musculoskeletal

Australian Orthopaedic Association Joint
 Replacement Register
www.aoa.org.au/jointregistry.asp

Australian Rheumatology Association Database
www.rheumatology.org.au/rheumatologists/aradatabase.htm

Reproductive

National Perinatal Statistics Unit (NPSU)
www.npsu.unsw.edu.au

Transplant

Australian and New Zealand Cardiothoracic Organ
 Transplantation Registry
www.anzcotr.org.au

Australian and New Zealand Dialysis and
 Transplantation Register
www.anzdata.org.au/anzdata/anzdatawelcome.htm

Australian and New Zealand Liver Transplant Registry
www.anzltr.org

Australian Bone Marrow Transplantation
 Recipient Register
www.bmntsw.com.au/abmtrr_about.php

Australian Corneal Graft Registry
http://som.flinders.edu.au/FUSA/ophthalmology/Aust_registry.htm

Australian Organ Donor Register
www.medicareaustralia.gov.au/provider/patients/aodr/index.shtml

Other disease-specific databases and registries

Australian Cystic Fibrosis Register
www.cysticfibrosis.org.au/projects/dataregistry/

Australian Motor Neuron Disease Registry
www.amndr.org

Australian National Creutzfeldt-Jakob
 Disease Register
<http://ancjdr.path.unimelb.edu.au/>

Bosentan Patient Registry
www.bosentanregistry.com.au

Appendix 3. Useful resources

More on statistics

The following resources provide further detail on statistical processes without too much complex jargon.

- *Medical statistics*, Kirkwood BR, Sterne JAC, Blackwell Science Publishing 2003
- *Statistics Notes in the British Medical Journal* 2005. Available from: www-users.york.ac.uk/~mb55/pubs/pbstnote.htm
- *Epidemiology for the uninitiated*, Coggon D, Rose G, Barker D. BMJ Publishing Group 1997. Available from: www.bmj.com/collections/epidem/epid.dtl.
- *Statistics at square one*, Swinscow T. BMJ Publishing Group; 1997. Available from: www.bmj.com/collections/statsbk/index.dtl
- *Statistics for dummies*, Deborah Rumsey, Wiley Publishing 2003.
- *Improvement Leaders' Guide: Measurement for Improvement*. NHS Institute for Innovation and Improvement, UK 2005. Available from: www.nodelaysachiever.nhs.uk/NR/rdonlyres/1EC0FD98-2BEE-43E6-A708-AFAE925B5F89/0/RG0073ImprovementLeadersGuideMeasurementforImprovement.pdf
- *The UCLA Center for Health Policy Research* Available from: www.healthpolicy.ucla.edu/HealthData/links.html

More on process mapping

- *Improvement Leaders' Guide: Process mapping, analysis and redesign* NHS Institute for Innovation and Improvement, UK 2005. Available from: www.institute.nhs.uk/images/documents/ILG/general/1_2PM_Nov05.pdf
- *Process mapping: a guide for health service staff* 2007, The Victorian Quality Council. Available from: www.health.vic.gov.au/qualitycouncil/downloads/process_mapping.pdf

More on measurement and quality improvement – general

- *Quality management in healthcare - principles and methods*. Lighter D, Fair D. 2nd ed, 2004, Jones and Bartlett Publishers, Sudbury.
- *Healthcare and Quality*. Institute of Medicine of the National Academies. Available from: www.iom.edu/?id=19174
- *Easy Guide to Clinical Practice Improvement*. NSW Health Department 2002. Available from: www.health.nsw.gov.au/quality/pdf/cpi_easyguide.pdf
- *Measurement for improvement toolkit*. Australian Commission on Safety and Quality in Medical Care, Commonwealth of Australia 2006. Available from: www.safetyandquality.org/internet/safety/publishing.nsf/Content/CommissionPubs
- *NHS The Improvement Network, East Midlands*. Available from: www.tin.nhs.uk/tools--techniques/implementation/tools-to-use-here
Includes more information on some of the improvement tools described in this guide.

More on surveys and questionnaires

- *Basics of developing questionnaires*. Available from: www.managementhelp.org/evaluatn/questnrs.htm
This site is a great place for beginners to start. The page covers all the basics to be aware of when starting to design a questionnaire.
- *Survey design* Available from: www.surveysystem.com/sdesign.htm
Another good introductory site. Great information on comparing the various methods of conducting a survey.
- *Guide to questionnaires and surveys* Available from: www.members.tripod.com/~frede_dast/conseil1_a.html
Contains many articles on a variety of topics around the creation of good surveys and questionnaires. This site's information is a bit more technical and in-depth, but is still understandable to the novice.

More on clinical audit

- *Clinical Audit – How to guides*. NHS United Bristol Healthcare. Available from: www.ubht.nhs.uk/healthcare-professionals/clinical-audit/how-to-guides.html.

More on qualitative methods

- *Qualitative research in healthcare* Catherine Pope, Nicholas Mays, Blackwell Publishing 2006

More on ethics and privacy

- *When does quality assurance in health care require independent ethical review?* Canberra: National Health & Medical Research Council; 2003. Available from: www.nhmrc.gov.au/publications/synopses/e46syn.htm.
- *State and Territory Privacy Laws*
Available from: www.privacy.gov.au/privacy_rights/laws/index.html#2
- *Guidelines to the National Privacy principles*
Available from: www.privacy.gov.au/publications/nppgl_01.html

More on presenting data

- *How to display data* Freeman, Walters and Campbell Blackwell Publishing 2008

Appendix 4. Glossary

While you don't have to be a statistician to use data effectively in quality improvement, you do need to become familiar with the basic terminology. This section of the guide outlines some common terms that you are likely to come across in your quality improvement work.

adverse event: an unintended injury or complication that results in disability, death or prolonged hospital stay, and is caused by health care management rather than the patient's disease.

affinity diagram: a tool that gathers large amounts of data (ideas, opinions, issues) and organises them into groupings based on their relationships. The affinity process is often used to group ideas generated by brainstorming.

audit: the measurement of care processes in relation to predetermined standards or guidelines.

bar graph: a diagram that compares bars of the same width but different heights according to the statistics or data they represent.

benchmarking: the practice of setting operating targets for a particular function by selecting the top performance levels, either within or outside a company's own industry. In a broader sense, benchmarking involves searching for and copying new ideas and best practices to improve processes, products and services.

bias: systematic favouritism present in the data collection process that results in lopsided, misleading results.

brainstorming: a technique to generate as many thoughts and ideas as possible within a defined time.

categorical variable: a set of data is said to be categorical if the values or observations belonging to it can be sorted according to category. Each value is chosen from a set of non-overlapping categories.

cause: the reason for a problem or defect.

cause and effect diagram: a tool for discovering all the possible causes for a particular effect.

census: collection of data from a whole population (as distinct from a sample).

common cause: a reason for a problem or defect that is inherent in the production process.

concept mapping: a technique for visualising the relationships among different concepts. Concepts are connected with labelled arrows, in a downward-branching hierarchical structure. The relationship between concepts is articulated in linking phrases, such as 'gives rise to', 'results in', 'is required by,' or 'contributes to'.

confidence interval: an estimate using a range of values (an interval) to predict the expected value of an unknown parameter, accompanied by a specific level of confidence, or probability, that the estimate will be correct.

continuous variable: a continuous variable is a numeric variable that can assume an infinite number of real values. For example, height, weight, age, distance, time and temperature.

control chart: a chart that includes a measure of central tendency such as the mean, and a measure of variability, such as the standard deviation, that provides information about the performance of a process and the presence of common cause or special cause variability.

correlation coefficient: a measure showing to what extent two variables vary in an interdependent way.

counts: a count of how many items or observations you have in your sample, for example the number of people receiving a particular treatment, or the number of people responding to a survey.

data: facts or figures from which conclusions can be drawn.

data dictionary: a document containing variable definitions.

dataset: the collection of data taken from your sample.

discrete variable: a discrete variable is one that cannot take on all values within the limits of the variable. For example, responses to a five-point rating scale can only take on the values 1, 2, 3, 4 and 5. The variable cannot have the value 1.7.

effectiveness: conformity to requirements; the degree to which the service is performed in the correct and desired manner.

focus group: an interviewing technique whereby respondents are interviewed in a group setting. It is used to stimulate the respondents to talk freely, encourage the free expression of ideas or explore attitudes and feelings about a subject.

frequency: the number of times an event or item occurs in a dataset.

histogram: a graphical version of a table that shows what proportion of cases fall into each of several or many specified categories.

histograph: a graph formed by joining the midpoints of histogram column tops. These graphs are used only when depicting data from the continuous variables shown on a histogram.

human factors: a discipline of study that deals with human-machine interface – including the psychological, social, physical, biological and safety characteristics of a user and the system the user is in.

incidence: the frequency with which something, such as a disease, appears in a particular population or area.

incident monitoring: capturing errors and near misses (errors that are detected before they reach the patient) for the purposes of quality improvement.

line graph: a visual depiction of how two variables are related to each other. It shows this information by drawing a continuous line between all the points on a grid.

linkage: interactions that effect coordination and completion of tasks.

margin of error: measure of the maximum amount by which sample results are expected to differ from the population.

mean (average): a measure of centre. The sum of all numbers divided by the total number of numbers.

median: a measure of centre. The point at which an equal number of data points lie above and below.

mind map: a diagram used to represent words, ideas, tasks or other items linked to and arranged radially around a central key word or idea. It is used to generate, visualise, structure and classify ideas, and as an aid in study, problem solving and decision making.

mode: a measure of centre. The mode is the value with the largest frequency.

normal distribution: a family of distributions that have the same general shape. They are symmetric with scores more concentrated in the middle than in the tails. Normal distributions are sometimes described as bell shaped.

numerical variable: numerical variables are expressed as numbers. These include discrete variables that can take on only a finite number of values (such as number of medications, number of services provided). See also continuous variable, discrete variable.

objectives: measurable statements that are consistent with the mission, vision and key drivers.

parameter: a number that summarises data collected from a whole population.

Pareto chart: a type of bar chart, prioritised in descending order from left to right, used to identify the vital few opportunities for improvement.

percentile: the percentage of values in the dataset that fall below a certain score.

performance measure: a measure that tells you how well your process is achieving its purpose.

PERT chart: the Program (or Project) Evaluation and Review Technique is a method to analyse the tasks involved in completing a given project, especially the time needed to complete each task, and identifying the minimum time needed to complete the total project.

population: the group of individuals relevant to your quality improvement project or study.

prevalence: the proportion of individuals in a population having a disease.

problem: the result of non-conformance to patient and other stakeholder requirements.

process management: improvement of work activities and work flow across functional or department boundaries.

process mapping: a way of collecting data and information about current processes that involves outlining and analysing the steps in your existing processes in order to confirm what is currently happening, identify problems within the process, and establish the causes of these problems.

qualitative data: includes virtually any information that can be captured that is not numerical in nature, such as information from focus groups, interviews or direct observation.

quality: the extent to which products and services meet or exceed customer requirements.

quality assurance (QA): retrospective review or inspection of services or processes that is intended to identify problems.

quality care: the extent to which health care services meet the patient's needs and produce the desired health outcome.

quality improvement: the continuous study and improvement of a process, system or organisation.

quality indicators: characteristics of products, services or processes that represent quality.

quantitative data: data measured or identified on a numerical scale. Quantitative or numerical data can be analysed using statistical methods, and results can be displayed using tables, charts, histograms and graphs.

range: a commonly used measure of variability, defined by the minimum or lowest observation and the maximum or highest observation.

rate: a ratio that describes one quantity in relation to a certain unit. For example the rate of hospital readmissions may be expressed as 4 per 100 people discharged.

ratio: a fraction, expressed in its simplest terms, that describes two groups relative to one another.

reliability: the ability to get repeatable results with subsequent measurements.

sample: a number selected from your population for study.

scatter diagram/scatter plot: a tool for analysing relationships between two variables. One variable is plotted on the horizontal axis and the other is plotted on the vertical axis. The pattern of their intersecting points can graphically show relationship patterns.

sentinel event: a sentinel event is a relatively infrequent, clear-cut event that occurs independently of a patient's condition; it commonly reflects hospital system and process deficiencies, and results in unnecessary outcomes for the patient (The Department of Human Services, Victoria).

standard deviation: a measure of variability of data. A measure of the average distance from the mean.

standard score: the number of standard deviations above or below the mean.

statistic: a number that summarises the data collected from the sample such as percentage, average, median, percentile.

survey: the collection of information about characteristics of interest from some or all units of a population, using well-defined concepts, methods and procedures, and the compilation of such information into a useful summary form.

validity: valid data or valid data collection tools measure what it is intended to be measured.

variable: variable is another term used in relation to measurement. Height, age, gender, amount of income, country of birth, language, diagnosis are all examples of variables. Variables may be classified in a number of ways – see categorical, numerical, continuous and discrete variable definitions.

Appendix 5. References

1. Australian Council for Safety and Quality in Health Care 2001, in Victorian Quality Council 2003, *Better Quality, Better Healthcare: A safety and quality framework for Victorian health services*. Department of Human Services. Available from: www.health.vic.gov.au/qualitycouncil/pub/improve/framework.htm [cited 2008]
2. National Health Performance Committee 2001, *National Health Performance Framework Report*. Available from: www.health.qld.gov.au/nathlthrpt/performance_framework/11381_doc.pdf [cited 2008]
3. Health Education and Management Innovations 2007, *Prevention of venous thromboembolism: best practice guidelines for Australia and New Zealand 4th ed*, Baulkham Hills.
4. Kirkwood BR, Sterne JAC 2003, *Medical statistics*, Blackwell Science Publishing
5. Statistics Notes in the British Medical Journal 2005, Available from: www-users.york.ac.uk/~mb55/pubs/pbstnote.htm [cited 2008]
6. Coggon D, Rose G, Barker D 1997, *Epidemiology for the uninitiated*. BMJ Publishing Group, Available from: www.bmj.com/collections/epidem/epid.dtl [cited 2008]
7. Lighter D, Fair D 2004, *Quality management in healthcare – principles and methods, 2nd ed.*, Jones and Bartlett Publishers, Sudbury.
8. Swinscow T1997 *Statistics at square one*. BMJ Publishing Group, Available from: www.bmj.com/collections/statsbk/index.dtl [cited 2008]
9. Streiner D, Norman G 2001, *Health measurement scales – a practical guide to their development and use, 2nd edition*, Oxford University Press.
10. Rumsey, D 2003, *Statistics for dummies*, Wiley Publishing.
11. National Emergency Medical Service for Children Data Analysis Resource Center, Available from: www.nedarc.org/ [cited 2008]
12. Pope C 2000, 'Mays Qualitative Research in Healthcare', *BMJ*, Jan 8; 320(7227):114–116, Available from: www.ncbi.nlm.nih.gov/pubmed/10625273?dopt=Abstract&holding=f1000,f1000m,isrctn [cited 2008].

