Communicating organisational outcomes using simple performance indicators: a case study using ACT elective surgery waiting lists

Catherine Elizabeth Rayner
October 2016

A thesis submitted for the degree of Doctor of Philosophy

of The Australian National University

I certify that this thesis is entirely my own original work.

    ————————————————————

Acknowledgements

As with any undertaking of this scope, there are many people whose patience and support have been vital to its completion. Right at the top of the list are my beloved family, Geoff and Alex. They have listened to me obsess about performance indicators at the dinner table for the last four years, sometimes asking intelligent questions and sometimes just letting me rant. The other family member who has been a close part of this is my father, John. As a visiting fellow in the centre where I was working, we shared many lunches and talks about both his research and mine.

My supervisory panel are a trio of entertaining individuals with whom it has been a pleasure to exchange ideas and refine arguments. As panel chair, Rod Lamberts has been terrific – supportive, encouraging and tolerant of all the interesting things that caught my attention (and distracted me) along the way. I am particularly grateful for his willingness to trust me to do the research as I saw fit and his acknowledgement of the pre-existing professional skills I brought to the project.

Abstract

This research explores the communication consequences of using simple, numerical information to convey information about the performance of large, public-sector organisations. To control the scope of the research, both practically and theoretically, the case study is based on publically available material, in other words, the material accessible by a lay, rather than a specialist audience. The formal publication and public reporting in the local newspaper of elective surgery waiting times in the Australian Capital Territory (ACT) public health system during 2010 and 2011 is used as a detailed case study.

The theoretical basis for the research is drawn from three main disciplinary streams: performance indicators and evaluation; public health and public health policy; and communicating about and with numerical information. Each of these broad areas is itself multi-disciplinary, with research findings published across many different outlets.

When I began this research in mid-2011 there was little overlap between the insights from the different discipline groups canvassed in this literature review. In particular, the application of communication theories to performance indicators had yet to be systematically explored. Research in the period 2011-2016 clarifies issues within each of the broad areas but there is still very little synthesis between them.

The overall goal of the research is twofold: to characterise and systematically describe the way numerical performance indicators are used; and to develop a framework for determining how difficult to understand a numerical performance indicator is likely to be. The resulting model for analysing the communication effectiveness of using different types of statistical constructs as performance indicators can be used both as a tool for designing performance indicators and as a tool for analysing communication problems related to numerical performance indicators.

Glossary of terms and abbreviations

ACT	The Australian Capital Territory, a self-governing territory of Australia and the seat of the national government.
AIHW	Australian Institute of Health and Welfare, a statutory body providing health and welfare information and statistics.
CGC	Commonwealth Grants Commission, a statutory body tasked with dividing Federally-collected revenue between the Australian States and Territories.
COAG	Council of Australian Governments
Contemporary History	Contemporary history uses a historical frame of reference to understand the world as it exists today. In the widest context of its use it is that part of history which is still within living memory.
GST	Goods and Services Tax. A Federally-collected point of sale tax which is a major revenue source for state and territory governments and distributed by the Commonwealth Grants Commission.
ICD-10	International Classification of Diseases version 10
Medicare	Australian Commonwealth-funded health insurance scheme providing universal free or subsidised healthcare.
NHDD	National Health Data Dictionary
NHRA	National Health Reform Agenda
NMDS	National Minimum Data Sets
NSW	New South Wales, state of Australia
NT	Northern Territory, self-governing territory of Australia
PIAAC	Programme for the International Assessment of Adult Competencies
Public Health System	The parts of Australia's health system that are publically funded. This includes all public hospitals and those

procedures carried out in private hospitals which are publically funded.

QALY   Quality Adjusted Life Years. A measure of the effectiveness of a health intervention.

QLD   Queensland, state of Australia

SA   South Australia, state of Australia

TAS   Tasmania, state of Australia

VIC   Victoria, state of Australia

WHO   World Health Organisation

Table of Contents

Table of Tables

Table of Figures

1 Introduction

This research explores the communication issues involved in using simple numbers, presented as performance indicators, to measure the performance of complex organisations. While it focuses on outcomes in the public policy domain, it is a communication study using tools and insights from a range of disciplines rather than conventional public policy research.

To control the scope of the research, both practically and theoretically, the case study is based on publically available material. This also helps to keep the focus on the taxpayer/citizen audience. The formal publication and public reporting in the local newspaper of elective surgery waiting times in the Australian Capital Territory (ACT) public health system during 2010 and 2011 is used as a detailed case study.

The overall goal of the research is twofold: to characterise and systematically describe the way numerical performance indicators are used; and to develop a framework for easily determining how difficult it is to understand a use of a performance indicator.

1.1 Rationale for the study

The original motivation for this research arose from my professional experience as a writer and data analyst. Over the past 20 years I have worked in a variety of settings providing analyses and interpretations of performance data. In each setting, the client kept asking for 'simple' or 'straightforward' answers to questions which became less and less simple the more we looked into them. The complexities of the systems and the limitations of the measurement data it was possible to collect had complicating effects on performance reporting in domains as diverse as military working hours and university research performance. As both an analyst and a communicator, I became fascinated with this tension between complex systems, data limitations and the frequently-expressed need for simple answers. As time passed I began to wonder if it could ever be possible to have performance indictors for complex activities that were meaningful, possible to collect and easy to communicate. In this research I wanted to use a formal, academic framework to investigate my anecdotal and experiential conclusions.

My overall goals were to develop some analytical tools for describing the communication aspects of performance indicator use and, if possible, to recommend some ways of mitigating communication problems with performance indicators.

There is a wide body of research about the development and use of performance indicators, much of it focussed on specific indicators for specific purposes. There is also research going back to the 1990s on how (and how not) to develop and use performance indicators as well as material on the limitations of even the best indicators.

The research relating to the shortcomings of performance indicators falls into two main categories:

- the question of how validly, in the statistical sense the performance indicators characterise the organisation being measured; and
- the side-effects and unintended consequences of introducing particular performance metrics into an organisation.

There is also a huge body of work on the communication of statistics and how humans deal with numerical information. While content-specific research is published in a wide range of disciplines, more general research is concentrated in the fields of risk communication, developmental anthropology and science communication.

The core of my research lies at the intersection of these two broad domains specifically in the context of public policy and government accountability. It seeks to address two questions:

1. What makes performance indicators hard to understand?
2. Is there anything we can do to make them easier to understand?

Answering these questions has the potential to improve the usefulness of public policy performance indicators, that is indicators relating to services provided to the community using taxpayer funds. One of the key audiences for this information is the taxpayer or citizen as both as user of the services and as someone who pays for them to be provided to the community. As described by Bird, Cox et al (Bird, Cox et al. 2005) one of the three core reasons for public sector performance measurement is that of accountability for the use of public resources.

By extending our understanding of the messages people receive from numerical performance indicators we should be able to change the ways in which they are presented and used to make the intended meaning more accessible to a wider audience.

1.2 Breaking down the problem

There are two, inter-related parts to communicating about the performance of something using numerical indicators. The first is communicating the relationship between the statistic being reported and the performance of the entity being measured. Without an agreed and understood definition of this relationship, there is no way for an audience to make sense of the information. The second part is the issue of understanding the statistics themselves, all the way from "what is a median" to what does the phrase "days waiting for elective surgery" really mean? In addition, there are the well-documented more basic issues of how people relate to and understand numerical information in general (Tversky and Kahneman 1974; Dunbar 1996; Gigerenzer and Edwards 2003; Gigerenzer, Gaissmaier et al. 2008; Peters, Hibbard et al. 2010).

1.3 Selection of the case study

In selecting a case study for this research, I wanted something that would illustrate and complement the theoretical aspects of the work as well as providing scope for detailed analysis of performance indicator communication. The specific characteristics for an appropriate case study were:

- It had to involve monitoring and, more importantly, reporting on the performance of a complex public institution or system of institutions;
- The institution or system needed to be Australian-based and the reporting done in English;
- The performance indicators used needed to be presented as fairly simple numbers;
- The indicators needed to be published regularly and available to the general public;
- The information about the background, development and implementation of the performance indicators needed to be available to the public; and
- It had to be of enough public interest for the performance indicators to be reported on and discussed in the news media and public forums.

The emphasis on public availability of background information as well as the performance indicators themselves has two purposes: it contains the scope of the research and focuses it on the issue of communicating with a broad, public audience

rather than with specialist groups. As an audience, 'the public' is a complex mix of different groups with different interests and there is little communication tailored to these smaller groups such as users of the public service being provided or workers within the system being measured. In some cases a sub-group of 'the public', such as surgery patients in the public hospital system, may receive information tailored to their circumstances as well as the information available to the general public. Based on my professional experience, it was likely that the two sets of information would not be the same.

I therefore decided to concentrate on information and materials readily available in the public domain. These were: official reporting; media reporting; reports by interest groups which use or respond to the official reporting; how much extra a determined lay person can find out and media reports based on the above sources.

In the end there were three possible candidates for a case study:

- NAPLAN results: The National Assessment Program – Literacy and Numeracy (NAPLAN) is an annual assessment for students in Years 3, 5, 7 and 9. It is used both to track individual student's progress and more broadly in the allocation of public funding to schools. It is also used outside its intended purposes to create league tables and comparisons between schools.
- ERA results: The Excellence in Research for Australia (ERA) process evaluates the quality of the research undertaken in Australian universities against national and international benchmarks.
- Public Hospital Elective Surgery Waiting Times: Elective surgery waiting times are widely used as one of the performance indicators for Australia's public hospital system.

The main reason I chose the elective surgery waiting times for my case study was that I was new to the content and so would be in the position of a determined lay person when I was looking for publically available information. In mid-2011, when I was looking for a case study, elective surgery waiting times were highly visible in the news media as new public hospital finding arrangements were being negotiated between the Australian Federal government and the governments of the States and Territories. To keep the range of the research manageable, I focussed primarily on the waiting lists for The Australian Capital Territory.

1.4 Elective surgery waiting times case study

Elective surgery waiting times serve as just one performance indicator for the Australian public hospital system. They are used for a range of purposes from high-level (Federal) allocation of certain types of health funding to day-to-day management of surgery at both the state/territory level and the individual hospital level. They are also widely used in lobbying for more public hospital resources, arguing that there is waste and inefficiency in the public hospital system, and to defend the performance of the system.

The major elements of the case study were:

- Documenting the detailed context, formal processes and official communication channels relating to elective surgery in the Australian public hospital system in general and the ACT in particular for the period 2010-2011.
- Creating a timeline of key policy and funding announcements for the period 2010-2011.
- Collecting reports relating to those processes, communications and announcements in the local ACT newspaper.
- Analysing the newspaper reports in the following ways:
  - o Descriptive statistics for the collection as a whole including looking for matches between number of relevant articles and key policy announcements;
  - o Article by article thematic analysis looking for both explicit and implicit content;
  - o Identifying each instance in the articles of elective surgery waiting time being used as a performance indicator for the public hospital system; and
  - o Instance by instance thematic analysis focussing on use of statistics and value judgements.

The first steps in understanding how elective surgery waiting times are used as performance indicators were to document the official contexts in which the indicators are used as well as the details of their collection, calculation and publication. This involved documenting the political and funding context in which the Australian

public hospital system operates, including where and how the waiting list data is used within this context. It also required documenting the process of data collection, from formal data definitions, through the overall administrative processes for elective surgery in the ACT public hospital system, to the first official public reporting of the indicators.

Once the official data is published, it is re-reported, interpreted and commented on in a range of forums. For the case study I focussed on articles from *The Canberra Times*, the ACT local newspaper, over the period 1 January 2010 to 31 December 2011. The time period was one of relative stability in policy direction with a strong focus on the funding and performance of the public hospital system.

1.5 Research Methods

Science Communication as a field of study makes use of a diverse range of techniques, primarily from the social sciences and the humanities as well as applied fields such as education, marketing and, of course, communication. *Chapter 3 - Political, Funding and Contextual Background to the Australian Public Hospital System* relies heavily on techniques from contemporary history, such as those described in *Contemporary History: Practice and Method* (Seldon 1988) and *The Contemporary History Handbook* (Brivati, Buxton et al. 1996).

The next part of the research concentrates on reporting and re-reporting of elective surgery performance (see *Chapter 4 - Newspaper article analysis*) It primarily uses thematic text analysis as described in Boyatis's *Transforming Qualitative Data: Thematic Analysis and Code Development* (Boyatzis 1998) and techniques relating to the analysis of graphical information (Tufte 2001).

There are many ways to extract meaning from a collection of texts, ranging from analysing each one as a separate entity to looking for large-scale patterns or themes within the entire collection. The techniques of thematic analysis offer a practical system for classifying and coding collections of texts. As Boyatzis points out in the preface to *Transforming Qualitative Information* "[thematic analysis is] something to be used to assist the researcher in the search for insight" (Boyatzis 1998). Part of the utility of thematic analysis is that it has the capacity to identify individual articles for more detailed analysis as well as revealing patterns and themes across a collection of texts.

At its core, thematic analysis involves developing a range of codes that can be applied to the collection, either to each text as a whole or to words and phrases within the texts. The codes can vary in complexity from simple quantitative descriptions (how long is the text, when and where was it published), through closed text-based questions (does the article mention elective surgery?) to highly conceptual questions (is the article concerned with the concepts of fairness and equity?). The skill in performing a thematic analysis lies in making the codes as unambiguous as possible so that different coders get the same results.

There are several different approaches to code development for thematic analysis each of which depends on how far in advance of coding the researcher wants to tightly specify their codes. If the coding exercise is part of an established, comparative program of research, the codes are likely to have been developed and standardised well in advance of their application. In cases where the exercise is primarily exploratory a data-driven, inductive approach to code development is more appropriate (Boyatzis 1998). This latter approach, sometimes referred to as coding for emergent themes involves starting with a broadly indicative set of codes derived from an initial reading of the texts which are then iteratively refined and added to with each detailed pass through the texts. The latter process was the one I used for analysing the material from *The Canberra Times*.

Coding for emergent themes is an iterative process which blurs the distinction between method and results. The process of developing the codes requires many readings of the text and experimentation with different sets of codes. The codes and their definitions are the results of detailed analysis of the texts for emergent themes and these sets of codes form the first part of the results of thematic coding. The second set of results comes from the application of the codes to the texts.

1.6 Structure of this thesis

The structure of the thesis has been driven by the structure of the research. The method and results are presented in two parts according to the primary research method used. This reflects the iterative and exploratory nature of the work: each analysis not only provided results but set the parameters for the next piece of work. The chapter structure is:

1. Introduction

2. Literature review

3. Method and results for the political, funding and contextual background to the Australian public hospital system

4. Method for newspaper article analysis

5. Results of newspaper article analysis

6. Discussion and analytical model

7. Conclusions

The results section of Chapter three and the main results in Chapter five have very little commentary and analysis. The analysis is concentrated in Chapter six which brings together both sets of results within the context of the literature reviewed in Chapter 2.

In the next chapter I will describe the pre-existing research underpinning this thesis and describe where my research fits into this research landscape.

2 Literature Review

2.1 Overview

The research in this thesis is concerned with the communication consequences of using simple, numerical information to judge and report on the performance of large, public-sector organisations. The organisation in the case study is the ACT public hospital system in the context of the wider Australian public health system. The theoretical bases for the research come from a wide range of disciplines, ranging from history to communication to health policy and risk perception.

The three broad areas of existing research used in this thesis are:

- Performance indicators and evaluation;
- Public health and public health policy; and
- Communicating about and with numerical information.

Each of these broad areas is itself multi-disciplinary, with research findings published across many different outlets.

2.2 Performance indicators and evaluation

2.2.1 What are performance indicators?

Broadly speaking, performance indicators are the signals we use to monitor the performance of complex systems. They can be as basic as a temperature gauge for a motor vehicle's engine or as complex as the national, quarterly, seasonally-adjusted unemployment figures. Typically, performance indicators are in two parts: a quantity that can be measured; and the reasoning that ties that measurement to the aspect of system performance that is being monitored.

A lot of the research relating to performance indicators is focussed on the validity or usefulness of specific indicators for particular applications. More general work on the theory behind the use of performance indicators is rarer, sometimes occurring as part of the discussion of a specific indicator or application.

A prime example of more general theory as part of a specific application is Donella Meadows' *Indicators and Information Systems for Sustainable Development* (Meadows 1998). While the report as a whole is focussed on indicators related to sustainable development, the opening chapter, *The nature of indicators, the importance of indicators*, is a

general discussion of the idea of performance indicators and their ubiquity. From the opening sentence, "Intuitively we all use indicators to monitor complex systems we care about or need to control" (p1, Meadows 1998) Meadows makes the point that we need indicators to help us deal with systems that are too complex for us to observe and monitor directly. In other words, indicators are by definition simpler than the system under consideration. Meadows is one of the earliest, if not the earliest to use the concept of the feedback loop between measuring what we value and coming to value what we measure. She describes this feedback loop as "common, inevitable, useful and full of pitfalls" (p2, Meadows 1998).

Another important point related to this feedback loop is the impact of indicators on human activities, and the sometimes unintended effects of choosing a particular indicator. For example, you could take your indicator of national success to be Gross Domestic Product and manage your nation to maximise performance against that indicator. If you manage your nation successfully according to your chosen indicator, you will increase Gross Domestic Product but not necessarily achieve anything else. Since the massive spending required for dealing with war and natural disasters usually increases Gross Domestic Product, it could be argued that both of these are active contributors to national success.

Performance indicators: good, bad and ugly is a review paper published in the Journal of the Royal Statistical Society. It was the result of a working party on performance monitoring in the British public service, chaired by Professor S. M. Bird (Bird, Cox et al. 2005). It opens with a general discussion of performance monitoring with particular reference to the [British] public sector and then moves on to look at the overall usefulness of performance measurement in the context of public sector performance management. The general findings, as summarised in the abstract, are "Performance Monitoring done well is broadly productive for those concerned. Done badly, it can be very costly and not only ineffective, but harmful and indeed destructive." (p1, Bird, Cox et al. 2005). While the authors describe some of the reasons for performance monitoring and give examples of what they consider to be good practice, it is clear that after intense study they remain wary of it.

Most performance indicators in the Australian public sector take the form of numbers derived from measuring one or more aspects of the system under consideration. Examples include Gross Domestic Product and unemployment figures

as measures of the national economy (Australian Bureau of Statistics 2016); crime statistics as measures of the effectiveness of law enforcement (Australian Federal Police 2015) and, the subject of this research, elective surgery waiting times as a measure of the performance of the public hospital system.

2.2.2   A brief history of public sector performance indicators

Performance management and performance indicators are not new phenomena. They have been a part of our society for many years in one form or another. Understanding their history, at least as it relates to the Australian public sector in the early years of the 21st century gives context for understanding how and why they are used today.

There are many ways to arrange historical events so as to tell a story, and compiling a history of performance management and performance indicators is no exception. The choice of events and influences that shape the present will be dictated by each historian's background, knowledge and interests (Jordanova 2006). For the purposes of this research, the history is primarily that of the comparatively recent past, the public sector, and countries with organisational, political and economic similarities to Australia – primarily the United Kingdom and the United States of America.

In *Performance Indicators: good, bad and ugly,* Professor Bird and her co-authors give three broad aims for the collection and analysis of public service performance data: finding out what works to deliver against the stated objectives; measuring the performance of individuals and organisations; and accounting to the public for the use of public funding and resources (Bird, Cox et al. 2005). The roots of these ideas go back a long way, and underlying their current forms is often the idea that it is both possible and desirable to use numerically expressed quantitative measurements as indicators for all three of these aims. The usefulness or otherwise of communicating using numerical information will be covered later in this chapter (see Chapter 2.4 on page 35).

Of the three broad aims given by Bird et al., the oldest is arguably that of accountability. The idea of reporting on how you have been spending someone's money or what you have been doing with your time is far from new. As soon as one person gave another some resources and asked them to do something there was the question of "Have you used my resources to do what I asked?". In Western culture references to accountability can be found as far back as the ancient Egyptian Great

Edict of Horemheb (c1300 BCE) which details punishments for corrupt and dishonest officials. The same theme can be found in the Christian Bible in the parables of the good steward (Matthew Chapter 24) and the dishonest steward (Luke Chapter 19).

In feudal times it was kings, overlords and the religious hierarchy requiring accountability but as society became more complex so did the rules governing accountability. The German philosopher Ferdinand Tönnies described two idealised models of social organisation: Gemeinschaft (commonly translated as 'community') which is characterized by personal social interactions, roles, values, and beliefs based on such interactions; and Gesellschaft (commonly translated as 'society') which is characterized by indirect interactions, impersonal roles, formal values, and beliefs based on such interactions. In the latter, there is a need for formal standards, measures and coinage to ensure that business and social activities can be carried out over larger distances and with one or more intermediate parties. This transition from Gemeinschaft to Gesellschaft is traced in Theodore Porter's book *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Porter 1995). In it he describes the social history of standardisation and measurement in post-revolutionary France and the United States of America (Porter 1995). This history includes parts of the history of accountability-style performance management: how well or badly resources have been used and how should they be allocated in the future.

The rise of science and the scientific method from the Renaissance onward led to a greater and greater capacity to make detailed quantitative measurements of the physical world. Isaac Newton's *Philosophiæ Naturalis Pricipia Mathematica* [Mathematical Principles of Natural Philosophy], published in 1687 (Newton 1687), expresses physical laws of motion and force in purely mathematical terms. This mathematical approach became standard in physics and chemistry and still underlies research in the physical sciences. It is commonly referred to as 'Newtoninism'.

This ability to measure and quantify the physical word was paralleled by the idea that to do so was desirable and would lead to a complete understanding of it and the rules by which it operated. By the late nineteenth century this was expressed by Lord Kelvin as "To measure is to know". In a much-quoted longer passage in one of his popular lectures, he says

> *In physical science the first essential step in the direction of learning any subject is to*
> *find principles of numerical reckoning and practicable methods for measuring some*
> *quality connected with it. I often say that when you can measure what you are speaking*
> *about, and express it in numbers, you know something about it; but when you cannot*
> *measure it, when you cannot express it in numbers, your knowledge is of a meagre and*
> *unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your*
> *thoughts advanced to the state of Science, whatever the matter may be. (Kelvin 1883)*

Modern references to this quote often leave out the first sentence, implying that the only valid way of knowing about *anything* is to quantify its attributes.

Some of the earliest formal applications of mathematical and scientific techniques to social and historical research, was done by Newton and his contemporaries. In his *Chronology of Ancient Kings*, Newton analysed historical texts, using astronomical calculations to find precise dates for events in the Classical and pre-Classical worlds that were recorded in ancient texts. His timelines contradicted the then accepted ancient Greek history by 500 years and ancient Egyptian history by 1000 years, this brought him into furious conflict with classical scholars who based their dates only on the texts themselves (Buchwald and Feingold 2013).

Many of these early attempts to apply the newly codified ideas of evidence and mathematical logic to areas of study other than the physical world related to theology and mythology: when exactly did Troy fall? When was the earth created? When did Solomon build the temple in Jerusalem? This led to some of the first calculations relating to human populations using lifespan, mortality rates, gender ratios and proportion of the population with children (Buchwald and Feingold 2013).

A major theme that Porter explores is the process of trying to extend measurement and quantification principles into the social sciences and society more generally. This coincided with the industrial revolution and the move to large-scale manufacturing. He details the development in France of the "physics of work" and its adoption in Britain by William Whewell, author of an 1841 textbook on the *Mechanics of Engineering*. Porter devotes a major chapter to describing the development of the techniques of formal cost-benefit analysis by the US Army Corps of Engineers. He charts and to some extent laments the move from applying cost-benefit analysis to specific projects with clearly understood costs to applying it to more nebulous social interventions where both costs and benefits are difficult to quantify.

2.2.3 **The beginnings of organisational performance management**

In the 1920s Walter A. Shewhart of the Bell Telephone Laboratories pioneered the concept of process quality control and the use of statistics as a tool for managing manufacturing processes. He moved from the idea of measurement at the level of projects to the idea of measurement of individual parts, processes and sub-processes. This work was picked up and greatly expanded upon by W. Edwards Deming, a mathematical physicist and statistician, from the late 1920s onwards.

Deming's innovation was to expand the use of Shewhart's process control tools for monitoring and managing production processes and apply them to the wider management and leadership processes within organisations. He treated the entire organisation as a system containing processes that could be measured, managed and improved.

As well as applying Shewhart's ideas more widely, Deming also worked to make them easy to understand, remember and implement without requiring statistical expertise. One way he did this was to express the philosophy behind his performance management approach in short, easy to remember sayings. The most famous of these, although it may be apocryphal, is "In God we trust, all others must bring data". Deming wrote widely on the topic of statistics and their application to production and management.

Although measurement and statistics underlie much of Deming's philosophy, he remained mindful of the fact that not everything could be reduced to a quantitative measure and that "The most important things are unknown or unknowable" (p97-98 Deming 1982 (reprinted 2000)). This insight of his often gets lost in the drive to measure and quantify.

As well as pioneering the use of statistics for managing business processes, W. Edwards Deming paid close attention to educating people working in businesses about statistics and their use. He divided his audience into four groups depending upon their role in the organisation:

- Management
- Statistical management
- Research
- Production workers

For each group he explained what performance statistics meant in the context of how that group used the information (Deming 1954). For instance, management needed to know the principles of statistical techniques and have an appreciation of them without necessarily being skilled in the techniques themselves while staff on the production line or in inspection needed an understanding of control charts and sampling techniques (Deming 1954). This is an early example of communication about performance measurement being carefully tailored to the needs of different audiences.

2.2.4   Public sector performance measurement

In her book *Public Policy Values* (Stewart 2009), Professor Jenny Stewart outlines the two major phases in the underlying policy of public management in the Anglo-American world. The first, from roughly the mid-nineteenth century to the 1980s had a strong emphasis on due process. Merit, impartiality and professionalism were seen as an antidote to the era of patronage that preceded it. The second, from the 1980s to the present saw a move from public administration to public management with an emphasis on measureable performance as the basis for evaluation and an increased desire for efficiency.

Stewart describes two main drivers for this change: a reaction against the perceived degree of public service autonomy and doubt as to the capacity of the public service of the time to work efficiently and effectively. The latter has been the impetus for much of the performance measurement and performance management in the Australian public sector.

2.2.5   Health sector-specific performance measurement

In 1995 a group of researchers from the Department of Accounting at the University of Waikato in New Zealand published a review of the use of measures and indicators in the management of public health services (Van Peursem, Pratt et al. 1995). Although published more than 20 years ago, the key observations still ring true and some of them were reinforced in *Performance indicators: good, bad and ugly* published ten years later (Bird, Cox et al. 2005). The New Zealand researchers were seeking to establish a means of classifying and describing different types of performance measures so that informed political decisions could be made about what to measure and what to report. This was in part a response to the main criticisms at the time of using performance indicators in the assessment of management:

> *The criticisms fall into two broad categories: first, that performance indicators are deceptive because they falsely convey an impression of objective truth; second, that by making certain aspects of performance visible, they marginalize other management activities.* (p36, Van Peursem, Pratt et al. 1995)

The paper highlights a key issue, common to all public policy outcome reporting, the public resources used are typically reported in economic terms but the outcomes are social and hence "cannot be conveniently measured in precise terms" (p36, Van Peursem, Pratt et al. 1995). A further aspect of this is framed as the relationship between outputs and outcomes or, to put it another way, between activities and achievements. For an outcome such as 'improving public health' both measurement and allocation of responsibility can be extremely difficult. Large-scale outcomes like this can be unpopular at institutional and individual levels because they run the risk of making mangers responsible for circumstances outside their control. The initial priority-setting exercise in the US state of Oregon is an example of an attempt to use this type of measure (see Chapter 2.3 *Public Health policy and Practice* on page 20).

Van Peursem, Pratt et al. also reflect on the idea of representational faithfulness, or the relationship between the concept of interest and what is measured, as they put it: "Measures are only humble representations, often used to achieve particular social or political ends" (p37, Van Peursem, Pratt et al. 1995). This problem of connecting measurable indicators to broad social outcomes became a recurring theme in the research for this thesis.

Two responses are offered to this problem: a range of performance measures needs to be used and considered holistically, rather than individually; and performance measures reported to the public should be used to focus attention towards an area of interest and inspire further enquiry. Indicators cannot stand alone as sources of information about management performance (p37, Van Peursem, Pratt et al. 1995). In the concluding remarks the authors "acknowledge that performance measures may be used in a political manner to further particular interests" (p60, Van Peursem, Pratt et al. 1995) and this issue will always exist. They propose three guidelines for applying performance measurement:

> *A balance of ordinal, nominal and ratio indicators should be produced to avoid the impression that precision has been achieved, as well as to provide a more balanced view.*

> *The way in which they are measured, as well as the measures themselves, needs to be*
> *an open, communicated process. It may be advisable to disclose that process and the*
> *participants who engage in it.*

> *It should be made clear that measures are an indication of a situation which may call*
> *for further enquiry. Indicators do not provide answers, they inspire questions, and this*
> *should be made clear* (p 60, Van Peursem, Pratt et al. 1995).

In their paper *Are performance indicators generic?* Kazandjian, Matthes et al. state that "Performance indicators for healthcare organizations represent a strategy for accountability worldwide" (p265, Kazandjian, Matthes et al. 2003). Their research, based in Maryland USA, looks at the possibility of developing a generic methodology for healthcare performance measurement and evaluation. In 2003 the Maryland Hospital Quality Indicator Project (QIP) had been running for 17 years and at the time was the only indicator-based performance management system used worldwide. As well as hospitals in the USA, 200 hospitals in nine European and Asian countries were part of the QIP. A key driver for using this system was the notion of social accountability, although the initial use of the QIP was primarily within hospitals and health care systems. The QIP is divided into two parts or dimensions: a measurement dimension which is generic and essentially value-free; and a value-laden evaluation dimension which is strongly influenced by the environment and framework within which the measurement is being reviewed (Kazandjian, Matthes et al. 2003). The two-part nature of the QIP increases its portability because the generic measurements can then be interpreted locally.

According to Lemieux-Charles, McGuire et al., the concept of performance in health systems is complicated by there being two different lines of accountability: professional (i.e. health-related) and managerial (Lemieux-Charles, McGuire et al. 2003). Their research looked at the development and use of performance indicators at different levels of the health system and the relationship between those levels. In their description of the theoretical background to their research they observe:

> *The development and use of performance indicators can be seen as both a quest for*
> *legitimacy and a quest for rationality.* (p760, Lemieux-Charles, McGuire et al. 2003)

They explored how these two motivations affected performance measurement at different levels within the health system, concluding that:

> *[performance management] at the institutional level is motivated by legitimacy while the technical/managerial level is motivated by rationality. Tensions exist between the two levels and between indicator development and use.* (p760, Lemieux-Charles, McGuire et al. 2003)

These two motivations, legitimacy and rationality have strong similarities with the ideas in Bird, Cox et al.'s later, more general paper on public sector performance measurement (Bird, Cox et al. 2005). Bird, Cox et al. give three main reasons for measuring public sector performance: research or understanding what works; management or identifying the functional competence of individuals and units; and democratic or providing a framework for Ministers to account for their stewardship of public resources (Bird, Cox et al. 2005). The legitimacy motivation is analogous to providing democratic accountability for the use of public resources while the rationality motivation contains both the research and managerial reasons for measuring public sector performance.

A 2005 paper from the Netherlands explores the theory and practice of using waiting time data as an indicator of health care system performance (Stoop, Vrangbæk et al. 2005). The focus of the paper is the problem of why the publication of waiting times in The Netherlands did not achieve one of its main goals, that of having consumers and general practitioners act on the information. This focus is quite different to that of the papers previously canvassed which were primarily focussed on the use of waiting time data as a way of reporting on the overall performance of the health system. The authors discuss how the need for the waiting time data to be interpreted affects its usability for purposes such as informing consumers. A key point of their argument is the need to distinguish between indicators used for internal purposes and indicators used for external purposes. In this case they define 'internal' as using data for management purposes and quality improvement and 'external' as allowing comparisons between health-care or health-care professionals. This again reflects the ideas of rationality and legitimacy. The use of performance indicators which is not explicitly covered by Stoop, Vrangbæk et al. is that of reporting on the performance of the overall health system. This omission may well be a by-product of the nature of the way the health system is organised in The Netherlands.

A system for measuring the performance of a health care organisation based on the balanced scorecard (BSC) approach was developed by a group of researchers based in Greece (Grigoroudis, Orfanoudaki et al. 2011). The BSC is a popular way of measuring business success by combining performance against a range of measures. Typically, the BSC consists of four major performance dimensions: financial, customer relations, internal business processes and innovation/learning. The researchers adapted the approach to account for the non-profit nature of the public health system but retained some financial measures so that efficiency was still part of the mix of measures. Other measures included the quality of the services provided; the satisfaction of internal and external customers; the self-improvement system of the organisation; and the organisation's ability to adapt and change.

In 2008 during the lead up to the introduction of performance-based funding for the Australian hospital system a Clinical Excellence Commission held a forum exploring how clinical practice variation could be monitored and the quality of hospital care measured and reported on (Leathey, Gilbert et al. 2008). The forum focussed on measures of the quality of clinical care such as the rate of hospital-acquired infections and the incidence of unplanned returns to the operating theatre. The goal of the forum was to develop a small set of safety and quality measures for hospitals in New South Wales (NSW).

There is now a body of research and professional commentary relating to the introduction of performance-based funding for the Australian public hospital system and the effects it has had (Australian Healthcare & Hospitals Association 2009; Curtis, Russell et al. 2010; Nocera 2010). The earliest piece, by The Australian Healthcare & Hospitals Association was published at the time that the idea of performance-based funding for public hospitals was being negotiated (Australian Healthcare & Hospitals Association 2009). It starts by describing a controversy about how some Victorian hospitals were reporting waiting times and goes on to call for well defined and uniform data collection across all parts of the public hospital system. Although published in the *International Journal of Health Care Quality Assurance*, it is a comment piece rather than the reporting of academic research.

Waiting lists and elective surgery: ordering the queue (Curtis, Russell et al. 2010) takes a broad look at how elective surgery waiting lists work with a focus on the assessment of need and timeliness. It begins by stating that waiting lists serve as a newsworthy focus for

discontent with the capacity of the Australian public health system to provide services. It goes on to discuss the increasing demand for elective surgery and the two major limitations on the ability of the public hospital system to respond to this increasing demand: funding and workforce availability. The paper compares Australian and international processes for prioritising access to elective surgery and suggests improvements to Australian practice. The authors identify four key elements of a prioritisation system for elective surgery: equity, transparency, certainty and scientifically valid prioritisation tools.

Antony Nocera looks at waiting lists as performance indicators from a different perspective, concentrating on the question of whether performance-based hospital funding works as a tool for reform or an incentive for fraud in the collection and reporting of data (Nocera 2010). The paper focuses strongly on reported instances of fraudulent performance reporting in both the NSW and Victorian hospital systems and the case of a surgeon in a Queensland hospital whose throughput of patients was high, which meant that elective surgery waiting time performance was good at his hospital, but who also had such high instances of injury and death that he was eventually charged with manslaughter and disbarred from practicing. The manslaughter conviction was eventually quashed in return for a plea of guilty to charges of fraud. The paper concludes with recommendations for nationally uniform legislation to make health service reporting standards consistent and to criminalise public sector data fraud. It also calls for 'realistic outcome measures that base hospital funding more on the quality and safety of patient care and less on patient throughput numbers (paragraph 3, Nocera 2010).

The themes of accountability and management occur throughout the literature relating to public sector performance measurement together with warnings and examples of the unintended consequences of basing performance judgements on these measurements. There is tension between what is valued for different definitions of performance as well as tension between what can be measured and which aspect of public sector performance is being judged.

2.3 Public health policy and practice

Elective surgery waiting times are used as a performance measure within the wider context of public health policy and practice. To understand their use as a performance measure, it is necessary to canvass some of the broader issues related to

the provision of public health services in general and public hospital services in particular. Depending upon where in the world you are and how the health system there is organised, 'public health' can have different meanings and therefore difference performance criteria. For example, in the United States of America, which has largely privatised the provision of health services and health care, 'public health' is taken to mean assessment, policy development, and assurance, not the direct provision of health care (Scutchfield, Bhandari et al. 2009). This contrasts with England where the public health service, known as the National Health Service (NHS) provides healthcare free at the point of service to all permanent residents. The NHS is paid for by general taxation (National Health Service 2015). These differences in the definition of public health must be kept in mind when reviewing research into performance measurement of public health systems.

Australia in the first decades of the 21st century had a wide-ranging, publically-funded health system covering all aspects of primary healthcare: general practitioners, hospitals, pathology and, to a lesser extent mental health services. In addition to this public health system there was a parallel private system, largely funded by private health insurance.

The preamble to the National Healthcare Agreement signed in August 2011 outlines the goals of Australia's health system and the role of public funding within that system:

> *This National Healthcare Agreement affirms the agreement of all governments that Australia's health system should:*
>
> o *be shaped around the health needs of individual patients, their families and communities;*
>
> o *focus on the prevention of disease and injury and the maintenance of health, not simply the treatment of illness;*
>
> o *support an integrated approach to the promotion of healthy lifestyles, prevention of illness and injury, and diagnosis and treatment of illness across the continuum of care; and*
>
> o *provide all Australians with timely access to quality health services based on their needs, not ability to pay, regardless of where they live in the country.*
> (p1 , COAG Reform Council 2011)

As the final bullet point from the National Healthcare agreement states, timeliness and quality are two key attributes required of Australia's public health services. Achieving this within a finite budget calls for some sort of rationing or prioritisation of services and access to them. In P*ublic Policy Values*, Stewart points out the key pitfall in this process:

> *Economists would argue that health expenditures should be allocated where they will benefit most (for example in terms of the value of life years saved), rather than to those [individuals] in greatest need. But empirical research suggests that most people subscribe to the 'rescue principle' where the most exigent conditions are treated first, regardless of cost.* (p78, Stewart 2009)

The concept of the rescue principle can be traced back to a priority-setting exercise carried out in Oregon State, USA in the late 1980s/early 1990s. The budget for public health was limited and the exercise had three clear goals:

- the desired outcome was health, rather than health services or insurance;
- there was a commitment to public processes and transparency of decision-making; and
- the exercise was to develop explicit health priorities to guide decision making.

Rather than excluding people from health coverage or underpaying to reduce access, the strategy was to eliminate particular health services according to "explicit priorities established by an independent commission through an accountable, public process" (p2, DiPrete and Coffman 2007).

Starting from the idea of clinical effectiveness, the Health Services Commission formed to carry out the exercise worked with hundreds of specialists to construct thousands of condition/treatment pairs. The commission also gathered information on public values concerning health care from public hearings, focus groups and surveys. Each condition/treatment pair was assigned a cost/utility value based on a formula which combined:

- Total cost of treatment,
- Cost of care with and without treatment,
- Years for which the treatment could be expected to benefit the patient, and

- Subjective weight for health limitations.

The condition/treatment pairs were ranked according to their cost/utility value and funded according to their ranking. The resulting list of what was to be funded from the public purse was deemed to be unacceptable by both physicians and non-physicians on the review panel. The core problem was that moderately effective, moderately expensive treatments for serious conditions were outranked by very cheap, very effective treatments for minor conditions. The missing factor was the importance of treating the condition in the first place.

A second attempt at a ranked list started by allocating condition/treatment pairs to one of seventeen categories, with the overall categories being ranked according to a combination of public values and then cost/effectiveness. The categories ranged from 1: Acute fatal condition, treatment prevents death with full recovery to 17: Fatal or non-fatal condition, treatment causes minimal or no improvement in quality of life. Services within these categories were prioritized by the commission according to effectiveness and cost. The final step in the process moved away from strictly quantitative methods and the commission members adjusted ranking of condition/treatment pairs was adjusted by hand "to assure that the prioritized list reflected their best judgment as clinicians and as representatives of those to be covered under the resulting benefit package" (p4, DiPrete and Coffman 2007).

The Oregon priority-setting exercise has attracted attention from many researchers into health policy. Ubel, Lowenstein et al canvass some of the explanations for the failure of the initial priority list based on cost-effectiveness in the opening of their 1996 paper *Individual Utilities Are Inconsistent with Rationing Choices: A Partial Explanation of Why Oregon's Cost-Effectiveness List Failed* (Ubel, Lowenstein et al. 1996). They review three main pieces of research which each give different reasons for the failure.

The first was by Kaplan who developed the 'quality of well-being scale' used to measure people's utilities for different cost/treatment pairs. Here 'utility' is used in the economic and game theory sense of the perceived ability to satisfy needs or wants. He argues that the list failed because the commission did not gather and present sufficient data about costs and outcomes (Kaplan 1992).

In the second body of research, David Hadorn, who acted as a consultant for Oregon, argues that all cost-effectiveness lists are doomed to fail, because cost-

effectiveness ratios underestimate people's desires to rescue those whose lives are endangered or who are seriously ill. Hadorn first published his analysis of the Oregon exercise in 1991 (Hadorn 1991), following up with a later review in 1996 (Hadorn 1996). The crux of Hadorn's work is the rescue principle, or Rule of Rescue, which can be summarized as "people's desire to rescue those whose lives are endangered or who are seriously ill" (quoted at p109, Ubel, Lowenstein et al. 1996).

While the Rule of Rescue is crucial to understanding public responses to issues of access to healthcare, it is not the only factor in play. In the third piece of research referred to by Ubel, Lowenstein et al, Nord et al take a distinctly Australian look at public preferences for how healthcare resources should be distributed (Nord, Richardson et al. 1995; Nord, Richardson et al. 1995). The first of these papers looks at public support for distributing health resources so as to maximize overall public health. According to the research, this policy "received very limited support when the consequence is a loss of equity and access to services for the elderly and for people with a limited potential for improving their health" (p1429, Nord, Richardson et al. 1995). In another paper, the authors looked at public concern for costs in the setting of health care priorities. They found that most of the respondents to their survey felt that it was "unfair to discriminate against patients who happen to have a high cost illness and that costs should therefore not be a major factor in prioritising" (p93, Nord, Richardson et al. 1995). Factors such as the effect this would have on the number of people who could be treated as well as the impact on their own chances of treatment did not have a major impact on the respondents' views. The overall conclusion was that "the cost-effectiveness approach to assigning priorities in health care may be imposing an excessively simple value system upon resource allocation decision-making" (p79, Nord, Richardson et al. 1995).

After canvassing this other research, Ubel, Lowenstein et al go on to explore another possible partial explanation for the failure of the initial Oregon exercise. In this survey-based research they first elicit respondent's utilities for various states of health and then used those responses in a second survey relating to the rationing of health services. The rationing choices were not only very different to what might have been predicted from the utility responses, they were also internally inconsistent with the same treatments being rated differently in different scenarios. The technique used to predict the rationing responses from the utility responses was to calculate the point at

which the importance of treating patients with two different conditions was equivalent. For example, the prediction from the utility questions was that treating 10 people with meningioma (a non-cancerous brain tumour) was of equivalent importance to treating 100 people with a ganglion cyst while the responses to the rationing questions had the equivalence point at 10 patients with meningioma and between 1,000 and 10,000 patients with a ganglion cyst. In the rationing part of the survey people consistently gave more weight to severe conditions regardless of the utility (i.e. impairment) caused by them. The research suggested that:

> *… people's answers to utility elicitations cannot be easily translated into social policy However, person-tradeoff elicitations, like those given in our rationing survey, cannot be substituted for established methods of utility elicitation until they can be performed in ways that yield acceptable internal consistency* (p108, Ubel, Lowenstein et al. 1996).

In the philanthropic arena, the behaviour relating to wanting to help identified individual rather than a population is referred to as the Collapse of Compassion (Cameron and Payne 2011). As the number of people in need of help increases, the degree of compassion people feel for them tends to decrease and the likelihood of donations also decreases.

Debate and research relating to the Rule of Rescue and variations on it continue to this day. Many reasons have been given for the phenomena relating to choosing to help identifiable individuals rather than populations regardless of the comparative benefits of the two course of action. some of the more prominent reasons advanced are:

- Humans, as social animals, have a limited number of people with whom they can have a favour-exchanging relationship. This number is ~150 and is based on the size of the neocortex (Dunbar 1996). In order for us to help someone they must, at least temporarily become a part of our social group (McRae 2011).
- The Rule of Rescue increases societal well-being by reinforcing the idea that we live in a society that places great value on life (McKie and Richardson 2003)

- People expect the needs of large groups to be potentially overwhelming, and hence engage in emotion regulation to protect themselves from being overwhelmed by distressing emotions (Slovic 2007; Cameron and Payne 2011).

- There are two processes involved in deciding to help someone or something: sympathy which is essentially emotional/irrational and deliberation which involves working out how deserving the object is of help and what help it needs. When these work together we get the ideal of helpful assistance to those who deserve it. Often the two processes work separately or one at a time rather than together (Lowenstein and Small 2007).

Although the Rule of Rescue and its counterparts show some of what drives public acceptance of health rationing decisions, if it is followed exclusively it can lead to a 'tragedy of the commons'. This is a concept pioneered by Garrett Hardin who in a 1968 article coined the phrase to describe the potential destruction of a common resource if everyone to whom it is available acts only in their own self-interest (Hardin 1968). Every small decision to use more of the common resource at an individual level can result in a resource disaster at the community level. Hardin's article has formed the basis for research and commentary relating to environmental and sustainability since it was first published. Although more recent research has cast doubt on the idea of humans acting as rational, self-interested entities, in particular the Nobel-prize wining work of Kahnemann and Tversky, (Tversky and Kahnemann 1974) the concept is still proving useful. A 2009 paper reworks the tragedy of the commons using large corporations, rather than individual people as the rational and self-interested agents (Rowland 2009) and found that, at this organisational level there was clear evidence that the tragedy of the commons occurs.

In public health today there is increasing debate about the cost to the community of devoting huge resources to life-saving intensive care at the very end of a person's life. The debate about what constitutes appropriate 'end of life' care is painful, emotional and value laden (Katelaris 2011; Australian Commission on Safety and Quality in Health Care 2015). One way to frame this debate is as a conflict between the Rule of Rescue and the desire to avoid a tragedy of the commons. The desire to delay death often leads to extremely expensive intensive care facilities being used to prolong life when there is little or no chance of even short-term remission and no hope of long-

term recovery. Because these resources are so costly, they are also of limited capacity. This means that there is a high likelihood that using them for end-of-life care will prevent some people from accessing them for life-saving care which will lead to full or partial recovery (Katelaris 2011).

A 2013 paper (Jecker 2013) examined the Rule of Rescue from an ethical perspective and argued that, despite its intuitive appeal, the Rule of Rescue lacks support from principles of ethics and justice and so should be rejected as a guide to resource allocation in many situations. In this paper, Jecker draws on the parable of the Good Samaritan to compare duties of justice which are considered obligatory and duties of charity which are optional. In defining the Rule of Rescue Jecker restricts it to providing help to those who have at least some chance of benefiting from it and then asks the question 'Is there an ethical obligation to provide assistance or is it morally admirable but not ethically obligatory? (p66-67, Jecker 2013). The paper distinguishes between agent-relative morality which appeals to moral reasons that are tied to a specific project, value or relationship and the more traditional moral theory construct of agent-neutrality where people act for reasons that are impartial and not tied to their personal connections. The author concludes that "rescue is a moral problem we need to assess rather than a moral mandate we can safely assume" (p80, Jecker 2013).

The Rule of Rescue, the Collapse of Compassion and the Tragedy of the Commons all involve humans acting in the short-term interest of individuals, if not at the expense of the greater good, at least without regard to the consequences for the greater good.

As an empirically observed pattern of human behaviour, the Rule of Rescue has strong communication implications for public health policy, even if its underlying causes are not clear-cut. Messages that stress benefits to many people or to the overall population are less likely to gain agreement and support than those which focus on dramatic help for a few who would otherwise be at immediate risk.

Australian academic Jenny Stewart frames her discussions of public policy in terms of values such as efficiency, equity and accountability. For each of these values there are arguments for and against maximising that quality in the design and implementation of public policy (Stewart 2009).

Stewart describes the value of 'efficiency' as asymmetric in that it is much easier to argue for than against it. She breaks cost efficiency into two components: productive or technical efficiency and allocative efficiency. The first describes the way that an organisation uses resources (inputs) to achieve its outputs. Improving productive efficiency means either producing more outputs with the same inputs or producing the same outputs with fewer inputs. Allocative efficiency describes the way resources (inputs) are distributed within an organisation or, more broadly, within society as a whole. Allocative efficiency in an economy is enhanced when resources are redistributed between uses in such a way that a higher level of consumer surplus is achieved. The idea behind allocative efficiency is that competitive markets do the best job of aligning resource use with personal preferences.

These underlying values are often grouped by Stewart into competing pairs as a way of describing and analysing choices in policy development and implementation (Stewart 2009). Typical value pairs, which need to be traded off in the public health domain are efficiency/equity; fairness/equity and efficiency/accountability.

The health sector is one where the tensions between efficiency and equity are very strong. This is particularly the case in Australia where there is a high-level goal of timely access to health care regardless of location and ability to pay (COAG Reform Council 2011). Providing this level of care is expensive, particularly for people in rural and remote areas. A large part of the work of the National Health Reform Agenda lies in determining 'efficient prices' for health services.

The Economics of Health Equity is an edited book exploring the concept of health equity from an economic perspective. In the Introduction it offers a range of options for the construction of equity in health care such as equal health outcomes, equal access to care for people with equal need for care, and equal use of care for people with equal need for care (Mcintyre and Mooney 2007).

The authors explicitly acknowledge that more tightly specified definitions of health equity are highly contextual and each comes with its own set of questions, problems and further refinements. They also clearly distinguish between horizontal equity or the equal treatment of equals and vertical equity or the unequal but equitable treatment of non-equals. Another way of looking at this is that horizontal equity involves equal resources or inputs while vertical equity aims for equal outcomes. This latter concept is far harder both to define and to act upon. For example, in the

education system horizontal equity would involve allocating the same resources to each student while vertical equity would allocate each student the resources they needed to achieve the same outcome. Vertical equity means that those with some form of disadvantage receive more resources than those who are not disadvantaged. Determining how great the starting inequities might be and how great a difference in policy response might be involves subjective judgements as well as simple formulaic calculations. The policy response usually involves some form of positive discrimination for which it can be politically difficult to make a case (Mcintyre and Mooney 2007).

An international study published in 2012 uses the techniques of multi-criteria decision analysis (MCDA) to determine the relative priorities given to efficiency and equity in the policy-setting of five countries (Brazil, Cuba, Nepal, Norway and Uganda) (Mirelman, Mentzakis et al. 2012). The overall goal of the research was to formalise priority-setting in public health and to find a tool to help determine the trade-offs between efficiency and equity in a formal and transparent manner. Of particular interest was the scenario of a more developed country with one set of assumptions about efficiency and equity assisting a less developed country with a different set of assumptions and the need to develop a shared understanding of the efficiency/equity trade-off to be used. The authors concluded that:

> *Greater use of MCDA in health priority setting would likely make national decisions more transparent and perhaps more rational and allow countries to characterize their efficiency and equity trade-offs in a manner that is consistent with their level of development and societal preferences* (p538, Mirelman, Mentzakis et al. 2012).

Again the concept of equity was found to be affected by context and culture.

On the surface, equity and fairness might be seen as essentially the same thing, but Stewart makes a distinction between them that has consequences from a communication perspective. She uses equity to describe formal judgements about the outcomes and aims of policy processes and fairness to describe various perceptions of those outcomes and aims. The gap between equity and fairness is greatest when the policy aim is based on vertical equity, that is achieving equal outcomes, which results in the uneven distribution of resources.

In my research I have narrowed equity down to formal intent, and fairness to perception of outcomes. As Stewart points out, there is no single, agreed perception of what is fair and despite its being subjective and variable, this has a large impact on the development and implementation of social policy:

> *"…social policy is as much about fairness as it is about equity. As it is implemented in the real world, social policy invokes a set of intricate relationships between different perceptions of fairness. In particular, perceptions of unfairness (that is, feelings against those receiving more than they are entitled to) play a decisive role in shaping responses to the distributions of goods and services that existing policies make possible."* (p66, Stewart 2009)

When it comes to trade-offs and accountability the major one is between accountability and efficiency. Measuring, monitoring and reporting on the activities of an organisation takes resources that could otherwise be used to carry out its core business. In their conclusions, Lemieux-Charles, McGuire et al. note that "The costs and values of engaging in performance systems are unclear" (p769, Lemieux-Charles, McGuire et al. 2003). This is reinforced in *Performance indicators: good, bad and ugly* where the authors comment on how costly and potential damaging poorly done performance management can be and that even when done well performance measurement is merely broadly productive (Bird, Cox et al. 2005).

In 2000 Barbara S. Romzek published an explicit analysis of the interplay between public sector reforms aimed at increasing efficiency, effectiveness and responsiveness in the United States, in the previous decade (Romzek 2000). The emphasis of the paper is primarily on personal accountability and the difficulty faced by public servants of being accountable to multiple sources of legitimate authority. She looks at the effect on accountability measures of changes in the public sector to make it more decentralised and responsive. A key point is the need to ensure that the chosen accountability measures relate to and reward the behaviour changes required by the reform process.

> *"If management reform rhetoric emphasizes employee discretion and autonomy, but performance reviews emphasize rules and process, then employees will emphasize rules and processes in their work efforts. This same dynamic applies at the corporate level also."* (p39, Romzek 2000)

This observation encapsulates the both problem of unintended consequences and the problem of mismatched expectations and measures. In my case study there is a clear illustration of this problem in the manipulation of waiting list data to meet arbitrary performance benchmarks. The cost of inappropriate accountability measures is not always financial, it can lead to poor achievement of the underlying outcomes the performance management processes are seeking to maximise.

A study from the UK looked at the association between how widely quality improvement measures were implemented in hospitals and those hospitals' performance against indicators of clinical quality (Weinex, Alexander et al. 2006). The research focussed on formal Quality Improvement (QI) which involved continuous monitoring and improvement of work process by specialist teams trained in basic statistics, equipped with problem-solving tools and given the authority to make decisions. QI is widely but not uniformly implemented in the UK, with 1,784 hospitals surveyed. The authors found that the effectiveness of QI in improving clinical quality indicators was associated with how it was implemented in each hospital. Hospitals with a high involvement of staff and senior managers reported positive outcomes on clinical performance indicators while those with broad involvement across many disparate units had negative outcomes. The involvement of physicians in the QI teams had no effect on performance against the clinical indicators. The possible reasons given for the negative association between broad involvement of hospital units in QI and performance against clinical indicators all involved resourcing for the QI effort and the need for technical and financial support for it to succeed. Broad involvement of many units was hypothesised to lead to a diluting of effort, resources and expertise meaning that there were many small QI projects none of which had sufficient resources to be successful (Weinex, Alexander et al. 2006).

As well as these general conflicts between different values in the setting of social policy, there are some specific conflicts arising in the health care sector. The report from the Academy of Royal Medical Colleges (UK) published in November 2014 *Protecting Resources, promoting value: a doctor's guide to cutting waste in clinical care* (Maughan and Ansell 2014) essentially argues that it is the ethical duty of medical practitioners to be more efficient by eliminating waste. The rationale given is that waste in one part of the health system can lead to delays or unavailability of treatment in other

parts of the system. Looking at the responses to the report in the BBC News (Brimelow 2014), as well as commentary by those in support of the concept, there was also a clear opinion from organisations such as the British Medical Association that patients' needs had higher priority than cost-saving and that a focus on costs would undermine the quality of patient care.

The Fall 2012 edition of *The Harvard Health Policy Review* contained a feature article criticising the Affordable Care Act in the USA primarily on the basis of its regulatory and compliance costs. Among other reporting and regulatory changes, the paper discusses the implementation of ICD-10, at the time a new, more complex classification of diseases for reporting, and argues that implementing it will cost many thousands of dollars per physician. Proponents of the new system argued that its finer granularity would allow for more accurate data capture and hence more effective and sophisticated management of patients. Opponents saw the finer granularity as burdensome and expensive, taking time away from direct patient care (Manchikanti and Hirsch 2012).

Another conflict within the heath care system is that between the needs of the population as a whole and the needs of individual patients. It is easily possible for the optimum treatment for an individual to be suboptimal for the wider population. This tension extends beyond the question of resource allocation as can be seen in the case of antibiotic overuse. Work by Broom, Broom and Kirby examined the reasons why, when antibiotic overuse is well known to be a problem, doctors are still prescribing them in such quantities (Broom, Broom et al. 2014). In this detailed qualitative research the authors found that "the rules of the game were heavily weighted in favour of the management of immediate clinical risks, reputation and concordance with peer practice vis-à-vis longer-term population consequences" (p81, Broom, Broom et al. 2014). Although this finding is in a different context and draws upon different theories, it echoes the Rule of Rescue findings of the mid-1990s (Nord, Richardson et al. 1995; Nord, Richardson et al. 1995; Hadorn 1996; McKie and Richardson 2003).

2.3.1   *Communicating public policy*

A critical aspect of public policy is its communication to the public. Citizens who pay for and rely upon public health, or any major public service, need to be informed about a range of aspects of its planning, policy and implementation. At the simplest

level, people need to know what services are available and how to gain access to them. There is also a need to communicate what is planned, why it is planned, and, the subject of this research, how well the system and its component parts perform. A component of communicating how well the system is performing is how well and appropriately the system is being measured. Two important sides to this communication are how government communicates with the public and how that communication is reported in the news media.

2.3.1.1 Government communication

Government Communication in Australia is a broad-ranging look at communication between Australian governments, the electorate, business and the media. The book contains a series of essays exploring how governments communicate and how media outlets report government activities (Young 2007). In the introduction, Young observes that

> *It [communication] is a dimension of every action or decision a government takes, from the way policies are made, promoted and enacted, to how government is organised and the relationships it builds with citizens, the media and other groups such as business and community organisations.* (pxxiii, Young 2007)

The book is primarily focussed on political communication and media-based public information campaigns rather than the communication of detailed information about policy and processes. It follows the development of formal communication teams within both the parliament and government departments, with an emphasis on relationships with the press. Of most relevance to my research are Chapter 3 *The Public Service and government communication: Pressures and dilemmas* by Brian Head (Head 2007) and Chapter 11 *Australian Governments and online communication* (Chen, Gibson et al. 2007).

Providing information to citizens and communities about government services and programs, giving free access to official information, and providing opportunities for feedback to the government are a vital part of a healthy democracy, with freedom of information flow in both directions regarded by bodies such as the OECD and Transparency International as being crucial for good government and accountability (Head 2007).

The author goes on to discuss the conflation of information provision and government marketing with the comment "The underlying intention and the cumulative effect is to promote an image of governmental competence and care." (p38, Head 2007). He warns of the potential dangers of this blurring of boundaries and the prospect of public servants finding themselves with conflicting interests as a consequence.

The conclusion of the chapter refers to "achiev[ing] the public-interest goal of informing and involving citizens and stakeholders, not only about the services offered by the government but also about the issues that are currently under consideration and open to consultation". (p49, Head 2007). The issue it does not canvass is that of reporting back to the community on the performance of existing services and facilities.

In *Australian Governments and online communication,* Chen, Gibson et al. (2007) discuss the impact of widespread internet access on the way governments communicate with citizens. The opportunities for faster, higher volume publication of information coupled with decentralised and interactive communication capabilities creates challenges as well as opportunities for a representative democracy. The chapter looks at two main types of online communication e-government, or using online tools to deliver services and provide access to information, and e-democracy, or using online tools to engage and consult with citizens and the community. At the time the chapter was written (2007) Australian governments were making far more use of the former. In Australia e-government has two main aspects: creating public value by developing new forms of services delivered electronically, and reducing government costs by automating 'compliance' activities such as the online lodgement of taxation dues, health insurance claims or business paperwork (Chen, Gibson et al. 2007). A third aspect, of particular interest to my current research is that of increasing the transparency of government processes, performance and policy development by making information relating to these issues freely available online.

2.3.1.2 **News media reporting**

Unless a member of the public has a need for specific policy or procedural information, in which case they are likely to go looking for it, the most probable channel for them to find out about government policy is through the news media. This means that the policy and related issues they find out about are those which

have been deemed to be newsworthy in some way. News organisations select what to publish based on editorial understanding of what types of articles are likely to appeal to the audience and therefore generate revenue.

Journalism textbooks give lists of criteria that influence how newsworthy something is. While these vary slightly between texts, the core criteria, often referred to as 'news values', remain essentially the same: impact, immediacy, proximity, prominence, novelty, conflict, and emotional content. The last value, emotion, is often referred to as 'human interest' and refers to the ability of a story to arouse emotional responses such as happiness, sadness and anger by evoking empathy, compassion and curiosity (Conley 2002; Harrower 2007). The presence of negative statements and references to conflict are seen to enhance news value. Research conducted in the UK showed that that negative news about health and medicine was more likely to be published in the mainstream media that positive news (Bartlett, Stern et al. 2002).

The influence of these news values should be evident in the selection and framing of elective surgery stories for publication in *The Canberra Times* I analysed as part of my research.

2.4 Communicating with and about numerical information

As discussed earlier (see Chapter 2.2.4 on page 15), public sector performance indicators in Australia tend to be expressed numerically. The main performance indicators from my case study are counts of people or surgeries performed and averages, usually medians, of the number of days waited for surgery. There is a huge body of research across multiple disciplines relating to how humans process and understand not just statistical information but numerical information.

One of the fundamental theories used in science communication is that of constructivism. Drawing heavily on the constructivist learning model, it proposes that when given new technical or scientific information audience members construct their own meaning(s) by fitting the new information into the framework provided by their previous knowledge (Yager 1991; Stocklmayer 2013). In order to apply the constructivist theory to communication with and about numerical information, we first need to understand what pre-existing constructs our audience might have when it comes to numerical information.

2.4.1 Learning about and understanding statistical concepts

In Australia, school attendance is compulsory up to Year 10 (~16 years old) with Mathematics as a core subject for the whole period. The Australian National Curriculum for Mathematics was endorsed and released in 2010. It sets out what is taught as common content across the country at each year level from pre-school (~5 years old) to Year 10 (~16 years old). A separate document specifies the advanced mathematics curriculum for the final two years of school. This means that the minimum level of mathematics the next generation of Australian adults will have been exposed to is, in theory, that of the common Year 10 curriculum. Having been taught a concept is, of course no guarantee of understanding it. Before the release of the Australian National Curriculum each State and Territory had its own broadly similar curriculum. Using the current curriculum as a guide, Table 2-1 below shows a selection of when various mathematical and statistical concepts are taught (Australian Curriculum Assessment and Reporting Authority (ACARA) 2010).

Table 2-1: Mathematical and statistical concepts taught in Australian Schools
(compiled from the Australian National Curriculum)

School Year and approx. age	Selection of the Mathematical and statistical concepts taught
Foundation and Year 1 ~6-7 years old	By the end of the Foundation year, students make connections between number names, numerals and quantities up to 10. They compare objects using mass, length and capacity.
	By the end of Year 1, students describe number sequences resulting from skip counting by 2s, 5s and 10s. They identify representations of one half. Students count to and from 100 and locate numbers on a number line. They carry out simple additions and subtractions using counting strategies. They partition numbers using place value.
	Students classify outcomes of simple familiar events (will happen, won't happen, might happen). They collect data by asking questions and draw simple data displays.
Year 6 ~12 years old	By the end of Year 6, students recognise the properties of prime, composite, square and triangular numbers. They describe the use of integers in everyday contexts. They solve problems involving all four operations with whole numbers. Students connect fractions, decimals and percentages as different representations of the same number. They solve problems involving the addition and subtraction of related fractions. Students make connections between the powers of 10 and the multiplication and division of decimals.
	Students list outcomes of chance experiments with equally likely outcomes and assign probabilities between 0 and 1.
	Students compare observed and expected frequencies. They interpret and compare a variety of data displays including those displays for two categorical variables. They evaluate secondary data displayed in the

School Year and approx. age	Selection of the Mathematical and statistical concepts taught
	media.
Year 7 ~13 years old	By the end of Year 7, students solve problems involving the comparison, addition and subtraction of integers. They make the connections between whole numbers and index notation and the relationship between perfect squares and square roots. They solve problems involving percentages and all four operations with fractions and decimals. Students represent numbers using variables. They connect the laws and properties for numbers to algebra. They interpret simple linear representations and model authentic information. Students use fractions, decimals and percentages, and their equivalences. They express one quantity as a fraction or percentage of another. Students solve simple linear equations and evaluate algebraic expressions after numerical substitution. Students identify issues involving the collection of continuous data. They describe the relationship between the median and mean in data displays. Students determine the sample space for simple experiments with equally likely outcomes and assign probabilities to those outcomes. They calculate mean, mode, median and range for data sets. They construct stem-and-leaf plots and dot-plots.
Year 10 (standard) ~16 years old	By the end of Year 10, students solve problems involving linear equations and inequalities. They make the connections between algebraic and graphical representations of relations. Students expand binomial expressions and factorise monic quadratic expressions. They find unknown values after substitution into formulas. They perform the four operations with simple algebraic fractions. Students solve simple quadratic equations and pairs of simultaneous equations. They compare data sets by referring to the shapes of the various data displays. They describe bivariate data where the independent variable is time. Students describe statistical relationships between two continuous variables. They evaluate statistical reports. Students list outcomes for multi-step chance experiments and assign probabilities for these experiments. They calculate quartiles and inter-quartile ranges.

As can be seen from Table 2-1, there is a strong focus in mathematics education on links to concrete objects and the everyday world. The early work on probability focuses on the concepts of 'likely' and 'unlikely', with formal, numerical odds being taught later. The concepts become distinctly more abstract in Year 7 with the introduction of formal statistical language. The idea of independent trials is not taught until Year 8 (~14 years old).

A major report into literacy and numeracy skills in Australia and their impact on labour market outcomes was published by the Productivity Commission in May 2014 (Shomos and Forbes 2014). It gives a different perspective to the national curriculum

on what statistical knowledge might be expected in the Australian population as it attempts to describe what is known rather than what has been taught. Using data from the Australian Bureau of Statistics' (ABS) *Programme for the International Assessment of Adult Competencies* (PIAAC) it uses six bands for describing a person's level of numeracy. A score of Level One and below means that someone has only the most basic of numeracy skills: counting, basic arithmetic simple percentages and simple graphs. From Level Two up to Level Five the amount of skills in statistics steadily increases from very simple (similar to mid-primary school in the National Curriculum) to very complex including the ability to critique, evaluate choices of models and representations of data. In 2011-12 almost 22% of Australians aged between 15 and 74 years had a numeracy level of one or below, meaning that the had essentially no understanding of statistics. A further 32.5% had a numeracy level of two, meaning that their understanding of statistics is limited to interpretation of relatively simple data and statistics in texts tables and graphs (Shomos and Forbes 2014). This leaves under half the adult population with the skills for the interpretation and basic analysis of data and statistics texts, tables and graphs. (Shomos and Forbes 2014)

Research into the phenomena of anchoring and adjustment suggests that when dealing with a range of probabilities people tend to choose a few points in the range to which they attach a verbal meaning, adjust the actual point they are given so that it matches one of the anchors and then act on the result (Lichtenstein and Slovic 1971; Tversky and Kahnemann 1974). For example, a range of probabilities from 0% to 100% in 5% increments might be reduced to three anchor points 0% = not going to happen, 50% = don't know, and 100% = will happen. Any amount in that range is then adjusted to the anchor point that seems nearest which may not be the closest numerically. The end points, with their definite outcomes are stronger 'attractors' that the indeterminate middle. This has strong implications for the use of waiting times as the clinical guidelines for different priorities of surgery will act as anchors.

2.4.2 *Day-to-day encounters with statistics*

In our day-to-day lives, most numerical information is strongly related to real objects: five loaves of bread, 30 litres of fuel at $1.85 litre, a certain number of dollars in our pay packet. More complicated uses of numbers like the interest on our mortgage or a

percentage score on an exam are still basically unambiguous and have strong links to the world we experience directly. The most common regular exposure to statistics for most Australians would be in the context of the daily weather forecast and the probability of rain. The Australian Bureau of Meteorology now uses words rather than percentage probabilities as the primary description of the chance of rain, as shown in Table 2-2 below. A possible reason for this is to take advantage of anchoring (Commonwealth of Australia Bureau of Meteorology 2015).

Table 2-2: Rainfall Probabilities and descriptions

Chance of rain	Terminology used
0%, 10%	No mention of rainfall in forecast
20%, 30%	Slight (20%) chance of rainfall in forecast Slight (30%) chance of rainfall in forecast
40%, 50%, 60%	Medium (40%) chance of rainfall in forecast Medium (50%) chance of rainfall in forecast Medium (60%) chance of rainfall in forecast
70%, 80%	High (70%) chance of rainfall in forecast High (80%) chance of rainfall in forecast
90%, 100%	Very high (90%) chance of rainfall in forecast Very high (near 100%) chance of rainfall in forecast

Other examples of statistical information in everyday life such as the unemployment rate are referred to intermittently and, unless there has been a problem with the data preparation[1], simply presented as unambiguous numbers. People who use more formal statistics regularly mostly do so in a specialised professional or education context.

---

[1] In 2014 the Australian Bureau of Statistics had to make major revisions to the seasonally adjusted unemployment rate after problems were found with how it had been calculated.

Broadly speaking, statistics are numbers which describe some of the properties of a group of objects which may or may not be real. In addition, the language of statistics assigns a precise, technical meaning to words such as 'average' and 'significant' which also have colloquial meanings. The blurring of the lines between real and idealised objects and between technical and colloquial language means that there can be a very large gap between the specialist's and layperson's interpretation of what a statistic means.

The way early mathematical training emphasises the links between numbers and quantities in the real world, combined with little regular exposure to formal statistics means that the heuristic shortcuts most people use every day for dealing with numbers and quantities are, by and large, invalid for dealing with statistical concepts. Take the simple example of the number of beds available in a hospital. The immediate, heuristic approach is to envisage exactly that number of freshly-made empty beds. What this image does not encompass is the idea that, in technical terms, a 'hospital bed' is a complex construct of people, equipment, space, time and other resources of which the physical bed is only a small part. It takes considerable effort, even for someone experienced with statistics to move past the immediate heuristic picture and unpick a statistic slowly and rationally.

2.4.3 Communicating statistics

As social animals, humans tend to respond well to stories conveyed by a trusted entity and which combine emotional and technical content. Discussions of this go back to Aristotle who, teaching in 350BCE, divided rhetoric, or persuasive communication, into three components:

- Ethos: appeal based on the character of the speaker.
- Logos: is appeal based on facts combined with logic or reason.
- Pathos: is appeal based on emotion.
  (Aristotle 350BCE)

Viewed through this lens, statistical information is primarily an appeal to logos with some ethos. The ethos is not necessarily positive in a persuasive sense – some research suggests that people do not trust statistical information in general while some do not trust the way official statistics are used, even if they trust the source of the statistics themselves (Simpson, Beninger et al. 2015). The addition to a statistical

report of a case study featuring individuals is a way to add the pathos or emotional component although it runs the risk of making things less clear. Just because the case study comes from the population represented by the statistic does not mean that the case study is itself representative of the population.

A subset of more recent research into communication, persuasion and decision-making looks at human abilities when dealing with facts, numbers and logic in a systematic fashion. The field of risk perception and risk communication is particularly rich in this type of research. Another area rich in this sort of research is science communication and studies of popular science/mathematics writing.

A series of books published between 1999 and 2005 by scientists and writers Ian Stewart and Jack Cohen in collaboration with popular novelist Terry Pratchett explore not only science, but some of the consequences of trying to look at the world in a scientific way. Although they are not, strictly speaking scholarly works, the cross-disciplinary mix of a mathematician, a biologist and someone whose livelihood depends on their ability to tell enthralling stories offers insights into the relationship between science and an ancient form of human communication: storytelling. The series, *The Science of Discworld*, uses the frame of a world based on magic to examine the role of science as a way of explaining our world. The concepts of randomness, complexity and the use of statistics are all discussed in the context of their role in scientific enquiry.

A key concept in the books is one the authors describe as *narrativium* or the urge to paint stories on the universe (Pratchett, Stewart et al. 1999). This concept is used to describe the processes by which humans develop understanding of what they observe and the reasons why we tend to see patterns that reflect the stories we are already telling ourselves. The tone of the books is often informal, bordering on frivolous, but the concepts come straight from research into the philosophy of science and the nature of scientific enquiry.

> *Humans add narrativium to their world. They insist on interpreting the universe as if*
> *its telling a story. This leads them to focus on facts that fit the story, while ignoring*
> *those that don't. But we mustn't let the coincidence, the clump, choose the sample space*
> *– when we do that we're ignoring the surrounding space of near coincidences*
> (p233, Pratchett, Stewart et al. 1999).

Each book explores a different aspect of science, each one relating back to storytelling and communication. A key observation is that:

> *Science is a structure created and maintained by people. And people choose what interests them, and what they consider to be significant and, quite often, they have thought narratively.*
>
> …
>
> *This human trait doesn't affect what the rules say – not much, anyway – but it does determine which rules we are willing to contemplate in the first place. Moreover, the rules of the universe have to be able to produce everything that we humans observe, which introduces a kind of narrative imperative into science too*
>
> (p11, Pratchett, Stewart et al. 1999)

A more formal look at similar topics is *Tribal Science* by Mike McRae, published in 2011 (McRae 2011). Extensively referenced and indexed, this book is a semi-populist analysis of human social thinking and its impact on our ability to 'do' science. McRae draws heavily on Dunbar's work in evolutionary anthropology (Dunbar 1996) to construct an argument for humans, as social animals, being evolved to deal with personalities and stories rather than numbers and abstract theories. He views the primary function of the brain as a tool for co-operating with our 'tribe', usually our friends and family, to make the most of the resources available in our environment.

McRae also quotes the four levels of evaluating evidence, developed by Deanna Kuhn, an educational psychologist. Kuhn's classification follows cognitive development from early childhood, each level being more sophisticated than the one before:

- Realist: Knowledge is certain and our perceptions are an accurate perception of the universe;
- Absolutist: Knowledge is viewed as right or wrong. We can see that it is possible for alternate positions to exist in the minds of others, but these positions are always objectively correct or incorrect;
- Multiplist: Knowledge is viewed as a construct of thinking, usually contingent and therefore often open to negotiation; and

- Evaluativist: The relative strength of opinions is judged according to a set of values and using thinking tools to determine how confident we should be in the validity of a belief.

The levels closely follow the development of Theory of Mind from the work of Jean Piaget in the second half of the 20th century. Theory of Mind posits two systems for tracking other people's belief states, one early-developing, cognitively efficient but inflexible and the other later-developing, cognitively demanding and flexible systems. Both models provide additional frameworks for looking at how humans use evidence to make decisions.

Dunbar's work on the evolution of language and the adaptation of human brains to facilitate the formation and maintenance of co-operative social groups gives one set of reasons for our typically poor ability to deal with numbers and probabilities (Dunbar 1996; Dunbar 1998). His work on the size of social groups also provides the basis for one of the explanations for the Rule of Rescue. Lida Cosmides has approached the issue from a different but complementary perspective (Cosmides, Barrett et al. 2010). Cosmides also argues that the human brain is optimised to process verbal and qualitative data rather than numerical or quantitative data. Although she gives group cohesion as the overall evolutionary driver, the detection of lying and sharing information about those who renege on social contracts are the specific activities requiring the adaptation.

The role of narrative and story-telling in science communication has been explored by several researchers. *The Relation of Story Structure to a Model of Conceptual Change in Science Learning*, (Klassen 2010) takes this a step further by describing a fundamental story structure of an initial state and a final state connected by a transformative event. Klassen then describes the use of this structure to emphasise the causal relationships taught in science. His observation "The idea that sequentially connected events are likely also causally connected is often a natural assumption." (p310, Klassen 2010) reinforces Pratchett et al.'s comments on the power of narrative imperative (Pratchett, Stewart et al. 1999).

Risks are frequently described in numerical and probabilistic terms which means that research into risk perception and risk communication includes material on the perception and communication of numerical and statistical information. One of the most important pieces of research into decision-making in the face of uncertainty was

published in 1974 by Amos Tversky and Daniel Kahnemann. *Judgement under uncertainty: heuristics and biases* (Tversky and Kahnemann 1974) challenged the common assumption underlying much economic theory, that humans act as rational, self-interested agents. They describe three heuristics, or shortcuts for making decisions in the face of uncertainty which effectively bypass careful, rational weighing of evidence: the representativeness heuristic; the availability heuristic; and the anchoring and adjustment heuristic (Tversky and Kahnemann 1974).

The representativeness heuristic involves relating the uncertainties about which you are making a judgement to a class of similar entities or a template and making a choice based on your perception of that class. The main systemic error or bias arising from this heuristic is that it is easy to overlook the inherent probability of different outcomes. In *Judgement under uncertainty* Tversky and Kahnemann give the example: "Steve is very shy and withdrawn, invariably helpful, but with little interest in other people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail." (p2, Tversky and Kahnemann 1974) When asked to assess the probability of Steve belonging to a range of professions including salesman, farmer and librarian, the representativeness heuristic led people to say that it was most likely that Steve was a librarian despite there being relatively few librarians compare to the other professions (Tversky and Kahnemann 1974). Ideas of representativeness also affect perceptions of randomness, if a sequence appears well-ordered it is less likely to be accepted as being randomly generated. The representative template for randomness is scattered, not ordered.

The availability heuristic describes situations where the perceived likelihood of something is influenced by the ease with which it is brought to mind. As things which occur with high frequency are usually easy to recall, this heuristic is often useful. The main bias associated with it is that if something is famous or very memorable, it is easy to recall and hence easy to overestimate how likely it is. Other factors affecting ease of recall are salience and how recently something has been brought to notice.

The anchoring and adjustment heuristic, briefly mentioned on page 38, refers to how people sometimes make estimates by starting from an initial value, or anchor, that is then adjusted to yield the final answer. The anchor may be suggested by the formulation of the decision, by partial computation or by existing knowledge. The

adjustments are usually insufficient, with different starting points leading to different estimates.

One way to look at these three heuristics is to see them as ways of turning situations involving uncertainty and probability into stories which are closely enough connected to our worldview that we can evaluate and act on them.

Tversky and Kahnemann continued their work on decision making, looking at cognitive and psychophysical influences on decision making in risky and riskless situations. Their next major piece of work introduced the idea of prospect theory as an alternative to the then prevailing model of expected utility theory (Kahnemann and Tversky 1979). They found two major effects which were at odds with the then prevailing theories: people underweighted potential outcomes that were merely probable when compared to outcomes that were certain (certainty effect) and generally discarded outcome components that were shared by all outcomes under consideration (isolation effect). The former leads to risk aversion in choices involving sure gains and risk seeking in choices involving sure losses. For example, the certainty of a reward of $3,000 would be preferred to an 80% chance of a reward of $4,000 and an 80% chance of a loss of $4,000 would be preferred to the certainty of losing $3,000. The latter leads to inconsistent choices depending upon how the choice is framed. Their alternative model assigned values to relative gains and losses rather than final outcomes.

Their final major publication, *Choices, values and frames* (Kahnemann and Tversky 1984) investigated the concepts of utility and value in two distinct senses: experience value, the degree of pleasure or pain, satisfaction or anguish in the actual experience of an outcome; and decision value, the contribution of an anticipated outcome to the attractiveness of an option or choice. In the economic model of the ideal, rational decision maker experience value and decision value are the same. Tversky and Kahnemann looked at factors which contribute to major differences between experience value and decision value. They found that expectations about how positive or negative an outcome would be and comparisons of outcomes within social groups were important factors. How a choice was framed affected not only the decision parameters but the experience of the consequences of the decision. (Kahnemann and Tversky 1984).

This seminal research led to Kahnemann being awarded the Nobel prize in economics in 2002. As this was six years after Tversky had died, the prize was not shared although Kahnemann himself refers to it as a joint award.

By describing the influences on decision-making in uncertain or risky circumstances, Kahnemann and Tversky showed the influence of communication and perception on how uncertain choices are interpreted and acted upon. Of central importance to this part of their work is the concept of framing and its effects.

After Tversky's death, Kahnemann continued his writing about decision-making in *Thinking, Fast and Slow*, published in 2011. He picks up on the importance of narrative, commenting on how well people respond to stories which are emotionally and associatively coherent (Kahnemann 2011). The key additional material in this later work is an examination of the role that emotion plays in the understanding of intuitive judgements and choices. His model describes two different types of thinking which underlie our judgements and choices. System One is fast, instinctive and requires little explicit paying attention or effort. There are two versions of System One, the first is innate and the second involves expert knowledge that can be accessed extremely fast. System Two is slow, involves careful reasoning from evidence and explicit attention to the task. System One is impulsive, System Two is controlled.

System One is fast because it uses heuristic shortcuts which are reliable enough to be useful but are still subject to a range of systemic biases. Using System 2 requires conscious effort and enough self control to realise that the first, instinctive response to a situation or question may not be the most useful or appropriate.

Another key researcher and writer in the field of risk communication is Paul Slovic. Over the past 40 years he has published a huge body of work relating to risk perception and risk communication. This ranges from early work on the gambler's fallacy, anchoring and adjustment (Lichtenstein and Slovic 1971) to more recent collaborations with researchers such as Dan Kahan on the relationships between science literacy, technical reasoning ability and acceptance of climate change (Kahan, Peters et al. 2012). Of particular relevance to communicating numerical information is his research group's work with Ellen Peters on the links between numeracy and risk-based decision making. More abstract work on this was published in 2006 (Peters, Västfjäll et al. 2006) and research explicitly about the effect of

numeracy on healthcare decisions in 2007 and then reprinted as a chapter in the 2010 book *The Feeling of Risk: New perspectives on Risk Perception* (Peters, Hibbard et al. 2010).

Numeracy and Decision Making looks at how the ability to understand and transform information expressed as probabilities relates to how well people perform on judgment and decision-making tasks. The research found that highly numerate people were less susceptible to framing effects and likely to draw stronger and more precise affective meaning from numerical information. This did not always lead to better decision-making (Peters, Västfjäll et al. 2006).

Numeracy Skill and the Communication, Comprehension and Use of Risk-Benefit Information (Peters, Hibbard et al. 2010) examines the nature of numeracy and the role it plays in healthcare decisions as well as what the best practices are for the presentation of numeric health information. It draws on the earlier research into numeracy, and then explores the effects of different ways of presenting information to people with different levels of numeracy. It also touches on the ethics of framing the information provided in ways that overtly influence patients' healthcare decisions. Their overall findings for the presentation of information included only showing directly salient information and reducing the amount of inference and calculation needed to process the information. Visual displays increase both comprehension and risk perception as does presenting absolute rather relative risks.

The idea of there being two different ways of processing information and interacting with other people has a strong basis in other psychology research. The concept of cognitive as opposed to affective processing was formally developed in Noam Chomsky's work in the late 1950s although the idea of 'thinking with your head' or 'thinking with your heart' has a much longer history in Western culture. Cultural references to the importance of helping individuals for emotional reasons such as friendship and loyalty can be found from ancient history to Star Trek. There are echoes of this distinction in Dunbar's work on social group sizes and the various bodies of research trying to explain the Rule of Rescue (see page 25).

An additional theory dealing with different levels of human reasoning is known as Fuzzy Trace theory. This theory, formulated in the early to mid-1990s, has its origins in domains of education research and cognitive development with particular emphasis on learning and memory formation. It posits two, parallel systems of

processing, storing and retrieving information: gist and verbatim. Gist representations of events are primarily qualitative and relate to underlying meanings; verbatim representations of events are precise, detailed and often include quantitative information (Reyna and Brainero 1995). When fuzzy trace theory is applied to the field of judgement and decision making, a key concept is that people usually rely on gist rather than verbatim information when weighing evidence for a decision (Reyna 2008). There are clear parallels between these two processing systems and Kahnemann's System One and System Two.

Using the constructivist model of communication to put the way mathematics and statistics are taught together with Kahnemann's concepts of System One and System Two thinking gives some insight into the problems with understanding statistics in general and numerical performance indicators in particular. The fast instinctive version of System One goes straight to the heuristic shortcuts described above, with all their potential for systemic errors. Even if someone is familiar enough with mathematics and statistics to have access to the expert version of System One, the way we learn about numbers has laid a foundation of unhelpful heuristics tied to the concrete rather than the abstract.

A common communication technique which draws strongly on the constructivist model, is to try to make an abstract or remote concept more understandable by relating it to the audience's immediate world of experience, usually by giving an explicit, and ideally familiar, example. In the case of statistical and abstract information this is potentially unhelpful as there may by no useful real-world actualisation of the abstraction. At worst it can be actively misleading as extrapolations that make sense with respect to the real world example might make no sense with respect to the abstract concept. The example of the vacant hospital bed mentioned earlier is an illustration of this. In our day-to-day world an empty bed is available for someone to use it. In the hospital context an empty physical bed is a small, although vital part of the resources constituting an available bed.

Two key statistical concepts that easily get lost in a bid to make information more accessible are that of dependent and independent variables and that of independent trials. A classic example of the first is the way standard population statistics are often represented in the media. At the time census data is published a description such as "the average Canberran is x years old, earns \$nnn per year and has a university

degree" is illustrated by an interview with someone who has all of these precise characteristics. If these variables represent independent characteristics, the composite picture is essentially meaningless. To give an extreme example, take the group of people with the characteristics described in Table 2-3 below and use the mode as the average for category variables and the mean as the average for ratio variables.

Table 2-3: Example of misleading averages

	Hair	Eyes	Sex	Age	Children
Person 1	Blonde	Brown	Female	10	0
Person 2	Dark	Blue	Male	55	4
Person 3	Dark	Brown	Female	35	2
Person 4	Blonde	Blue	Male	65	2
Person 5	Dark	Blue	Male	30	2
Person 6	Dark	Blue	Male	30	2
Person 7	Blonde	Brown	Female	45	1
Average	Dark	Blue	Female	37.1	1.86
Most common combination	Dark	Blue	Male	30	0

In this example, the combination of the averages is not the same as the most common combination and both, seemingly representative constructs have little similarity to most members of the population.

The nonsense that can arise when statistical concepts are treated as real was the basis for a chapter in the children's book *The Phantom Tollbooth* by Norton Juster (Juster 1961). The book's protagonist has a lengthy conversation with 0.58 of a child who is part of an average family consisting of a mother, a father and 2.58 children. Although written to entertain, the conversation reinforces important points about the problems with seeing statistical constructs as part of the real world.

The idea of independent trials is counter-intuitive and as difficult to explain convincingly as it is to understand. Although an infinite number of tosses of a fair coin will land heads up half the time, this is by no means the case over short or even finite numbers of tosses. The coin retains no knowledge of which way up it landed the previous time so, even if it lands heads up ten times in succession, the odds of it landing heads up an 11th time are still 50:50. The same notion applies to shuffling cards or buying a lottery ticket, each unique combination of cards is equally likely

every time they are dealt and not winning the lottery every week for ten years has no effect on the likelihood of winning it this week.

Although these two concepts might seem to be peripheral to using averages as performance indicators, they are fundamental to understanding what we can and cannot know from statistical information. The most critical point is that the statistics for a population do not give any specific information about an individual within that population. This means that performance measures based on population statistics only apply at that level. The median waiting time for a particular type of surgery does not indicate when an individual will be operated on.

Probabilities are often expressed as percentages, for example the chance of rain (see Table 2-2 on page 39). While this is a valid representation, there are inherent difficulties involved in comparing percentages of different sized populations or expressing percentage increases. The idea that a tiny percentage of a very large number is itself a large number can be hard to remember, just as it can be hard to believe that a large percentage of a tiny number is still tiny. When there were scandals associated with the treatment of prisoners by the Americans in the Second Gulf War, one of their commanders proudly declared that 99% of troops were behaving appropriately. Looked at mathematically, there were 200,000 troops deployed at the time which meant the literal meaning of his statement was that some 2,000 troops were not behaving appropriately.

In his papers *Simple tools for understanding risks: from innumeracy to insight* (Gigerenzer and Edwards 2003) and *Helping Doctors and Patients make sense of Health Statistics* (Gigerenzer, Gaissmaier et al. 2008), Gigerenzer looks at the problem of understanding what a positive test for a particular cancer means, given parameters such as false-positive; prevalence and detection rate, all expressed as percentages. His solution is to convert everything into a number of occurrences out of 1000 and then compare those numbers – effectively converting everything to the same denominator.

The second problem is both simpler and harder to deal with. The idea that a 100% increase means double the original number is fairly straightforward, but a 250% increase meaning 3½ times the original number takes System 2 thinking, no matter how familiar you are with this way of expressing amounts.

There is an extra set of conceptual difficulties when you are dealing with numbers relating to rates of change and changes in rates of change e.g. "the inflation rate has doubled". In this case the problem is intimately concerned with the difficulty of actualising the concept at all, compounded by the need to deal with statistical concepts such as averages.

The task of clearly explaining statistics to a broad lay audience is difficult. In the UK the Royal Statistical Society has a series of annual awards for Statistical Excellence in Journalism (Royal Statistical Society 2016). The awards reward and recognise "integrity in their presentation, explanation and use of statistics – avoiding distortion and highlighting the extent of uncertainties." (paragraph 4, Royal Statistical Society 2016).

While researchers such as Gigerenzer offer techniques for dealing with some types of statistical and numerical information and there is a large body of research into the types of difficulties humans have dealing with numbers and statistics, there is no universal, simple set of tools to bridge the gap between numbers and constructed meaning.

2.4.4 Graphs and pictorial communication

A common communication technique for numerical and statistical information is to turn the numbers into a pictorial display such as a graph. Much of the reporting and commentary surrounding elective surgery waiting lists, such as the ACT Public Health Services Quarterly Performance Reports and the AMA's annual Public Hospital Report Card, uses graphs to show comparisons between jurisdictions and other categories as well as changes over time (ACT Health Services Directorate 2010; ACT Health Services Directorate 2010; ACT Health Services Directorate 2010; ACT Health Services Directorate 2011; ACT Health Services Directorate 2011; ACT Health Services Directorate 2011; Australian Medical Association 2011; Australian Medical Association 2011).

In 2001 the second edition of *The Visual Display of Quantitative Information*, the seminal work by statistician and information designer Edward Tufte, was published (Tufte 2001). This beautifully produced book arose from a collaborative series of seminars given by Tufte and statistician John Tukey. The book explores and illustrates the history and practice of using abstract, non-representational pictures to show numbers

(Tufte 2001). According to Tufte "At their best, graphics are instruments for reasoning about quantitative information" (p9, Tufte 2001). The book is divided into two parts, the first reviews the history of data graphics from their inception in the mid-eighteenth century; the seconds is devoted to the then current theory and best practice in their design. At the end of each chapter is a series of points summarising the principles explored in it.

Tufte is passionate about the need to be truthful in the graphical presentation of data and is scathing about what he refers to as 'lying graphics'. He begins by stating that the visual representation of the data and the numerical representation of the data must be consistent and goes on to explore what he means by 'consistent'. (Tufte 2001). He gives some principles for enhancing graphical integrity (i.e. creating graphics that do not lie):

> *The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented.*
>
> *Clear, detailed and thorough labelling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.*
>
> *Show data variation not design variation* (p77, Tufte 2001)

2.4.5 Knowledge transfer

A field of research which has links to research into communication, problem solving and organisational theory is that of knowledge transfer, sometimes referred to as knowledge management. It looks at the practical problems concerning the transfer of knowledge from one part of an organisation to another with a strong emphasis on the knowledge being able to be used and acted upon. Work in this field began in the final years of the 20th century with the rise of what has become known as the knowledge economy. Of most relevance to this thesis is the exploration of what makes knowledge useable and able to inform decisions.

The publication of Thomas A. Stewart's *Intellectual Capital: the new wealth of organisations* (Stewart 1997) in 1997 marked the beginning of research and writing focussed on making the most of the knowledge generated within an organisation. Stewart substantially revised and updated his work in response to the huge uptake of internet

communications at the turn of the 20ᵗʰ century publishing *The Wealth of Knowledge: Intellectual Capital and the Twenty-first-century Organization* in 2001 (Stewart 2001).

A key premise of Stewart's work is that a lot of the knowledge within an organisation is tacit: it is not written down and is usually hard to articulate. This tacit knowledge is critical to an understanding of processes and concepts in an organisation and, in Stewart's opinion, is best communicated by people talking to each other. He gives a light-hearted but still serious example of this in a footnote:

> *It is worth mentioning here, however, the piece of hardware that has been documented as the most effective means of sharing knowledge. It is called a coffeepot. Coffeepots are cross-functional and non-hierarchical. They encourage informal discussion. Serendipitous things happen around them.*
> (p90 Stewart 2001)

In *The Knowing-Doing Gap: how companies turn knowledge into action* Pfeffer and Sutton (Pfeffer and Sutton 2000) examine the reasons why organisations persist with processes and behaviours they know are ineffective. When describing the limitations of systems for collecting storing and retrieving knowledge Pfeffer and Sutton come to a similar conclusion to Stewart's:

> *Knowledge management systems rarely reflect the fact that essential knowledge, including technical knowledge, is transferred between people by stories, gossip, and by watching one another work.*
> (p18 (Pfeffer and Sutton 2000)

When it comes to performance reporting, this tacit or working knowledge encompasses all the assumptions and chains of logic connecting a measurement with a conclusion about performance.

This need for a human element in knowledge transfer reinforces research from other fields from Aristotle's description of pathos as one of the three components of persuasive communication (Aristotle 350BCE) to modern research about the need for people to be able to relate new knowledge to their personal worldview (Yager 1991; Stocklmayer 2013) and the vital importance of storytelling to effective communication (Klassen 2010, Pratchett, Stewart et al. 1999).

2.5 What the literature says about my research topic

When I began my research in mid-2011 there was little overlap between the insights from the different discipline groups canvassed in this literature review. In particular, the application of communication theories to performance indicators had yet to be systematically explored. Research in the period 2011-2016 clarified issues within each of the broad areas but there is still very little synthesis between them.

In the area of performance measurement in general and public sector performance measurement in particular measures are used for a range of sometimes incompatible purposes. Across the literature the two high-level purposes for performance measurement are accounting for the use of public resources and managing public services. These purposes are complicated by the need to have performance indicators based on data that can be accurately, reliably and cost-effectively measured as well as the need to avoid any unintended negative consequences caused by a focus on the measure rather than the overall performance of the system. Underlying performance measurement is an understanding of and agreement on what 'good' performance looks like. In the public health sphere, the three goals of equity, efficiency and perceived fairness are all components of good performance. The difficulty is finding the optimal balance between them. Communicating with and about numerical information has been researched from a variety of standpoints ranging from evolutionary anthropology to risk perception to cognitive psychology. The overall consensus seems to be that it is inherently difficult for humans to use numerical information effectively, particularly for decision making.

My research synthesises theories from the area of performance measurement and evaluation, public health policy and practice, communicating with and about numerical information and to analyse a case study based on communication related to a complex performance indicator in the public health domain.

Having used this chapter to place my work in the wider research landscape, I will use the next chapter to describe the Australian political and funding context within which elective surgery waiting times are used as performance indicators.

3   Political, funding and contextual background to the Australian public hospital system

This chapter describes the political and funding context within which elective surgery waiting times are used as performance indicators for the Australian public hospital system, and the processes for collecting, collating and reporting on elective surgery waiting times. It starts with an overview of the methods used to construct these descriptions from a range of publically-available sources. Chapter 6 of this thesis will draw together critiques of these descriptions and how they are publically communicated.

3.1   Methods

3.1.1   Overview

Reconstructing the political and funding context for the Australian public hospital system in 2010 and 2011 uses the techniques of contemporary history, a term used to describe the study of the recent past, particularly that within living memory (Seldon 1988). As a discipline, it uses mainstream historical techniques such as archival searches, document analysis, and analysis of the oral record where it exists. Its principle advantage over the study of the more distant past is that there is more likelihood that source material will be readily available. Its principle complications are that new material is constantly being added to the evidence base and the possibility of conflict between sources. This conflict is not just between written and oral sources but also between written sources created by different people for different reasons.

3.1.2   Timelines and descriptions

No performance indicator has meaning without context. For it to be meaningful it needs comparators and an understanding of the environment in which it is used. This means that to understand the use of elective surgery waiting times as performance indicators for the Australian public hospital system, they must be put into the context of both the hospital system and the wider political and funding context for the Australian public health system. The context is complicated by responsibility for public health belonging primarily to Australian States and Territories but with much of the funding coming from the Commonwealth level of Government. Details of the

relationship between the different levels of government in Australia and the detailed effect of this on the public hospital system are given in Chapter 3.2.1 starting on page 61.

Information describing the public health system and its governance is available from a range of public documents including:

- Formal agreements between the Commonwealth government and the State/Territory Governments
- Legislation, both State/Territory and Federal
- Policy documents
- Formal Reports
- National and international standards

Many of these documents contain a short summary of the recent history of the Australian public health system as well as details relating to the specific purpose of the document.

The inter-governmental agreements and legislation are the most formal documentation of how accountability and performance measure are used within the public health system but they are by no means the only sources of information. The agreements take many months to negotiate and so are slow to react to changing circumstances. Legislation is similarly slow to change, given the need for formal drafting and multiple parliamentary votes. Neither type of document typically includes operational or purely administrative information. This more detailed information is contained in a range of supporting agreements and administrative orders. The next level of documentation is local policy and implementation documentation which further refines the higher-level guidelines in accordance with state/territory needs.

Table 3-1: Documentation relating to the different levels of governance in the public hospital system

Level of governance	Type of documentation
World Health Organisation	Internationally agreed classification of health interventions
International Standards Organisation	Internationally agreed standards for data collection and management.
Australian institute of Health and Welfare	Nationally agreed standards for data collection and classification of health interventions,

Level of governance	Type of documentation
Australian Commonwealth Government	Commonwealth legislation • Taxation, including Medicare levy and Goods and Services Tax (GST) • National health policy, mainly relating to primary care
Council of Australian Governments	Agreements between the Commonwealth government and State/Territory governments • National Health Reform Agenda • GST distribution Performance reporting of achievement against goals in reform agenda documents
State/Territory governments	State/territory legislation State/territory health policy Consolidated data from hospitals within the State/Territory
Individual Public Hospitals	Day-to-day management, policy and procedural documentation Initial data gathering

When it comes to data definitions and reporting, the highest level specifications are the international standards relating to the collection and reporting of health-related data produced by the World Health Organisation (WHO). These standards underlie the Australian standards and specifications managed by the Australian Institute of Health and Welfare (AIHW). These are described in more detail in Chapter 3.3.4 Data Collection on page 91.

My initial source of documents was the Council of Australian Governments (COAG) Reform Council website (www.coagreformcouncil.gov.au/) and this remained one of my principal sources. The COAG Reform Council is a part of the arrangements for financial relations between the Commonwealth government and the States/Territories. The time of this research coincided with the National Health Reform Agenda, a process managed by COAG to improve the entire Australian public health system and reallocate funding responsibilities between the Commonwealth and State/Territory Governments. The formal reports and submissions relating to the reform agenda contained references leading back to the underpinning legislation, data definitions and high-level policy documents relating to the public hospital system in Australia. From these I was able to identify those that dealt specifically with elective surgery waiting times.

As well as this top-down approach starting at the Commonwealth level, I also started at the local level, searching the ACT Health Directorate website for policy and procedures relating to elective surgery in public hospitals. Links at this level led both up to the wider Commonwealth/State level as well as down to detailed local procedures and implementation guidelines.

From the information in the source material I synthesised two accounts relating to the background and context in which elective surgery waiting times are used as performance indicators for the public hospital system:

- **The Australian public hospital system**

  This describes how the Australian public hospital system is managed and funded within the context of Commonwealth/State relations and the effects of the National Health Reform Agenda. It describes how and when elective surgery waiting times are used as performance indicators in this context.

- **Elective Surgery Waiting List data**

  This describes the lifecycle of the waiting list data from formal data definitions, through the overall administrative processes for elective surgery in the ACT public hospital system, to the first official public reporting of the indicators.

During 2010-2011, the time period covered by this research, the political and funding environment for the Australian public hospital system was in a constant state of flux. There was major reform of how funding and accountability were allocated between the States/Territories and the Commonwealth and instability in the Commonwealth government which eventually lead to a change in that level of government in 2013. The two possible approaches for this part of my research project were to:

- choose a point in time and provide an accurate, detailed description of the system as it was then; or
- document the evolving nature of the overall environment during the research period.

I chose the latter, primarily because the publication of elective surgery waiting time data was one of the key indicators used to drive and monitor the reform agenda. The

period covered is the same as that covered by the articles in *The Canberra Times*, January 2010 to December 2011.

Creating an account of historical events from the documentary record was not straightforward. The process ended up as a feedback loop from discovering information, fitting it into the narrative, finding gaps or contradictions in the narrative and searching for further information. This process is admirably described by Ludmilla Jordanova in her book *History in Practice*:

> *It is unhelpful to think about historical research as a simple sequence of tasks that should be performed in a given order. In fact, historians constantly move between the main types of activity they perform, namely, engaging with sources, delineating a problem, setting it in broad contexts, developing a framework and constructing arguments in written form.*
> (p159, Jordanova 2006)

Although all the source documents contributing to the narrative were in some sense "official" each had been created by a specific entity for a specific purpose. This meant that each had its own particular perspective on elective surgery waiting times, public hospitals, the wider public health system, and the overall Australian political system. The emphasis of the documents could range from a focus on financial matters, to specifications of reporting requirements to details of how the waiting lists themselves were to be managed. Some documents, such as official performance reports, were primarily descriptive; others such as policy proposals produced early in the reform process were primarily persuasive.

Constructing a coherent overview from these disparate sources meant that I had to reconcile a range of subtly different ways of looking at the issues involved. This process is covered at length by Jordanova in *Chapter 7: Historians' Skills* of *History in Practice* (Jordanova 2006). As she observes, "…in reality the important skills lie in tracking down information and knowing how to deploy it thoughtfully rather than in remembering it." (p151, Jordanova 2006). This quote emphasises the importance of the link between straight content and the interpretation of that content within a particular context.

Where possible I used information that I could find in two or more sources. When I had to make a decision about how to represent something that was shown differently

in different sources I chose the description that fitted best with the material around it. I kept firmly in mind that I was researching this from a communication perspective, so I also gave priority to any source material that shed light on how the performance indicators were to be communicated and interpreted. As an example, when covering the establishment of the National Health Performance Authority, I focussed on its reporting role and the State/Territory negotiations about how data was to be presented and compared, rather than the details of the debate about how the authority was to be funded. This business of selecting sources based on context is outlined in Anthony Seldon's Preface to Contemporary History:

> *Above all the contemporary historian must beware the seductive temptation of regarding*
> *any data as objective and final: there are always vital contextual questions to be asked.*
> (p2, Seldon 1988)

Inevitably I dealt with both primary and secondary sources. As mentioned above, many of the later reports and supporting documentation gave short summaries of one aspect or another of the background and history of the elective surgery waiting list data. The common points of these summaries formed the basis of my timelines. Cross-checking these secondary accounts against the available primary documents gave me a way of validating some details.

As well as the formal documentary record there is also an unwritten history stored in the memories of the people involved at all levels in the provision of elective surgery in the Australian public hospital system. As Nicholas Cox points out, "We can only find on file what was put there at the time" (p82, Cox 1988). There will always be details of processes and decisions that are not captured in either the confidential or the public record because they arose from informal discussions and exchanges. What ever is written later about the basis of these decisions will always be a selection and interpretation of what happened. Since my focus is on the public aspect of communication about elective surgery waiting lists, not the communication between informed individuals within the system, I deliberately used only sources available to a determined member of the public. As I became more informed I discovered more sources but still kept to those on the public record.

The most common medium for these public documents was official and semi-official internet sites. The rise of the internet as a means of making information available to the public has greatly changed the nature of research into contemporary history.

While searching government websites might be quicker than searching hard-copy archives, sending off for printed reports and submitting Freedom of Information requests it comes with its own set of limitations. As Brian Head points out in *The Public Service and government communication: Pressures and dilemmas* (Head 2007) it is all too easy for information provision to become entangled with government marketing. This makes it all the more important to look for information from non-government sources and independent statutory bodies as well as that provided directly by government departments.

Using only public sources meant that I dealt with the same information as that presented to the public and could focus on analysing it and its communication consequences. Collecting the oral history associated with the background to the use of elective surgery waiting times as performance indicators would be a major piece of research in its own right and is beyond the scope of this research project.

3.2 The Australian public hospital system

This part of Chapter 3 provides a description of the political and funding context within which elective surgery waiting times are used as performance indicators for the Australian public hospital system. The description is synthesised from a range of sources using the methods outlines in Chapter 3.1.

3.2.1 Overview of governance and funding

Australia has a universal public health system the broad principles of which are described in the introduction to the National Healthcare agreement.

> *This National Healthcare Agreement affirms the agreement of all governments that Australia's health system should:*
>
> - *be shaped around the health needs of individual patients, their families and communities;*
>
> - *focus on the prevention of disease and injury and the maintenance of health, not simply the treatment of illness;*
>
> - *support an integrated approach to the promotion of healthy lifestyles, prevention of illness and injury, and diagnosis and treatment of illness across the continuum of care; and*

- *provide all Australians with timely access to quality health services based on their needs, not ability to pay, regardless of where they live in the country.* (paragraph 4, COAG Reform Council 2011)

The funding and governance of the Australian public hospital system was under review in the period covered by the case study, 2010-2011. The debate effectively started with a big push for national hospital reform by the Commonwealth Labor government in 2010. This account will go up to the end of the reform process but will not include any changes made by the Commonwealth Liberal government elected in mid-2013.

To understand the funding and governance of the Australian Public Hospital System it is necessary to understand a little of Australia's broad political system and its history.

Australia was formed as an independent nation on 1 January 1901 when six British colonies, New South Wales, Queensland, South Australia, Tasmania, Victoria and Western Australia formed a federation. Legislation was passed in the British Parliament to allow this federation to govern in its own right as the Commonwealth of Australia (Australian Government n.d.). This national government is referred to as either the Commonwealth government or the Federal government. Elections for this level of government are always referred to as Federal elections. The colonies changed from being separate colonies answering to Britain to states of the new Commonwealth. By 2010 the Commonwealth of Australia was made up of the six original states, three self-governing territories (Australian Capital Territory, Norfolk Island and Northern Territory), and seven territories governed directly by the Commonwealth. Sections 51 and 122 of the Australian Constitution define the issues upon which the Commonwealth government can pass laws for the benefit of the nation (Australian Government n.d.). In practice the two onshore territories, the Australian Capital Territory and the Northern Territory, act as states with similar powers and responsibilities.

Matters governed by Commonwealth law include defence and foreign affairs, trade commerce and currency, immigration, postal services, telecommunications and broadcasting, air travel, and most social security and pensions (Parliament of New South Wales n.d.). Personal income tax is also under the jurisdiction of the Commonwealth. The major responsibilities of the states and territories are schools;

hospitals; conservation and environment; roads, railways and public transport; public works; agriculture and fishing; industrial relations; community services; sport and recreation; consumer affairs; police; prisons; and emergency services (Parliament of New South Wales n.d.). The Commonwealth is involved in some State responsibilities such as health and education, as a provider of funding. Negotiations between the Commonwealth and State/Territory governments are done through the Council of Australian Governments (COAG).

The main source of revenue for the States is the Goods and Services Tax (GST), a broad-based tax of 10% on most goods, services and other items sold or consumed in Australia (Australian Taxation Office 2012). It is collected at the point of sale and distributed to the states and territories according to the recommendations of the Commonwealth Grants Commission (CGC) (Commonwealth Grants Commission 2015). The goal of the CGC recommendations is to ensure that each State has the capacity to provide services at national average levels at average levels of efficiency. The underlying concept is that citizens in different States should have access to equal standards of government services. This process is referred to as Horizontal Fiscal Equalisation (HFE) (Department of the Parliamentary Library 2002). The States and Territories have full control over how they spend the GST distributed to them, they are not strictly bound to providing those services used in the modelling and calculations related to the distribution process.

Medicare is a universal health insurance scheme funded by the Commonwealth. it provides free or subsidised healthcare services to the Australian population. It is partly funded by revenue raised by a levy of 2% on all taxable income, the balance comes from consolidated Commonwealth revenue. Medicare collection is done by the Australian Taxation Office and Medicare payments are administered by the Commonwealth Department of Human Services. According to their website, Medicare covers:

- *free or subsidised treatment by health professionals such as doctors, specialists, optometrists and in specific circumstances dentists and other allied health practitioners*

- *free treatment and accommodation as a public Medicare patient in a public hospital, and*

- *75% of the Medicare Schedule fee for services and procedures if you are a private patient in a public or private hospital. This does not include hospital accommodation and items such as theatre fees and medicines.*
  (paragraphs 2-5, Department of Human Services 2015)

The benefits payable under the legislation covering Medicare are documented in the *Medicare Benefits Schedule* (Australian Government Department of Health 2014). Typically, the Medicare benefit for a health service is set at a percentage of the Scheduled Fee, an amount determined by the Government to be appropriate for the service. There is no requirement for private practitioners to limit their charges to the schedule fee and many charge far more. The principle challenge for public hospitals when dealing with Medicare is that payment is made after the fact on the basis of services provided.

The complexity of the funding and governance of the Australian health system and the fragmentation of responsibilities was identified as a major weakness and source of inefficiency in a 2015 OECD report (OECD 2015).

There are four main political parties at both the Commonwealth and State levels of government in Australia:

- Australian Labor Party.
  Centre-left party, formally linked to the Australian labour movement.
  Held government at the Commonwealth level 2007-2013.
- Liberal Party
  Centre-right party, typically in favour of lower taxes and smaller government.
  Typically governs in coalition with the National Party.
- National Party
  Founded to represent the interests of rural Australia, originally called the Country Party.
  Typically governs in coalition with the Liberal Party.
- The Greens
  Formed more recently than the other parties. Frequently holds the balance of power in government, particularly in the upper houses such as the Senate.

Relations between the Commonwealth and the States/Territories are complicated by the potential for different parties to hold power at the two levels of government. The States and Territory governments are often dominated by the party not in power at the Commonwealth level, particularly if one party has held power at the Commonwealth level for several electoral cycles. At the beginning of 2010 the Labor party was in its first term of government at the Commonwealth level after several terms of Liberal/National party government. Table 3-2 below shows which political parties were in power at the State and Commonwealth level between January 2010 and December 2011. The changes of government in Victoria in November 2010 and New South Wales in March 2011 changed the balance of the negotiations between the Commonwealth and the States.

Table 3-2: Political parties in power in Australia 2010 and 2011

Jurisdiction	Time Period	Governing Party
Commonwealth	January 2009-December 2011	Australian Labor Party
Australian Capital Territory	January 2009-December 2011	Australian Labor Party
New South Wales	January 2009-March 2011 March 2011-December 2011	Australian Labor Party Liberal/National Party
Northern Territory	January 2009-December 2011	Australian Labor Party
Queensland	January 2009-December 2011	Australian Labor Party
South Australia	January 2009-December 2011	Australian Labor Party
Tasmania	January 2009-December 2011	Australian Labor Party
Victoria	January 2009-November 2010 November 2010 to December 2011	Australian Labor Party Liberal/National Party
Western Australia	January 2009-December 2011	Liberal Party

3.2.2 COAG Hospital Reform Agenda timeline

Public hospital waiting times for elective surgery were seen as an emerging problem in early years of the 21st century. In December 2009 The National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan (COAG Reform

Council 2009) was signed by the Commonwealth State and Territory Governments. Under this agreement, reward funding was tied to States/Territories meeting elective surgery waiting time targets.

The hospital reform process gained new momentum at the beginning of 2010 with a Federal election due that year. Public health costs were rising faster than the State and Territories' ability to meet them and the public health system was perceived to be in decline. According to a series of speeches marking Australia Day 2010 by then Prime Minister, Kevin Rudd, health care costs had risen by 11% in the previous five years while state revenues had only risen by 4% in the same period (Massola 2010).

In March 2010 the Commonwealth began negotiations through COAG with the States and Territories to reform the public hospital system, moving from a State-based system to a more national system. A key element of the initial proposal was that the States and Territories forgo a proportion of their GST distribution in return for the Commonwealth explicitly meeting more of the costs associated with the public hospital system. The percentages of additional funding; the amount of GST to be redistributed and the extent to which the states would remain in charge of their hospital systems all changed as the negotiations continued. The initial proposal was for 65% of the funding for public hospitals to come directly from the Commonwealth in return for the States and Territories forgoing 30% of their GST income.

At the COAG meeting in April 2010 all States and Territories except for Western Australia, the sole Liberal party-governed state, agreed in a communiqué to new healthcare funding and governance arrangements with the Commonwealth which would come into effect at the beginning of July 2011. Under the changes the Commonwealth would become the dominant funder of the public hospital system which would be run by a new system of regional networks. The Commonwealth would fund 60% of the national efficient price (as calculated and agreed by states and territories) for public hospital services provided to public patients, capital, research and training in public hospitals. This would initially be paid for by a redirection of some of the GST paid to the States and Territories to pay explicitly for healthcare, based on the calculated cost of providing services. Over the longer term the agreement meant that the Commonwealth would cover the cost increases in the provision of public hospital services (COAG 2010).

The lead up to the Federal election held in August 2010 featured proposals and counter-proposals by the two major political parties on the issue of public hospitals with both sides claiming to have a plan that would fix the system. The Labor party was re-elected but had to form a minority government, weakening its ability to pass the required legislation to implement any changes to the public hospital system.

In December 2010 the Commonwealth launched a website called MyHospitals which published data for public and private hospitals about what services were available, and what the waiting times were for elective surgery and emergency department treatment. This early version of the website was widely criticised by some States and Territories for using out of date or inappropriate data (Cronin 2010).

In the lead up to the COAG meeting in February 2011 both Victoria and Western Australia were resisting the need to hand over a proportion of their GST revenue as part of the reform of the public hospital system. Western Australia had not agreed to the original proposal in April 2010 and Victoria was threatening to withdraw its agreement. A new proposal was taken to COAG where States and Territories would keep their GST revenue and the Commonwealth would match all new funding for hospitals. The proposal outlined greater transparency for funding decisions and hospital management. Key parts of the transparency provisions were the establishment of a National Health Performance Authority and an improved and expanded MyHospitals website (COAG 2011). This agreement, known as the National Health Reform Agreement -- National Partnership Agreement on Improving Public Hospital Services (COAG Reform Council 2011) was signed on 13 February 2011.

Under the agreement, the National Health Performance Authority was to be established by Commonwealth Legislation as of 1 July 2011. Its role would be to "…develop and produce reports on the performance of hospitals and health care services, including primary health care services" (COAG 2011). Although the Heads of Agreement was signed by all States and Territories, Victoria continued to oppose the establishment of the National Health Performance Authority. This opposition was finally withdrawn in June 2011. A major point of contention was over the question of who would be considered to be the manager of a state hospital system. Victoria fought for and gained the right for states to be notified of under-

performance and given the opportunity to address it before the information was made public.

In July 2011 the Commonwealth government set new targets, tied to reward funding, for elective surgery waiting times. There were two parts to the new targets: the first was for all elective surgery to be carried out within clinically recommended time frames; the second was aimed at reducing the number of people who had already waited longer than clinically recommended. Table 3-3 below summarises the timeline for the COAG Hospital Reform Agenda.

Table 3-3: Basic Timeline for COAG Hospital Reform Agenda 2010-2011

Date	Event
December 2009	National Partnership Agreement on Elective Surgery Waiting List Reduction Plan signed by Commonwealth, State and Territory governments.
January 2010	Prime Minister Kevin Rudd flags health, education and infrastructure as the three main priorities for the Commonwealth government in 2010.
March 2010	Commonwealth government releases plan for reform of public health. Proposed that 65% of funding for public hospitals would be paid directly by the Commonwealth with the finding coming from redirection of 30% of GST revenue.
	Strong objection from the States and Territories over the trade-off
	Start of Federal election campaign with both major parties nominating public health as a major issue.
April 2010	All States and Territories except Western Australia agree to a new healthcare funding package with the Commonwealth.
November 2010	Change of government in Victoria from Labor Party to Liberal/National Party coalition.
December 2010	MyHospitals website reporting performance measures for public hospital throughout Australia launched.
February	New COAG Heads of Agreement signed by the Commonwealth

Date	Event
2011	and States and Territories.
March 2011	Change of government in New South Wales from Labor Party to Liberal/National Party coalition.
July 2011	National Health Performance Authority established.
	Final version of *National Health Reform Agreement – National Partnership Agreement on Improving Public Hospital Services* signed by the Commonwealth, State and Territory governments.
August 2011	Publication of final performance report for the National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan.

3.2.3 Key public health funding and governance agreements

When the *National Health Reform Agreement -- National Partnership Agreement on Improving Public Hospital Services* (COAG Reform Council 2011*)* was signed in July 2011 the funding of Australia's public hospital system was governed by a network of agreements between Australia's State and Territory governments and the Commonwealth government. Each agreement had its own funding, outcomes and performance monitoring arrangements. The major agreements were:

- *Intergovernmental Agreement on Federal Financial Relations* (COAG 2011). Contained schedules relating to performance reporting and payment arrangements for all funding flowing from the Commonwealth to the States and Territories.

- *National Healthcare Agreement* (COAG Reform Council 2011). An agreement between the Commonwealth government and the States and Territories relating to the provision of all types of public health care. The agreement was regularly re-negotiated based on emerging needs and performance under previous agreements.

- *National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan* (COAG Reform Council 2009). The third of three agreements between the Commonwealth government and the States and Territories explicitly relating to reducing elective surgery

waiting times. They ran from 2009 to 2011 and had detailed performance measures with financial rewards for meeting elective surgery performance targets. This agreement was due to run out shortly after the signing of the *National Health Reform Agreement – National Partnership Agreement on Improving Public Hospital Services* which replaced some of its provisions.

- *National Health Reform Agreement – National Partnership Agreement on Improving Public Hospital Services* (COAG Reform Council 2011).
  An agreement between the Commonwealth government and the State and Territory governments intended to improve public patient access to elective surgery, emergency department and sub-acute care in public hospitals.

3.2.4 *National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan*

This agreement was an initiative of the Council of Australian Governments (COAG) under the *National Healthcare Agreement*. The waiting list Agreement was implemented in three stages, each covered by its own National Partnership Agreement between the Australian States and the Commonwealth.

The formal outcome of the three waiting list Agreements was to "reduce the number of Australians waiting longer than clinically recommended times for elective surgery by improving efficiency and capacity in public hospitals." (paragraph 3, COAG Reform Council 2009).

Funding under the agreements was to be in three stages:

- Stage One provided $150 million to bring about an immediate reduction in the number of people waiting longer than the clinically recommended time for elective surgery.
- Stage Two provided $150 million for system and infrastructure improvements that will improve elective surgery performance in the long-term.
- Stage Three provided funding of up to $300 million for:
  - Part 1: Meeting jurisdiction specific elective surgery volume targets;
  - Part 2: Exceeding the jurisdiction specific elective surgery volume targets set in Part 1; and
  - Part 3: Improved elective surgery waiting list management to achieve the following outcomes:

- a reduction in the number of patients ready for care who have waited longer than clinically recommended;
- maintain or improve the median and 90th percentile; and
- maintain or improve the percentage of patients seen within the clinically recommended time by urgency category.

(COAG Reform Council 2009)

The Waiting List Reduction Plan agreements were designed to contribute to the following specific objectives:

a) *an efficient and effective public hospital system that is able to adapt to the pressures of rising health costs and increasing demand;*

b) *improved health outcomes and patient experience and satisfaction;*

c) *integration between the hospital system and other health services;*

d) *targeting of services; and*

e) *smooth patient transitions between health settings through assessment, referral and follow up at key points throughout the healthcare system.*

(paragraph 10, COAG Reform Council 2009)

The objectives and outcome of the agreements were to be achieved by reducing waiting times for elective surgery in public hospitals; increasing the number of elective surgery procedures undertaken in public hospitals and improving the management of waiting lists to achieve greater efficiency.

Under the agreements, there were three assessment and reporting periods:

- Period 1: 1 July 2009 to 31 December 2009
- Period 2: 1 January 2010 to 30 June 2010
- Period 3: 1 July 2010 to 31 December 2010

Each target in Stage Three of the plan had reward funding attached to it and if a jurisdiction did not meet a target it was paid a percentage of the reward based on the percentage of the target that was achieved.

Table 3-4: State and Territory Maximum Reward Funding July 2009—December 2010 (adapted from Table A2, COAG Reform Council 2009)

State	Period 1	Period 2	Period 3
NSW	$21,104,359	$14,509,247	$14,509,247
VIC	$16,021,709	$11,014,925	$11,014,925
QLD	$12,529,910	$8,614,313	$8,614,313
WA	$6,284,992	$4,320,932	$4,320,932
SA	$5,074,090	$3,488,437	$3,488,437
TAS	$1,554,009	$1,068,382	$1,068,382
ACT	$937,246	$644,356	$644,356
NT	$493,685	$339,408	$339,408
Australia	**$64,000,000**	**$44,000,000**	**$44,000,000**

The final assessment report for the National Partnership Agreement on the elective Surgery Waiting List Reduction Plan was published in August 2011 (COAG Reform Council 2011). Tables 3-5 to Table 3-10 below show the performance of the states and territories against the three performance components for Stage Three of the plan. The tables are adaptations of those on pages 9-17 of the final assessment report (COAG Reform Council 2011). In the final assessment report the tables were inconsistently formatted with some have state by state data in rows and some having it in columns. This inconsistency made it extremely difficult to compare data from different tables. My adaptation consistently puts the independent variable (state) in column 1 with the dependent variable (usually target and performance) in columns two and three.

Table 3-5: Performance against target number of surgeries by State (Part 1, 1 June 2009 – 31 December 2010)

State	Target	Performance
NSW	306,228	310,809
VIC	207, 079	231,593
QLD	186,980	195,580
WA	108,406	116,763
SA	66,017	70,368
TAS	21,668	24,616
ACT	14,619	15,242
NT	8,392	9,177

Australia	**919,389**	**974,148**

All jurisdictions met their elective surgery volume, or Part 1, targets for each of the three reporting periods in Stage Three of the plan and so were paid the maximum reward funding for that part of the agreement.

The assessment for Part 2 of the agreement was based on two factors, whether or not the jurisdictions achieved their Part 1 targets and the cost-weighted volume of elective surgery admissions for all three periods (1 June 2009 – 31 December 2010). In the event that a jurisdiction did not meet a Part 1 target their unallocated reward funding was to be added to the pool for Part 2 reward funding. Since all states and territories met their volume targets, there was no unallocated Part 1 reward funding. Table 3-6 below shows the volume and proportion of cost-weighted elective surgery admissions by states which was used to distribute the Part 2 funding. Part 2 funding was also linked to Part 3 funding under a complicated set of rules. If a jurisdiction met its targets for Part 3A, it was paid 100% of its Part 2 funding. If it did not meet its Part 3A targets, it was assessed against its Part 3C targets to determine a reduced Part 2 payment. If it did not meet its Part 3C targets there was still a possibility for a further reduced payment if it met its targets for Part 3B.

Table 3-6: Volume and proportion of cost-weighted elective surgery admissions by State (1 June 2009 – 31 December 2010)

State	Cost-weighted Volume	Proportion
NSW	450,273	30.9%
VIC	344,152	23.6%
QLD	324,555	22.3%
WA	143,681	9.9%
SA	125,595	8.6%
TAS	37,036	2.5%
ACT	24,208	1.7%
NT	8,750	0.6%
Australia	**1,458,250**	**100.0%**

As detailed above, Part 3 of the assessment related to the management of waiting lists and had three sets of targets within it. For the first target of reducing the number of patients ready for care who have waited longer than clinically recommended some

jurisdictions had numerical targets and some had percentage targets. Jurisdictions which already had a relatively low rate of patients waiting longer than clinically recommended (<10%) were essentially required to maintain this level of performance to the end of the assessment. Table 3-7 below shows state by state results for the first component of the Part three assessment.

Table 3-7: Patients ready for care who have waited longer than the clinically recommended time by state at 31 December 2010

State	Target	Performance	Target Achieved
NSW	<420	178	Yes
VIC	≤10%	14.2%	No
QLD	≤10%	8.5%	Yes
WA	<10%	10.2%	No
SA	<506	322	Yes
TAS	≤1,334	4,169	No
ACT	≤566	2,006	No
NT	≤291	619	No

The second Part three target was to maintain or improve the median and 90th percentile waiting times for people who had been removed from the waiting list. Removal from the waiting list usually means that the patient has had their surgery although there are other reasons such as moving to the private hospital system, moving to a different state or being no longer able to maintain ready for care status due to decreasing health. Table 3-8 and Table 3-9 below show state by state performance against these targets for the assessment period.

Table 3-8: Median number of days patients have been waiting at removal from waiting list by state at 31 December 2010

State	Target	Performance	Target Achieved
NSW	≤40	44	No
VIC	≤33	34	No
QLD	≤26	27	No
WA	≤28	28	Yes
SA	≤39	35	Yes
TAS	≤48	35	Yes
ACT	≤73	70	Yes

NT	≤42	34	Yes

Table 3-9: 90th percentile number of days patient have been waiting at removal from waiting list by state at 31 December 2010

State	Target	Performance	Target Achieved
NSW	≤279	330	No
VIC	≤216	178	Yes
QLD	≤132	148	No
WA	≤184	153	Yes
SA	≤218	194	Yes
TAS	≤491	333	Yes
ACT	≤377	381	No
NT	≤307	226	Yes

The third component of the Part three target was maintaining or improving the percentage of patients seen within the clinically recommended time for each of three urgency categories:

- Category 1 – urgent, surgery within 30 days;
- Category 2 – semi-urgent, admission within 90 days; and
- Category 3 – non-urgent admission within 1 year

Table 3-10 shows state by state performance for each urgency category against these targets.

Table 3-10: Percentage of patients seen within the clinically recommended time by state and urgency category at 31 December 2010

State	Category 1			Category 2			Category 3		
	Target %	Perve.	Target Met	Target %	Perf.	Target Met	Target %	Perf.	Target Met
NSW	≥93	93%	Yes	≥79	90	Yes	≥96	91	No
VIC	100	100%	Yes	≥70	76	Yes	≥91	93	Yes
QLD	≥83	85%	Yes	≥82	74	No	≥89	93	Yes
WA	≥88	89%	Yes	≥77	82	Yes	≥96	97	Yes
SA	≥80	88%	Yes	≥78	90	Yes	≥89	95	Yes
TAS	≥72	74%	Yes	≥46	62	Yes	≥62	73	Yes
ACT	≥94	89%	No	≥45	46	Yes	≥74	75	Yes
NT	≥78%	85%	Yes	≥58	61	Yes	≥76	80	Yes

3.2.5   *National Health Reform Agreement -- National Partnership Agreement on Improving Public Hospital Services*

The *National Health Reform Agreement – National Partnership Agreement on Improving Public Hospital Services* (COAG Reform Council 2011) was signed in July 2011 and followed on from the National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan (COAG Reform Council 2009). It covered public patient access to three major aspects of the public hospital system:

- elective surgery,
- emergency department care, and
- subacute care.

Subacute care is care relating to a patient functioning and quality of life rather than care related to a specific medical condition. Examples of subacute care are rehabilitation care, palliative care and geriatric rehabilitation and management care.

The agreement will expire on 30 June 2017 unless terminated or extended by the parties to the agreement. It follows on from the National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan (COAG Reform Council 2009).

The desired output of the agreement relating to elective surgery is "a higher proportion of elective surgery patients seen within clinically recommended times, and a reduction in the number of patients waiting beyond the clinically recommended time." (paragraph 16a, COAG Reform Council 2011) Under the agreement, $200 million in reward funding was made available to States and Territories for meeting performance goals. The goals were in two parts:

- Part 1 involved stepped improvement in the proportion of patients in each urgency category seen within the clinically recommended timeframes (See Table 3-11 on page 77). The improvements were relative to a baseline of 2010 performance unless agreed differently on a case by case basis.
- Part 2 involved a progressive reduction in the number of patients who were overdue for surgery with particular emphasis on those who had waited the longest (See Table 3-13 on page 78). This target was expressed in terms of the average number of days surgery was overdue, calculated by summing all days overdue for an urgency category and dividing by the number of patients.

The three smallest jurisdictions, Tasmania, the Australian Capital Territory and the Northern Territory were given an extra year to meet their targets because they had less opportunity for efficiency gains related to scale. Table 3-11 below shows the elective surgery improvement targets for each state and Table 3-12 below shows the associated reward funding for this part of the agreement. Table 3-11 combines three table from the agreement document with reformatting to allow easier comparison of data between states.

Table 3-11: Elective surgery on-time improvement targets by urgency category 2012-2016 (adapted from Tables A5, A6 and A7, COAG Reform Council 2011)

State	Cat.	Baseline (2010)	End 2012	End 2013	End 2014	End 2015	End 2016
NSW	1	92.3%	96.0%	100.0%	100.0%	100.0%	
	2	86.6%	90.0%	93.0%	97.0%	100.0%	
	3	89.4%	92.0%	95.0%	97.0%	100.0%	
VIC	1	100.0%	100.0%	100.0%	100.0%	100.0%	
	2	72.5%	79.0%	86.0%	93.0%	100.0%	
	3	91.9%	94.0%	96.0%	98.0%	100.0%	
QLD	1	83.0%	89.0%	100.0%	100.0%	100.0%	
	2	74.8%	81.0%	87.0%	94.0%	100.0%	
	3	88.1%	91.0%	94.0%	97.0%	100.0%	
WA	1	87.4%	94.0%	100.0%	100.0%	100.0%	
	2	79.2%	84.0%	90.0%	95.0%	100.0%	
	3	97.2%	98.0%	99.0%	99.0%	100.0%	
SA	1	87.5%	94.0%	100.0%	100.0%	100.0%	
	2	87.6%	91.0%	94.0%	97.0%	100.0%	
	3	95.5%	97.0%	98.0%	99.0%	100.0%	
TAS	1	75.4%	84.0%	92.0%	100.0%	100.0%	100.0%
	2	59.3%	67.0%	76.0%	84.0%	92.0%	100.0%
	3	76.8%	81.0%	86.0%	91.0%	95.0%	100.0%
ACT	1	91.8%	95.0%	97.0%	100.0%	100.0%	100.0%
	2	44.1%	55.0%	66.0%	78.0%	89.0%	100.0%
	3	76.9%	82.0%	86.0%	91.0%	95.0%	100.0%
NT	1	79.1%	83.0%	94.0%	100.0%	100.0%	100.0%
	2	56.9%	59.0%	74.0%	83.0%	91.0%	100.0%
	3	81.6%	84.0%	89.0%	93.0%	96.0%	100.0%

For all States and Territories the urgency category with the lowest baseline proportion of cases seen on time and hence the greatest rate of improvement required over the period of the agreement was Category 2 Non-urgent.

Table 3-12: Notional reward funding for Part 1 of the National Elective Surgery Targets (Table A3, COAG Reform Council 2011)

	2012-2013	2013-2014	2014-2015	2015-2016	2016-2017	Total
NSW	$7.9 m	$7.9 m	$7.9 m	$7.9 m		**$31.5 m**
VIC	$6.2 m	$6.2 m	$6.2 m	$6.2 m		**$24.7 m**
QLD	$5.3 m	$5.3 m	$5.3 m	$5.3 m		**$21.1 m**
WA	$2.7 m	$2.7 m	$2.7 m	$2.7 m		**$10.9 m**
SA	$1.7 m	$1.7 m	$1.7 m	$1.7 m		**$6.8 m**
TAS	$0.4 m	$0.4 m	$0.4 m	$0.4 m	$0.4 m	**$2.2 m**
ACT	$0.3 m	$0.3 m	$0.3 m	$0.3 m	$0.3 m	**$1.7 m**
NT	$0.2 m	$0.2 m	$0.2 m	$0.2 m	$0.2 m	**$1.2 m**

Reward payments for Part 1 were split evenly between the three urgency categories. To receive the reward payment a State of Territory must have:

- At least maintained the volume of elective surgery and not dropped below the 2010 baseline; and
- Achieved an improvement of at least half the difference between the previous year's target and the current year's target.

If a jurisdiction achieves an increase of 98% or higher on the previous year's target it would receive 100% of the reward funding. If a jurisdiction exceeded the target for a period it would receive 100% of the reward payment.

Table 3-13: Elective surgery targets for average number of overdue surgeries by urgency category 2012-2016 (Table A8, COAG Reform Council 2011)

State	Cat.	Baseline (2010)	End 2012	End 2013	End 2014	End 2015	End 2016
NSW	1	0	0				
	2	39	29	20	10	0	
	3	130	98	65	33	0	
VIC	1	0	0				
	2	129	97	65	32	0	
	3	165	124	83	41	0	
QLD	1	18	0				

State	Cat.	Baseline (2010)	End 2012	End 2013	End 2014	End 2015	End 2016
	2	89	67	45	22	0	
	3	81	61	41	20	0	
WA	1	27	0				
	2	90	68	45	23	0	
	3	87	65	44	22	0	
SA	1	31	0				
	2	30	23	15	8	0	
	3	45	34	23	11	0	
TAS	1	138	69	0			
	2	356	285	214	142	71	0
	3	440	352	264	176	88	0
ACT	1	45	23	0			
	2	179	143	107	72	36	0
	3	246	197	148	98	49	0
NT	1	67	34	0			
	2	97	78	58	39	19	0
	3	144	115	86	58	29	0

The reward funding for Part 2 of the agreement was based on the proportional reduction in the number of patients who have waited longer than the recommended time. Urgency Category 1 cases were all to be seen within clinically recommended times by the end of 2012. In addition, the 10% of patients within each urgency category who have waited the longest must have their procedures in the reporting year (COAG Reform Council 2011).

Table 3-14: Notional reward funding for Part 2 of the National Elective Surgery Targets

	2012-2013	2013-2014	2014-2015	2015-2016	2016-2017	Total
NSW	$7.9 m	$7.9 m	$7.9 m	$7.9 m		**$31.5 m**
VIC	$6.2 m	$6.2 m	$6.2 m	$6.2 m		**$24.7 m**
QLD	$5.3 m	$5.3 m	$5.3 m	$5.3 m		**$21.1 m**
WA	$2.7 m	$2.7 m	$2.7 m	$2.7 m		**$10.9 m**
SA	$1.7 m	$1.7 m	$1.7 m	$1.7 m		**$6.8 m**
TAS	$0.4 m	$0.4 m	$0.4 m	$0.4 m	$0.4 m	**$2.2 m**
ACT	$0.3 m	$0.3 m	$0.3 m	$0.3 m	$0.3 m	**$1.7 m**
NT	$0.2 m	$0.2 m	$0.2 m	$0.2 m	$0.2 m	**$1.2 m**

3.3 Collection and reporting of elective surgery waiting list data

This part of Chapter 3 provides a description of the processes by which elective surgery waiting time data is defined, collected and reported. The description is synthesised from a range of sources using the methods outlined in Chapter 3.1.

Data specific to elective surgery waiting times starts with data used to manage elective surgery waiting lists at the hospital level which is extracted and used in several levels of reporting, each of which feeds into the level above it. The number of reporting levels depends on the size of each state health system. As a very small jurisdiction, the ACT has only two levels of reporting: hospital and territory-wide. Larger jurisdictions may have an interim level of reporting based on local areas (e.g. south-east New South Wales). State and Territory report feed into the national reporting. Key data definitions are set at the national level to ensure as much consistency as possible in reporting between different jurisdictions.

A factor which complicates reporting about elective surgery is that it is carried out in the private hospital system as well as the public hospital system. This means that waiting times reported for the public hospital system as part of the National Health Reform Agenda reflect only some of the activity in the overall health system relating to elective surgery. Procedures carried out in the private system but funded by the public system are classified as part of the public system. There is no clear or consistent mechanism for identifying patients who move between the systems or the reasons for their doing so.

3.3.1 Data definitions
Elective Surgery

There is no explicit definition of "Elective Surgery" in the *National Healthcare Agreement* and the various iterations of the *National Partnership Agreement Elective Surgery Waiting List Reduction Plan*. Instead there is a reference to the *National Health Data Dictionary* (Australian Institute of Health and Welfare 2010).

The (Australian) National Health Data Dictionary (NHDD) (Australian Institute of Health and Welfare 2010) is produced and maintained by the Australian Institute of Health and Welfare (AIHW). The dictionary's purpose is to provide national standards for the health reporting framework. The standards are designed to support

both statistical and clinical analysis of data with the aim of providing meaningful input into community discussion and public policy debate on health issues in Australia. The current version of the NHDD was released in August 2010. This version conforms to international standard ISO/IEC 11179 Information Technology Metadata Registries (MDR), 2004 (International Standards Organisation 2004).

According to the NHDD:

> *Elective Surgery is defined as any surgery that a patient's doctor or health professional considers to be necessary but which can be delayed by at least 24 hours.*

The Medicare Benefits Schedule is an exhaustive list of services provided with full or partial public funding. The services relating to elective surgery are spread between three of the main categories:

- Category 1: Professional Attendance,
- Category 2: Diagnostic Procedures and Investigation, and
- Category 4: Therapeutic Procedures.

Surgical procedures are listed as Group T8 of Category 4 in the schedule, taking up over 170 pages. They are divided into 16 Subgroups according to the broad type of surgery:

- Subgroup 1: General
- Subgroup 2: Colorectal
- Subgroup 3: Vascular
- Subgroup 4 - Gynaecological
- Subgroup 5 - Urological
- Subgroup 6 - Cardio-Thoracic
- Subgroup 7 - Neurosurgical
- Subgroup 8 - Ear, Nose And Throat
- Subgroup 9 - Ophthalmology
- Subgroup 10 - Operations For Osteomyelitis
- Subgroup 11 – Paediatric
- Subgroup 12 - Amputations
- Subgroup 13 - Plastic And Reconstructive Surgery
- Subgroup 14 - Hand Surgery

- Subgroup 15 – Orthopaedic
- Subgroup 16 - Radiofrequency Ablation

Subgroups may have multiple groups within them, for example Subgroup 3: Vascular is composed of 13 smaller groups and Subgroup 15: Orthopaedic has 21 subgroups. The subgroups themselves can be made up of tens of individual procedures, each with a calculated agreed efficient cost and rebate, (pages 341-514, Australian Government Department of Health 2014).

The way the procedures are divided into categories and listed means that the costs associated with a particular instance of surgery may be spread throughout the Medicare Benefits Schedule.

Urgency Category

There are three clinical urgency categories used in public hospitals to classify elective surgery patients. In the NHDD these are defined as:

- Urgency Category 1: Admission within 30 days desirable for a condition that has the potential to deteriorate quickly to the point where it may become an emergency.
  This is sometimes referred to as "Urgent"
- Urgency Category 2: Admission within 90 days desirable for a condition causing some pain, dysfunction or disability but which is not likely to deteriorate quickly or become an emergency.
  This is sometimes referred to as "Semi-urgent"
- Urgency Category 3: Admission at some time in the future acceptable for a condition causing minimal or no pain, dysfunction or disability, which is unlikely to deteriorate quickly and which does not have the potential to become an emergency.
  This is sometimes referred to as "Non-urgent"
- In some documents Urgency Category 3 is defined as:
  Admission within 365 days for a condition causing minimal or no pain, dysfunction or disability, which is unlikely to deteriorate quickly and which does not have the potential to become an emergency.

The second definition is the one used for reporting on Category 3 patients who have waited longer than clinically recommended.

Ready for Care/Not Ready for Care

A patient on the Elective Surgery Waiting List must be classified as being:

- Ready for Care: prepared to be admitted for the awaited procedure or treatment or to begin the pre-admission process.

- Not Ready for Care: not in a position to be admitted or to begin the pre-admission process for the awaited procedure or treatment. Reasons include:
  - Unfit for surgery (clinical): the patient's health has temporarily declined to the point where it is inadvisable to proceed with the awaited procedure
  - Staged procedure (clinical): there is a planned clinical pathway that requires a predictable series of treatments, each depending upon the successful completion of the last.
  - Deferred Procedures (personal) the patient is not yet prepared to be admitted for the proposed procedure/treatment for social, work or other commitments.

There is a limit to the cumulative number of days a patient can be not ready for care for personal reasons. The limits for the urgency categories are 30 days for Category 1; 90 days for Category 2 and 180 days for Category 3. If a patient chooses to exceed the limit for their urgency category they can be removed from the waiting list.

Calculation of a patient's waiting time includes only the time a patient spends "Ready for Care".

3.3.2 Differences in interpreting data definitions

At the heart of the agreements using elective surgery waiting times as performance indicators is the comparison of performance between different states and territories. For these comparisons to be valid, clinicians in each state and territory need to interpret the definitions underlying the statistics in exactly the same way.

At the time the National Health Reform Agenda began in 2010 the details of how data definitions were interpreted differed between different states and territories. This became clear during the early negotiations about performance indicators and reward funding with the allocation of urgency categories being an area where the differences were most apparent.

In 2012 AIHW and the Australian College of Surgeons started the process of refining the definitions relating to elective surgery urgency categories and developing "agreed national elective surgery urgency category definitions (including for patients not ready for care) that will enable consistent application across all states and territories" (Slide 2, Australian Institute of Health and Welfare and Royal Australasian College of Surgeons 2012).

Although determining a patient's urgency category is a clinical decision, a state by state analysis of the distribution of urgency categories for elective surgery admissions showed differences between jurisdictions that were greater than could be accounted for by random fluctuations. Table 3-15 below shows the variations in the proportion of elective surgery admissions in each urgency category for each state and territory (Slide 4, Australian Institute of Health and Welfare and Royal Australasian College of Surgeons 2012).

Table 3-15: Elective surgery admissions by urgency category and state

State	Category 1	Category 2	Category 3	Total
NSW	26%	31%	43%	100%
VIC	28%	48%	24%	100%
QLD	37%	47%	16%	100%
WA	26%	35%	39%	100%
SA	33%	34%	32%	100%
TAS	41%	43%	17%	100%
ACT	29%	49%	42%	100%
NT	43%	38%	19%	100%
Total	30%	40%	30%	100%

These differences could also be seen at the level of particular surgical specialisations. For example, the differences between the allocation of urgency categories for hip and knee replacement surgery showed a wide variation between different states and territories. In New South Wales and South Australia 25% of hip replacements were given an urgency category of 2 while in the ACT 85% were given that urgency category (Slide 6, Australian Institute of Health and Welfare and Royal Australasian College of Surgeons 2012).

3.3.3 Elective surgery process overview

The processes for managing elective surgery are complex and different entities within the public health system interact with these processes in different ways. Some of the more important points of view are:

- Medical: patients, carers and primary care providers such as general practitioners; surgeons and other specialists with the authority to refer someone for surgery;
- Management: hospital managers and administrators;
- Performance Reporting: those involved in the collation and reporting of data relating to elective surgery; and
- Political: those concerned with the politics of providing public hospital services. This can include lobby and advocacy groups as well as elected political representatives.

From the medical perspective, the typical processes for a patient joining, being on and leaving the public hospital system elective surgery waiting list are:

1. A patient goes to their general practitioner with a problem.

2. After assessment the general practitioner refers the patient to an appropriate surgical specialist. It can take from days, to weeks, to months to get an appointment with a specialist. In the ACT patients are referred to a specific specialist with their own waiting list. In other states such as Queensland, patients are referred to a pool of specialists with a joint waiting list appropriate to their medical condition.

3. The specialist assesses the patient and, if deemed necessary, assigns an urgency category and refers the patient to the hospital for surgery. Approximately 30% of patients referred to a surgical specialist are not recommended for surgery and so do not join the elective surgery waiting list.

4. The patient remains on the waiting list as long as they are deemed ready for care (see above). If their medical condition changes their surgeon conducts a clinical review and alters the urgency category if necessary.

5. The patient is removed from the waiting list once they have had their surgery. If a patient's condition declines to the extent that they are deemed to be permanently 'Not Ready for Care' they are removed from the waiting list.

The Australian Medical Association (AMA) argues that the time between being referred to a specialist and seeing a specialist should be counted as time waiting for elective surgery (Australian Medical Association 2011). The counter argument to this is that 30% of patients referred to specialists do not end up requiring surgery. The first task of the specialist is to assess whether or not surgery is needed.

From the management perspective of the hospital, the process centres around managing the access of many patients with different needs, ready for care statuses and surgical requirements. The task is complex, with the need to manage resources such as theatre space; availability of people such as surgeons with appropriate specialisations, nursing staff, and anaesthetists; and bed space in wards.

The key policy document used in the ACT public hospital system is the ACT Elective Surgery Access Policy. The information below comes from the 2011 version of the policy which was based in the original 2007 policy with modifications incorporating feedback from stakeholders (ACT Health 2011). The policy is underpinned by 12 principles:

- *Referrals for elective surgery are clinically appropriate and represent a suitable treatment for the patient's condition*

- *Patients are provided with easy to understand information about access to elective surgery and their rights and responsibilities*

- *Public patients are the shared responsibility of the hospital, the referring surgeon and the relevant specialty*

- *Patients waiting for elective surgery are fully informed about, and have consented to, the procedure/treatment*

- *All documentation is complete, legible and accurate*

- *Waiting list management services are provided in an efficient, transparent and patient centred manner*

- *The elective surgery waiting list is managed to ensure patients are treated equitably within clinically appropriate timeframes and with priority given to patients with an urgent clinical need*

- *The scheduling of surgery is undertaken in consideration of available capacity*

- *Hospitals minimise the impact and inconvenience to patients whose surgery they postpone*

- *The elective surgery waiting list is managed to promote the most effective use of available resources*

- *When a surgical specialty is unable to cope with increased demand, the hospital will be informed to escalate options for the patient*

- *There is valid, reliable and accountable reporting of access to elective surgery.*
  (p4 , ACT Health 2011)

After the more general introductory sections, the policy is presented as a series of four clinical processes (Referring Patients for Elective Surgery, Elective Surgery Categorisation, Timeliness of Elective Surgery, Doctor's Leave) and seven administrative processes (Elective Surgery Accountable Officer, Registering Patients on the Waiting List, Managing Patient Status, Scheduling Patients for Surgery (Admission Process), Postponement of Surgery, Removing Patients from the Waiting List) (ACT Health 2011). Each process lists the principles which apply to it, a

statement of policy specific to the process, guidelines for implementing the policy and statements of responsibility for each stakeholder who is part of the process (ACT Health 2011). Table 3-16 below shows which principles underpin each process. One principle, 'When a surgical specialty is unable to cope with increased demand, the hospital will be informed to escalate options for the patient' is not explicitly linked to any of the processes in the policy. While the office holder to whom the issue is to be escalated is specified, there is no information about what actions are available to ensure that an escalated case is treated. The only action specified is in the case of repeated problems with the timeliness of surgery for a particular surgeon in which case the action is a review of their session allocations (p14, ACT Health 2011).

Table 3-16: Principles underpinning elective surgery processes

Principle \ Process	Referring Patients for Elective Surgery	Elective Surgery Categorisation	Timeliness of Elective Surgery	Doctor's Leave	Elective Surgery Accountable Officer	Registering Patients on the Waiting List	Managing Patient Status	Scheduling Patients for Surgery	Postponement of Surgery	Removing Patients from the Waiting List	Record Keeping
Referrals for elective surgery are clinically appropriate and represent a suitable treatment for the patient's condition	X										
Patients are provided with easy to understand information about access to elective surgery and their rights and responsibilities	X				X	X	X		X		
Public patients are the shared responsibility of the hospital, the referring surgeon and the relevant specialty	X		X								
Patients waiting for elective surgery are fully informed about, and have consented to, the procedure/treatment	X										
All documentation is complete, legible and accurate	X					X					
Waiting list management services are provided in an efficient, transparent and patient centred manner					X	X	X			X	
The elective surgery waiting list is managed to		X	X		X			X		X	

Principle \ Process	Referring Patients for Elective Surgery	Elective Surgery Categorisation	Timeliness of Elective Surgery	Doctor's Leave	Elective Surgery Accountable Officer	Registering Patients on the Waiting List	Managing Patient Status	Scheduling Patients for Surgery	Postponement of Surgery	Removing Patients from the Waiting List	Record Keeping
ensure patients are treated equitably within clinically appropriate timeframes											
The scheduling of surgery is undertaken in consideration of available capacity			X					X	X		
Hospitals minimise the impact and inconvenience to patients whose surgery they postpone									X		
The elective surgery waiting list is managed to promote the most effective use of available resources				X	X						
When a surgical specialty is unable to cope with increased demand, the hospital will be informed to escalate options for the patient											
There is valid, reliable and accountable reporting of access to elective surgery.											X

There are two types of review for patients on the waiting list:

- Clinical Review: this is required to change the patient's urgency category
- Administrative Review: this is used to manage the access of patients to a surgery date appropriate to their urgency category.

A Clinical Review is conducted by the referring surgeon or other appropriate clinician and is triggered by a deterioration in the patient's condition. The list of appropriate clinicians includes the relevant registrar, senior registered nurse, named delegate of the referring surgeon or a nominated officer (ACT Health 2011).

Rules for triggering an Administrative Review are laid out in the section relating to timeliness of elective surgery in the ACT Elective Surgery Access Policy 2011 (p13, ACT Health 2011) and reproduced below:

Table 3-17: Triggers for administrative review of patient status

Patient Category	Administrative Review	Trigger	Action
Category 1: admission within 30 days	On receipt of Request for Admission	No surgical booking date. Surgery not possible within 30 days.	Discuss dates and other options with surgeon. Book surgery. If no dates, escalate.
Category 2: admission within 90 days	On receipt of Request for Admission	Surgery not possible within 90 days. No date scheduled after 60 days	Confirm patient still requires surgery. Discuss dates and other options with surgeon. Book surgery. If no dates, escalate.
Category 3: admission within 365 days	Bi-monthly	No date scheduled after 270 days	Confirm patient still requires surgery. Discuss dates and other options with surgeon. Book surgery. If no dates, escalate.

Table 3-17 above illustrates the tension between the clinical requirements described by the urgency category and the capability of a hospital to fulfil those requirements.

3.3.4 Data Collection

The performance data relating to elective surgery is drawn from the detailed information collected by individual public hospitals. For each patient a range of data is collected as part of the routine administration relating to elective surgery. The key

data set for each patient is given below with details from the procedures for the ACT hospital system. As discussed in Chapter 3.3.5 *Formal Reporting* on page 92 there are subtle variations between jurisdictions in the detailed interpretation of data definitions.

Key patient data used to calculate elective surgery waiting time performance indicators:

- Date added to waiting list
  A patient is registered as being on the waiting list within three days of a Request for Admission (RFA) form being received by the hospital. The date of registration is the date the form was received. Patients are advised, in writing, that they have been added to the list.

- Surgical Specialty
  This data is collected at the time a patient is referred for surgery.

- Urgency Category and date assigned
  The initial Urgency Category is assigned as part of the process of referral for elective surgery. Urgency Category is updated as required by clinical review which may occur as a routine part of patient care or be trigged by an administrative review (See Table 3-17 on page 91)

- Ready for Care status and date assigned
  A patient can only be referred for surgery and put on the waiting list if they are 'ready for care'. This is updated as required by clinical review

- Date of leaving waiting list

- Reason for leaving waiting list

Other data, such as the number of days a patient has been on the waiting list can be calculated from the other fields.

3.3.5 Formal Reporting

There were several layers of formal reporting relating to elective surgery, starting with individual hospitals and extending to national waiting time data.

The first step in the reporting cycle was for individual public hospitals to extract data from their systems and report to their state-level health department. Next state-level health departments aggregated data from the hospitals in their jurisdiction and sent it to the Australian Institute of Health and Welfare (AIHW) up until July 2011. After

then it was sent to the National Health Performance Authority (NHPA). Some jurisdictions also published their own performance reports publically. For larger states these public reports were prepared for regions within the state as well as for the state as a whole.

The central organisation (AIHW or NHPA) calculated the performance indicators for each state from the state-level data. Under the agreement establishing the NHPA the results of these calculations were then sent back to the states for verification and amendment or explanation of special circumstances where necessary. Data and performance indicators for individual hospitals were published on the MyHospitals website (National Health Performance Authority 2016) and formal performance reports submitted to COAG (COAG Reform Council 2011).

Some of the key indicators were:

- Median waiting times
    - by urgency and specialisation for a hospital
    - aggregated by urgency and specialisation for a state/territory
    - all specialisations aggregated by urgency for a state/territory
- Number of people on waiting list:
    - by urgency and specialisation for a hospital
    - aggregated by urgency and specialisation for a state/territory
    - all specialisations aggregated by urgency for a state/territory
- Number of people receiving surgery
    - by urgency and specialisation for a hospital
    - aggregated by urgency and specialisation for a state/territory
    - all specialisations aggregated by urgency for a state/territory
    - all urgency categories combined for a state/territory
    - aggregated by urgency category for Australia as a whole
    - all urgency categories combined for Australia as a whole
- Percentage of patients seen within clinically recommended time
    - by urgency category for a state/territory
    - all urgency categories combined for a state/territory
    - aggregated by urgency category for Australia as a whole
    - all urgency categories combined for Australia as a whole

The performance reports to COAG formed the basis for the negotiation between the Commonwealth and the states/territories of reward funding under the National Health Reform Agreement (see Chapter 3.2.4 on page 70).

This chapter described the collection and synthesis of source material relating to the political, funding and administrative background to the use of elective surgery waiting times as performance indicators for the Australian public hospital system. It presented the material as two descriptions: one based around the overall political and funding context and one focussing on administrative procedures and data collection. In the context of the Australian public hospital system, elective surgery waiting times are used as performance measures in agreements between the Federal government and the states and territories. In an effort to improve the public hospital system in general and the provision of elective surgery in particular, states and territories were set targets relating to elective surgery with reward funding available to those jurisdictions which met the targets. The next chapter will cover the methods used to collect and analyse elective surgery-related newspaper articles from The Canberra Times for the period 1 January 2010 to 31 December 2011.

4 Newspaper analysis methods

This chapter describes the methods used to collect and analyse material relating to elective surgery published in a daily newspaper. For collecting material from the newspaper I used techniques relating to the study of contemporary history outlined in Chapter 3.1 on page 55. To uncover patterns in the articles and later in the instances of elective surgery waiting time being used as a performance indicator I used techniques from Richard Boatzis's *Transforming Qualitative Data: Thematic Analysis and Code Development* (Boyatzis 1998). I chose Boyatzis's techniques because they allow for an iterative approach to coding which allowed me to continually refine my codes as I learned more about the data. A consequence of this iterative approach is that there has been some blurring of the boundaries between method, results and analysis for this part of my research.

4.1 Overview

Public hospital services in general and the provision of elective surgery in particular are reported in the news media as matters of public interest. For closer analysis I chose the ACT's local newspaper, *The Canberra Times* as a widely read and readily accessible source of reporting and commentary. According to the website of its publishing company, Fairfax Media (Fairfax Media 2013) *The Canberra Times* has a circulation of 28,614 as at 2013. This is down from the 2011 figure of 31,521 for weekday editions (Media Spy 2013). It is the only ACT-based daily newspaper and it is a broadsheet containing a mix of local, national and international news. Since the ACT is the seat of the Australian Commonwealth Government, with a high percentage of its population working in government-related jobs, *The Canberra Times* has a strong focus on political aspects of the news.

4.2 Collecting articles from *The Canberra Times*

When I started my research in June 2011, *The Canberra Times*, maintained an online archive of the articles it published each day. Although the newspaper website was restructured several times over the next few months the archive remained available and easily accessible until early 2012, after this time it was essentially unavailable. During that six months, I downloaded a full list of all the articles published in 2010 and 2011, a time of great change in the funding arrangements for Australia's public hospital system.

The online archive was not a perfect reflection of what had appeared in the print version of *The Canberra Times*. Some articles referred at the end of the text to more complete versions in the print edition of the newspaper. Rather than trying to incorporate material from both the online and print version of the newspaper I have confined my data to the articles in the newspaper's online archive. This had the advantages of having the material all in an electronic format from the start, and saving the work of cross-checking between print and online material and deciding which version should take precedence.

From the data in the online archive I compiled a spreadsheet listing all articles and their metadata from the website. Between June 2011 and January 2012, the archive listed articles by month with publication date, headline, byline and first sentence. There was also a short indication of which area of the newspaper the article had appeared in (e.g. SPORT-GENERAL). Although there was a category of NEWS-HEALTH it was only applied to one article in the 2010 collection and 79 articles in the 2011 collection, many of which were not related to health funding, the public hospital system, or elective surgery waiting lists.

Typically the structure of a newspaper article follows a clear pattern of topic sentence; overview paragraph; detail and comment. This meant that I could use a combination of the headline and the first sentence of each article to manually identify those relating to public hospital funding and elective surgery waiting lists. I was deliberately broad in my interpretation of which articles were relevant because it was safer to collect more widely than necessary and then narrow the selection after examining the full text of the articles rather than risk not collecting relevant material. For example, I collected an article relating to the establishment of a $15m GP clinic in the ACT as it was not clear from the headline and first sentence whether or not it would relate to my areas of interest. The article was subsequently excluded as the full text showed it to be outside the scope of my research. In all, five articles were excluded from the initial full text collection because, on reading the full text, they were not related to elective surgery. I ended up identifying 96 articles relating to elective surgery.

Each article flagged as being relevant was assigned a unique identifier to assist with referring to them and tracking changes to the collection and the full text downloaded from the online archive. The identifier had a three-part structure:

1. A: indicated that the unit of analysis was an article
2. Two-digit sequential identifier
3. Year of publication (4-digits)

For example, the first article relating to elective surgery published in 2010 would have an identifier of A012010. A full list of all the articles is in Appendix A.

There were some articles that looked like duplicates of those from a previous day. By comparing the body text I found that some were exact duplicates with different publication times and some were partial duplicates with extra or amended material. For example A252011 on 7 September 2011 is an expanded version of A242011 from 6 September 2011. I removed the earlier versions of the word-for-word duplicates. There were also pairs of articles about the same topic with different emphasis, A322011 and A332011 were published on 7 and 8 December 2011 and both deal with a deal between the ACT government and Calvary private hospital. The earlier article focuses on the health aspects of the arrangement while the later article focuses on the political aspects.

As well as the documents containing the full text of the articles selected for analysis I created a summary dataset containing metadata derived from each article to allow for a wide range of potential analyses:

- Date
- Day of Week
- Title
- Reporter
- Word count
- More information in print version?
- Short summary of major points and comments on its relationship to other articles and events.

The three datasets (listing of all articles from 2010 and 2011; full text of articles relating to elective surgery; and metadata for articles relating to elective surgery) had a range of uses:

- Compiling the lists of articles gave dates for the release of formal reports relating to elective surgery waiting lists. This helped to determine reporting

cycles and relevant agencies as well as providing information about potential additional source documents.

• The article lists gave an indication of what was deemed to be newsworthy about elective surgery waiting lists.

• The full text of the articles gave a record of the reported reaction to reports and developments by various interested parties: politicians, consumer groups, health professionals and patients.

• The full text of the articles gave examples of the language journalists and their sources chose to use when referring to elective surgery waiting lists and the issues surrounding them.

• The metadata for the elective surgery articles provided a structured format suitable for recording the results of thematic analysis.

4.2.1 Descriptive statistics

There were 16,282 articles in the online archive for *The Canberra Times* over the period 1 January 2010 to 31 December 2011. They were unevenly split between the two years, with 41% in 2010 and 59% in 2011, suggesting a change in the way articles were selected for the archive sometime during that period. Graphing the month by month figures as shown in Figure 4.1 below shows the difference between the two years. Looking closely at the graph, there may have been two changes in the archive, a sharp one in May 2010 and a slower change resulting in an increase in the number of articles through much of 2011.
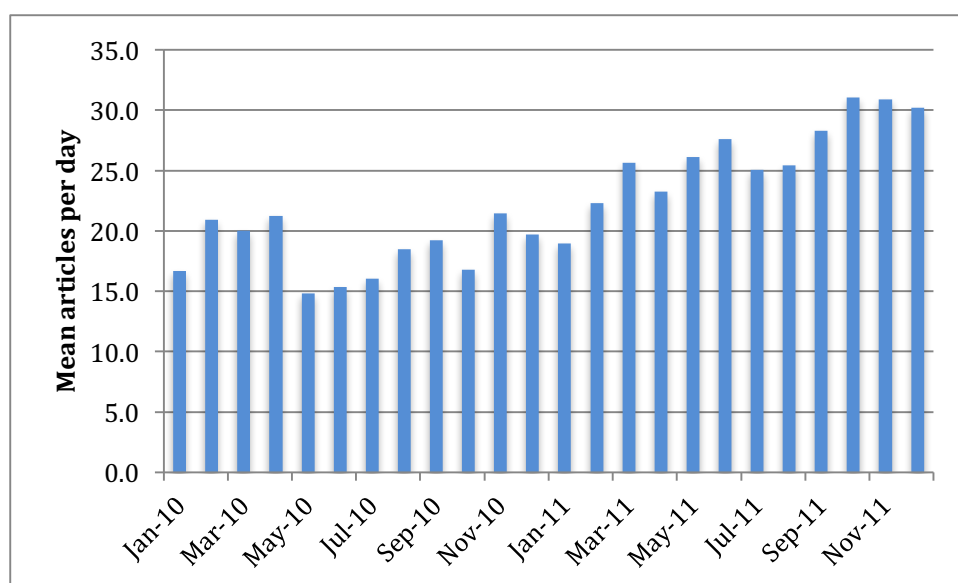
Figure 4.1: Monthly breakdown of mean Canberra Times articles per day

These potential discontinuities in the underlying data meant that is was unsuitable for large-scale statistical analysis. It was suitable for the two main purposes of this research:

- Looking for patterns in the appearance of elective-surgery related articles, either as a raw number each month or as a proportion of the articles published; and
- Using thematic analysis to describe and characterise how elective surgery waiting times are used as performance indicators in newspaper articles.

Because the number of articles selected as relating to elective surgery (96) was small compared to the overall number of articles, the day to day fluctuations appeared disproportionately large. To smooth out these small-scale variations, I grouped the articles according to the month in which they were published. The patterns were very similar whether you looked at the raw number of selected articles each month (Figure 4.2) or those articles as a percentage of the total number of articles published.
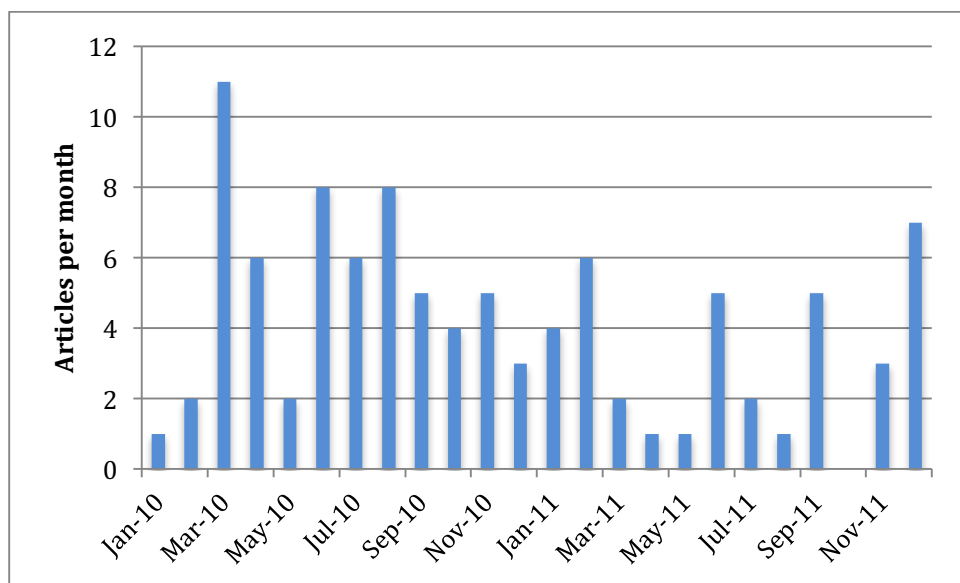
Figure 4.2: Number of elective surgery articles by month

4.3 Matching articles to timelines

Using the number of articles relating to elective surgery as a crude indicator of the newsworthiness of the issue, I compared the distribution of the selected articles to the timelines I had constructed relating to the political and funding context (see Chapter 3.2). I was looking to answer three questions:

1. Were major political and funding events relating to elective surgery deemed newsworthy enough to be reported on?

2. Did reporting on a major political or funding event coincide with other articles about elective surgery?

3. Were articles relating to elective surgery published independently of reporting of major political or funding events?

Answering these questions would provide information about the context within which elective surgery waiting times were being used as performance indicators in communication with the public.

4.4 Article-level thematic analysis

4.4.1 Code development and checking

Based on a combination of my background reading and a first-pass reading of the articles as I was collecting them, I chose the following preliminary coding criteria and read each article in detail to determine which code(s) to assign. Each criterion was phrased as a question which I answered for the article as a whole. Question numbers begin with 'A' to indicate that they related to article-level coding.

Table 4-1: Initial coding criteria for newspaper articles

	Question	Codes
A-1	Is the article explicitly about elective surgery or about the hospital system more generally?	Yes No
A-2	Is the article about an individual's experiences or about elective surgery in general	Individual System
A-3	Is the article about the Australian hospital system or about the ACT hospital system?	Federal ACT
A-4	Does the article explicitly mention 'fairness' or 'equity' in the provision of elective surgery in the public hospital system	Fairness Equity
A-5	Is elective surgery defined in the article?	Yes No

As can be seen, initially these codes were simple binary pairs. After coding the first six month's worth of articles I realised that this choice had been based on incorrect and simplistic assumptions: some articles included both parts of a pair while others had three or more possible answers. Following the concepts of coding for emergent themes as described by Boyatzis (Boyatzis 1998) I modified my codes and started coding again. This process of coding, refining and recoding was repeated until I had a set of coding criteria that were robust enough to use. An indication that a question had been sufficiently refined was that coding became much faster, with few stops to debate which code applied. While I was developing the codes I primarily worked question by question, going through the entire article collection for one question at a time.

Once I had made a full pass through the article collection, coding each article against all the questions, I did a second pass through the articles several weeks later, this time working question by question and not looking at the previous answers. These two coding passes were then subjected to checks first for internal consistency and then for consistency between each pass.

The internal consistency checks for each pass were:

- Check all codes used exactly match those in the final specification
- Check and fill blanks in the coding
- Cross check coding of 'Equity' between Questions A-4 and A-4a (Equity/Fairness) and Question A-7 (Equity/Efficiency).
- Cross-check Question A-2 (Individual/System) and Question A-2a (Positive/Negative).

I then compared the results of the two coding passes, flagging every instance where the codes did not match. There were two main types of error:

- simple error arising from mistakes in either recording the answer or finding explicit content. Some of the simple errors were caused by the interaction between Question A-4a and Question A-7.
- errors arising from a difference in interpretation of the question each time I applied the codes.

Table 4-2 below shows the breakdown of matches, errors and different interpretations for each question.

Table 4-2: Results from checking article-level coding

	Question	Matches	Simple Errors	Different Interpretation
A-1	Is the article explicitly about elective surgery or about the hospital system more generally?	92	4	0
A-2	Is the article about an individual's experiences or about elective surgery in general	89	0	7
A2a	Are stories about individuals positive or negative?	92	4	0
A-3	Is the article about the Australian hospital system or about the ACT hospital system?	84	5	7
A-4	Does the article explicitly refer to either 'fairness' or 'equity' in the provision of elective surgery in the public hospital system?	94	2	0
A-4a	Does the article use the concepts of fairness or equity the public hospital system?	69	0	27
A-5	Is elective surgery defined in the article?	96	0	0
A-5a	Is the idea of 'urgency category' used and either defined or implied in the article?	79	1	16
A-6	Are indicators and their utility discussed	82	0	14
A-7	Does the article refer to the concepts of efficiency and/or equity in the provision of elective surgery in the public hospital system?	80	9	7

For all the codes where a difference in interpretation caused the mismatch, I re-read the article and the coding criteria to decide which was the appropriate code. I then re-wrote the coding criteria more explicitly and reapplied it to all the articles, not just those with a mismatch in codes. This process was repeated until I got the same answers each time I applied the codes.

As an independent check of the robustness of the specifications for the coding criteria, I gave detailed code definitions in the format used for Appendix B and full text of all the articles from The Canberra Times to my supervisor. He chose five articles at random from each year and assigned codes to them based only on the definitions supplied.

Comparing my supervisor's coding with mine showed two systemic differences:

- Questions A4, A4a and A7 referred to perceptions of elective surgery, and there was no explicit mention of elective surgery in the articles. Supervisor marked as N/A. This was simply rectified by revising the specification to cover other services in the public hospital system.

- My supervisor interpreted equity far more broadly than I did and fairness somewhat more broadly. As this was an interpretive issue relating to a concept, I tightened the specifications and re-checked the coding of questions A4a and A7 for every article.

4.4.2 Evolution of coding specifications

Developing the codes was an iterative process with each coding pass and each check feeding into the specifications. The coding specifications for each question also evolved as I learned more about the detailed content of the articles I was working on and became more familiar with the underlying political and policy frameworks. Table 4.1 on page 100 shows my initial coding criteria and Table 4.3 on page 113 shows the final form it evolved into as I worked. For each question, I have documented its progression from its original to its final form. The wording given for each question is its final version. The format of this final version draws heavily on Boyatzis' structure of a useful, meaningful code (Boyatzis 1998).

Question A-1: Is the article explicitly about elective surgery or about the hospital system more generally?

This question did not change during the coding process.

The articles selected for analysis included those referring to the public hospital system, its structure, governance and funding as well as those directly referring to elective surgery. There were also articles about the private hospital system which referred to its interaction with the public system.

This coding question allowed the identification of smaller subset of articles relating explicitly to elective surgery in the public hospital system. It includes anything to do with elective surgery funding; elective surgery waiting times; elective surgery administration; elective surgery facilities; as well as stories relating to individuals and their elective surgery experiences. It does not include stories about surgery in private

hospitals, even if it was 'elective', unless it was done as part of the public hospital system.

All errors found in checking were simple mistakes.

Summary of final coding criteria

Yes

The article refers to elective surgery funding; elective surgery waiting times; elective surgery administration; elective surgery facilities; as well as stories relating to individuals and their elective surgery experiences. Articles relating to procedures performed privately (i.e. privately funded) are not included. Articles referring to treating public patients in private facilities are included.

No

The article is about the broader political, administrative and funding context surrounding elective surgery in the public health system. There is no direct mention of elective surgery.

Question A-2: Is the article about an individual's experiences or about elective surgery in general?

This question did not change during the coding process.

This question is intended to uncover the balance between system-wide reporting and the reporting of stories about individuals. During coding it emerged that there were some stories which, while mainly about the system, used a personal story as an illustration of the impact of the system on people. This showed up as interpretation errors in the checking between coding passes. The criteria were refined so that any mention of an individual's experience in an article mainly about the system and any mention of how the system works in an article mainly about an individual were coded as 'Both'.

Summary of final coding criteria

Individual

The article is centred on an individual's story and any mention of the hospital system is directly related to them and their experiences. Any generalisations about the hospital system in an article about an individual trigger the answer "Both".

System

The article does not mention any individual or their experiences relating to elective surgery.

Both

The article deals with both elective surgery in general and an individual's experiences. This category includes stories using an individual's experience to illustrate a more general story.

Question A-2a: Are stories about individuals positive or negative?

This question was added after the coding the first six month's worth of articles.

There are many ways of using stories about individuals to illustrate the workings of a complex system. Stories could be positive (e.g. the extra funding is great, I've finally had my surgery) or negative (even with all the extra funding and work on reducing waiting lists hasn't made any difference to me). This question aimed to find out if there was a dominant way of framing personal stories with respect to the public hospital system. This meant that if a story about a positive outcome for a patient was negative about the public hospital system it was classified it as 'negative'.

All errors found in checking were simple mistakes.

Summary of final coding criteria

Positive

The article is about a positive outcome for a patient and is positive about elective surgery in the public hospital system.

Negative

The article is either about a negative outcome for the patient and/or is negative about elective surgery in the public hospital system.

Both

The article meets the criteria for both 'positive' and 'negative'.

N/A

The answer to Question A-2 was "System"

Question A-3: Is the article about the Australian hospital system or about the ACT hospital system?

The period 2010-2011 was a time of great change in the funding and policy

framework for Australian public hospitals. These high-level negotiations were widely reported both at the Federal level as well as their likely impact on the ACT. In addition there were stories specifically about elective surgery in the ACT hospital system and the management of health services in the ACT. The initial coding for this question was a simple binary choice: Was the article about Commonwealth-level policy and funding issues or specifically about the ACT health system? After coding articles from the first few months, I added a third category 'Both' to cover articles that dealt with the interaction of the Commonwealth and ACT systems. The terms 'Federal' and 'Commonwealth' are interchangeable for describing national-level government and policy in Australia. I used 'Federal' for coding because it is the shorter term.

Once I had coded the articles from 2010 it became clear that there was another category to be coded. The close interaction between the southern NSW health services and the ACT health services meant that there were issues relating specifically to the ACT region, particularly Queanbeyan and Yass but potentially as far away as Cooma. I renamed the 'Both' category to 'Federal/State' and used it for articles dealing with the interaction of Federal funding and policy with the ACT or ACT Region.

In checking the differences in coding between the passes it became clear that it was difficult to know when to distinguish between the ACT and the ACT region and how they fitted in with articles dealing with the relationship between the Federal government and the ACT. The coding criteria were rewritten much more tightly and an extra category, ACT Region, added.

Summary of final coding criteria

Federal

Federal-level policy and funding, relationship between the States and the Federal Government

ACT

Specifically relating to the ACT, its budget and governance

ACT Region

Relating to services and policy that covers the ACT and the surrounding region, including the treatment of NSW patients in ACT hospitals

Federal/ACT

Relations between the ACT and the federal governments, includes Federal-level initiatives covering the ACT region.

Question A-4: Does the article explicitly refer to 'fairness' or 'equity' in the provision of elective surgery or other services in the public hospital system?

This question was modified after external checking of coding.

In Professor Jenny Stewart's book Public Policy Values (Stewart 2009) she devotes a chapter to the differences between equity and fairness, advancing the idea that fairness is the politics of equity and a far more subjective concept than equity. In the background policy document relating to access to elective surgery, equity is listed as a key principle. This question is to determine if the concepts of 'fairness' and equity emerged in newspaper reporting about elective surgery. The coding was done by searching for the specific words 'fair', 'fairness', 'equity' and 'equitable'. Of these four terms, the only one found was 'equitable' and that occurred in two articles on consecutive days about a report from the ACT Auditor General.

There were no mismatches in pass comparison because the coding was based on an automated search.

The wording of the question was tightened by adding "or other services" after the independent check by my supervisor.

Summary of final coding criteria

Fairness

The word 'fair' or words with 'fair' as a root are used in the article

Equity

The word 'equity' or words with 'equity' as a root are used in the article

Both

The word 'fair' or words with 'fair' as a root are used in the article and the word 'equity' or words with 'equity' as a root are used in the article.

Neither

Neither the word 'fair' or words with 'fair' as a root are used in the article nor the word 'equity' or words with 'equity' as a root are used in the article.

Question A-4a: Does the article use the concepts of fairness or equity in the provision of elective surgery or other services in the public hospital system?

This question was added after the first pass of coding all the articles

The paucity of explicit mentions of equity and fairness led to the development and inclusion of this concept-based question. Using Stewart's work on these concepts (Stewart 2009) I wanted to distinguish between the intent of a policy and the perception of that policy once it was put into practice. 'Equity' was taken to refer to underlying policy goals for the distribution of resources while 'fairness' related to the perception of people getting (or not getting) a level of service that was fair and reasonable. In other words, 'fairness' was characterised by an emotional reaction to the state's attempts to be equitable. In Stewart's words,

> *…social policy is as much about fairness as it is about equity. As it is implemented in the real world, social policy invokes a set of intricate relationships between different perceptions of fairness. In particular, perceptions of unfairness (that is, feelings against those receiving more than they are entitled to) play a decisive role in shaping responses to the distributions of goods and services that existing policies make possible.* (p66, Stewart 2009)

As a concept-based question this was much more difficult to specify unambiguously than the more concrete questions. The first comparison of my own coding passes had 27 out of a possible 96 mismatches, all due to interpretation rather than simple error. The distinction I initially failed to specify tightly enough was that between the organisational intent of a policy or procedure and the perceived outcome of implementation.

Cross-coding by my supervisor showed that, even with tighter wording there was scope for individuals to differ in the breadth with which they interpreted the definitions. After a final rewriting I developed more detailed specifications which I could use consistently.

Summary of final coding criteria

Fairness

The article concerns the perception of the fairness of implementing policies/procedures. Typically the article relates to individual cases and specific

events. Code as 'fairness' if the implementation of the policy/process is perceived as (un)fair.

Equity

The article concerns the process of calculating or setting policy guidelines for equitable distribution of resources, typically done at a high level and in the abstract. In this case 'equity' includes equity of access to resources. Code as 'equity' if the *intent* of the policy/process is perceived as (in)equitable.

Both

Both the concept of 'fair' as defined above, and the concept of 'equity' as defined above are used in the article

Neither

Neither the concept of 'fair' as defined above, nor the concept of equity as defined above is used in the article.

Question A-5: Is elective surgery defined in the article?

This question did not change during the coding process.

In the formal definition of elective surgery, the word 'elective' has a meaning far removed from the standard dictionary definitions (See Chapter 6.1: The Language of Waiting Lists). Without the formal definition, it is extremely difficult for a lay person to interpret the articles relating to elective surgery. The main exception to this is people who have direct experience of elective surgery in the public hospital systems, either as patients, carers, advocates or health-care professionals.

There were no mismatches in any of the checking for this question.

Summary of final coding criteria

Yes

The term 'elective surgery' is used and defined in the article'

No

The term 'elective surgery' is not defined in the article.

Question A-5a: Is the idea of 'urgency category' used and either defined or implied in the article?

This question was added after the first pass of coding all the articles.

The reason for looking at urgency category is that the category sets formal expectations of how long a patient will wait. In some articles there was reference to

how long a patient had thought they would have to wait but no mention of a formal urgency category. The coding specifications were designed initially to capture explicit references to the term 'urgency category' and were expanded by the addition of the code 'Implied' to capture less precise references to expected waiting times that were based on the urgency categories.

In the comparison of my coding passes all uses of 'implied' were unmatched, with many more instances of 'implied' in the second pass. Analysis of the mismatches showed two main problems: cases where one of the formal labels was used but a vague, rather than an explicit, timeframe was referred to; and mention of how long people had been waiting without there being an indication of how long they should have waited. The definitions were reworded more tightly to explicitly cover the instances where there were mismatches.

Summary of final coding criteria

No

No mention of waiting times for elective surgery.

Used

Either Category 1, Category 2, Category 3 or Urgent, semi-urgent, non-urgent mentioned but not explicitly defined. Includes using the label with a vague timeframe.

Defined

One of the labels from the list for 'Used' appears, modified by a statement of the maximum recommended number of days waiting.

Implied

A (possibly vague) timeframe for when surgery should have happened is mentioned by itself, without a label from the list for 'Used'.

Question A-6: Is the validity or utility of indicators discussed?

This question was added after the first pass of coding all the articles.

The need for this question emerged during the first pass of coding the articles. Those explicitly relating to elective surgery used several different measures, both precise and imprecise, to refer to the issue of people waiting for elective surgery. Two of the measures mentioned are: average (median) waiting time at the time of surgery; and average (median) waiting time for people on the waiting list at a particular reporting date. When there was a reference to people who waited longer than the

recommended time it was sometimes not clear which measure was being used. The interaction between the different measures meant that improving performance according to one measure could mean that performance according to another indicator suffered. This complexity meant that different stakeholders were able to select different indicators to match their chosen narrative.

The initial version of this code was intended to gather enough data to classify which stakeholder was using each particular measure with a view to investigating possible reasons for their choice. This proved to be too specific and didn't uncover any additional articles to those already noted. The question was broadened to ask if the indicators themselves, rather than the performance of the system, were a topic of discussion or analysis.

This was the most difficult code to specify. The comparison of my coding passes showed 14/96 mismatches, all arising from differences in interpretation of the coding criteria. The rewording ended up specifying cases for applying each code rather than giving principles for applying the codes.

Summary of final coding criteria

Yes

Article mentions disagreement about what indicators mean; discussion of/disagreement about data collection, quality and coding; discussion of which indicator to use; discussion of validity of indicators; discussion of utility of indicators; need for and establishment of National Health Performance Authority.

No

Article reports on: discussion of performance; using indicators to describe performance; changes in value of indicators.
Article does not mention performance indicators.

Question A-7: Does the article refer to the concepts of efficiency and/or equity in the provision of elective surgery or other services in the public hospital system?

This question was added after the first pass of coding all the articles.

In matters of public policy there is an inevitable trade-off between efficiency and equity, it is effectively impossible to maximise both. Stewart refers to them as one of the key value pairs in decision making about the provision of services to the public

(p80, Stewart 2009). In the first pass of coding I found only two explicit references to equity but the tension between efficiency and equity seemed to be a distinct undercurrent in the articles. To see whether the undercurrent was real or something I projected onto the source material, I added this code and looked for direct evidence of the interaction of the efficiency/equity value pair. Following my experience of coding for Question A-4, I went straight to looking for the concepts of 'efficiency' and 'equity' rather than searching explicitly for the words. The concept of 'efficiency' I took to include 'productivity'. Although they are not identical in a technical sense, their meaning is close enough for the purposes of this code. I classified concerns about access under the heading of 'equity'.

The comparison between my two coding passes showed some interpretation differences, mainly arising from more explicitly classifying 'access' issues as equity issues. The simple errors arose primarily from earlier ambiguities in coding for Question 4a. After re-wording both this question and Question 4a I re-coded all the articles for these questions. The differences between my final coding and my supervisor's arose from his more broad interpretation of the word 'equity'.

Summary of final coding criteria

Note: Commenting on a change in waiting times by itself does not trigger this question. There must be at least an implied judgement about the efficiency or equity represented by the change.

Efficiency

The article mentions efficiency; cost/benefit; value for money or similar values.

Equity

If Question A-4 or question A-4a is coded as 'equity' or 'both', this question must also be coded as 'yes' for 'equity'. The criteria for coding as 'equity' are identical to those used for questions 4 and 4a.

Both

The concept of efficiency is discussed and the answer to question 4a is either "equity' or 'Both'.

Neither

Neither efficiency not equity are referred to in the article.

If the answer to question 4a is 'Neither' then the only possible answers to this question are 'efficiency' and 'neither'.

Table 4-3: Final coding criteria for newspaper articles

	Question	**Codes**
A-1	Is the article explicitly about elective surgery or about the hospital system more generally?	Yes No
A-2	Is the article about an individual's experiences or about elective surgery/the hospital system in general	Individual System Both
A-2a	Are stories about individuals positive or negative?	Positive Negative N/A
A-3	Is the article about the Australian hospital system or about the ACT hospital system?	Federal ACT ACT region Federal/ACT
A-4	Does the article explicitly refer to either 'fairness' or 'equity' in the provision of elective surgery in the public hospital system?	Fairness Equity Both Neither
A-4a	Does the article use the concepts of fairness or equity in the provision of elective surgery in the public hospital system?	Fairness Equity Both Neither
A-5	Is elective surgery defined in the article?	Yes No
A-5a	Is the idea of 'urgency category' used and either defined or implied in the article?	No Used Defined Implied
A-6	Is the validity or utility of indicators discussed?	Yes No
A-7	Does the article refer to the concepts of efficiency and/or equity in the provision of elective surgery in the public hospital system?	Efficiency Equity Both Neither

For ease of reference a summary of the final coding criteria and rationale for each question can be found in Appendix B.

Thematic coding is a powerful tool for extracting structured meaning from a collection of texts but the more concept-based the code, the more time-consuming the code development and application. Despite reading descriptions of how long it

could take I was surprised at just how time-consuming it turned out to be. If I were to do it again I would break down my proposed coding questions according to their complexity and use the coding time estimates derived from Table 4-4 and Table 4-6 and summarised in Table 7-1 on page 187 to work out how long it would take. By generalising question according to complexity I have developed a simple tool that can be used to work out what type of questions and how many can be used in a given time for a piece of thematic analysis.

Developing codes for explicit content such as question *A-4 Does the article explicitly refer to either 'fairness' or 'equity' in the provision of elective surgery in the public hospital system?* were relatively fast to develop and code. They typically required only two or three passes through the collection of 96 articles to ensure that there was a code for each article (i.e. the code set was complete) and that only one code applied to each article (i.e. the codes were discrete). Coding time for these questions eventually dropped from eight hours per pass through the article collection to under .2 hours.

Codes where decisions needed to be made about content such as *A-3 Is the article about the Australian hospital system or about the ACT hospital system?* were far more time-consuming. Working from the initial binary form of the code to the final four-part code took many passes through the collection of 96 articles, making sure that the codes were both complete and discrete. It took multiple iterations of writing the code descriptions so that the differences between the codes was clear enough for them to be applied reliably. Coding time for these questions started at each pass through the collection taking several days and ended up at just under a day.

The most time-consuming codes to develop and apply were those relating to abstract concepts, such as *A-2 Are stories about individuals positive or negative?* and *A-4a Does the article use the concepts of fairness or equity in the provision of elective surgery in the public hospital system?* This type of code took many rewrites of the definitions and passes through the article collection before they could be applied reliably. This shows clearly in the analysis of error types for the pass comparisons in Table 4-2 above. The definitions for these codes are typically longer and more complex than those relating to codes for explicit content. Coding time started at four days per pass through the article collection and eventually dropped to one day per pass. Table 4-4 below shows the approximate time taken to code each question based on the number of articles, number of passes required and the time taken for each pass. The long times

represented the passes through the collection while the code was being developed and the short times represent the final passes used as the basis for cross-checking.

Table 4-4: Coding times for article-level thematic analysis

Type of question	Long coding time per question	Short coding time per question	Total Coding time per question	Number of questions	Total coding time
Explicit content	3 passes @ 5 min/article = 24 hours	2 passes @ 1 min/article = 3.2 hours	27.2 hours	3	82 hours
Decision-based content	4 passes @ 15 min/article = 96 hours	2 passes @ 5 min/article = 16 hours	112 hours	3	336 hours
Implicit content	6 passes @ 20 min/article = 192 hours	2 passes @ 5min/article = 16 hours	208 hours	4	832 hours
Total					**1250 hours (156.25 days)**

4.5 Indicator use analysis

The second level of analysis I applied to the material from *The Canberra Times* involved looking more deeply within the articles at how elective surgery waiting times were being used as a performance indicator for the public health system.

4.5.1 Article identification

The first step was take the 58 articles identified in the coding for question A-1 as explicitly referring to elective surgery and identify those where waiting times were used as performance indicators. The key determinant of this was that a waiting time was mentioned or implied and it was compared to something else. Following the pattern of earlier coding, I framed this as a question with a closed set of answers. The question numbers for this set begin with 'PI' to indicate that they are related to uses of waiting times as a performance indicator.

Question PI-1: Are waiting times used as a performance indicator in the article?

- **Yes**

  A waiting time, either a system average or an individual's is compared to something else.

  E.g. using the waiting time in connection with an urgency category or other clinical guidelines; comparing a waiting time in one place to that in another; commenting on whether the waiting time seems short or long compared to the patient's expectations.

- **No**

  A waiting time is used in isolation without reference to a comparator.

  Waiting times are not mentioned.

Each article identified was annotated with a brief reason for its selection e.g. 'Reference to actual waiting times, urgency categories and guidelines'. Of the 58 articles referring explicitly to elective surgery, 43 contained at least one use of elective surgery waiting time as a performance indicator.

4.5.2 Instance identification

Since this analysis was based on each instance of performance indicator use, not each article, I created a new data set with identifiers for each use of an indicator. E.g. I01242010 was the first reference to an indicator in article A242010. The basic metadata for each instance of indicator use was:

- Instance Identifier

- Article identifier

- Article title

- Gist of indicator use: this is a short summary for each instance of which indicators were used and how. E.g. 'Current average waiting times for urgent surgery compared to longest guideline'

If the same indicator was used twice but expressed differently it was recorded as two different instances of indicator use. E.g. "The ACT has the longest waiting times for Category 3 surgery" and "In the ACT, 44.1 per cent of Category 2 patients received their treatment on time, compared to next-placed Northern Territory." was coded as two different instances of indicator use.

Where two indicators had been conflated within a statement they were coded separately. For example "Canberrans have the longest queues for elective surgery in the nation, with more than 10per cent waiting more than 365 days for treatment." was coded as two instances of indicator use: the number of Canberrans waiting compared to other states and the proportion of Canberrans waiting more than 365 days.

Often the same indicator was used to illustrate a series of articles, for example after the release of a report on hospital performance there would be several articles discussing the report from different viewpoints. The use of the same indicator in the same way but in different articles was coded as separate instances. There was a total of 116 instances of elective surgery waiting times being used as a performance indicator in the 43 articles identified as containing at least one instance.

In addition to the basic metadata, I included the four fields from the original thematic coding of the articles which dealt most closely with indicator use:

- Is elective surgery explicitly mentioned?
- Is the article about an individual or the hospital system?
- Is 'urgency category' used, defined or implied?
- Are indicators discussed?

4.5.3 Instance coding

The next level of categorisation looked at each instance of using waiting times as a performance indicator for details of what statistics were used to express waiting time as a performance indicator, what assumptions were embedded in that expression and what, if any, comments there were about performance. The final attribute I looked for was the match between the statistical meaning and the comments about performance.

I took each instance of indicator use and, working primarily from the 'Gist of indicator use' field, coded for:

- type of statistic
- bases for comparison
- the entity being measured
- statistical assumptions

- communication assumptions
- value judgement about performance
- match between statistical information and value judgement

Again these were framed as questions with a closed set of possible answers

Question PI-2: What type of statistic is used to describe elective surgery waiting time?

This question is focussed on how elective surgery waiting time is expressed. Statistically speaking the simplest expressions are counts, followed by individual waiting times. Next in complexity is the median for a single urgency category followed by the median across all urgency categories.

- **Counts:**
  Performance indicators relating to elective surgery that are expressed as counts include: people added to waiting list; people on the waiting list; and surgeries performed. While the latter is not explicitly a waiting time, is it used as part of the suite of performance measures for elective surgery.

- **Individual waiting time**
  This is usually expressed relative to the clinical guidelines and is implied within phrases such as 'waiting too long for urgent surgery'.

- **Median waiting times**
  Medians rather than means are used to describe state and national average waiting times. The two medians used are: the median for a single urgency category and the median for all people waiting for surgery in the jurisdiction of interest. The latter is statistically difficult as if refers to performance against an unknown mix of urgency categories, each of which has a different recommended waiting time. This difficulty is flagged in the coding for 'Statistical Assumptions' below.

Question PI-3: With what is the waiting time statistic compared?

In order to be an indicator of performance, a waiting time statistic needs to be compared to something either implicitly or explicitly. In many cases several comparators were combined, for instance the ACT's performance relative to the clinical guidelines in a particular time period could be compared to its performance

in an earlier time period or the performance of another state. The most complex cases were a combination of three comparators.

- **Count of people seen within time comparator**

  Used as a second or third comparator when an earlier comparator is time-based e.g. a guideline for an urgency category. Statistically difficult if comparison is between two populations of widely different sizes. This difficulty is dealt with in the coding for 'Statistical Assumptions' below.

- **Doctor's reported recommendation**

  Optimal waiting time according to treating doctor, as reported by patient. No clear reference to guidelines or urgency categories. E.g. 'Told he needed surgery within a fortnight' [I01232010].

- **Guideline for appropriate urgency category**

  There is enough information to determine that the waiting time statistic is being compared to the guideline for a specific, and by inference, appropriate urgency category. This is not limited to instances where an urgency category is specified. The phrase 'waiting too long' is taken to mean waiting longer than the guidelines for the patient's urgency category.

- **Guideline for least urgent category (1 year)**

  The waiting time statistic, either 'Individual Waiting Time' or a median is compared to one year, which is also the guideline for Category 3 (non-urgent) surgery.

- **Median for the appropriate urgency category**

  This is used when the waiting time statistic is 'Individual waiting time' and an urgency category is given or able to be inferred.

- **Median for a mix of urgency categories**

  A median for elective surgery waiting times across all urgency categories. Since the proportion of each urgency category in the statistical population is not given, this is statistically difficult. This difficulty is dealt with in the coding for 'Statistical Assumptions' below.

- **Percentage of people seen within time comparator**

  Used as a second or third comparator when an earlier comparator is time-based e.g. a guideline for an urgency category. Comparing percentages of

different sized populations can be difficult to understand. This difficulty is dealt with in the coding for 'Communication Assumptions' below.

- **Result from previous time period**

  The same statistic, and other comparators if any, is compared between two different time periods.

- **Result in other state(s)**

  The same statistic, and other comparators if any, is compared between two different state(s) or against national results. Used when the ACT's rank (e.g. worst in Australia) is mentioned, even if no state jurisdiction is referred to.

Question PI-4: Which part of the public health system's performance is being measured?

Elective surgery waiting time data is used as a performance indicator for both the state-level and national public hospital system. The reporting in *The Canberra Times* is primarily focussed on the ACT with other jurisdictions most often used as comparators.

- **ACT**

  The waiting list statistic is for patients using the ACT public hospital. system for elective surgery.

- **National**

  The waiting list statistic is for all patients in Australia using the public hospital system for elective surgery.

- **Other State(s) – excludes ACT**

  The waiting list statistic is for patients using a state public hospital system other that the ACT's.

Question PI-5: What statistical assumptions are needed for the comparison to be statistically valid?

When waiting time statistics are used in a newspaper report, if assumptions are required to ensure that they are statistically valid they are typically not included. This question codifies these unwritten assumptions when they occur. The assumptions fall into two broad categories: those required to ensure that like is compared with like and those relating to what inferences can be drawn from the statistic.

Comparison assumptions:

- **All comparators assumed**

  This barely qualifies as a performance indicator, a median is reported with no explicit comparators.

- **Doctor's recommendation is a firm guideline**

  When a doctor gives a recommended time within which surgery should occur but this is not related to a formal urgency category, the assumption is that the recommendation takes the place of the category guideline.

- **Each urgency category compared to itself**

  If a combination of urgency categories or an imprecise description of urgency category such as 'less urgent' is compared between time periods or state(s) the assumption is that each urgency category is compered to itself.

- **Guidelines less than time already waited**

  Used for comparisons of 'patients/people waiting too long'. The assumption is that the time people have waited is greater than the clinical guideline for their urgency category.

- **Rate of people joining list is constant**

  Comparisons of the number of people on the waiting list, either by time period or state(s) is only a performance indicator if the rate of people joining the list is the same for both parts of the comparison.

- **Time to be removed from the waiting list is the same as waiting time**

  There are two periods of time that are referred to as 'waiting time': the total time from joining the list until surgery occurs and the time from joining the list until the reporting date. These are sometimes confused or conflated.

- **Urgency category mix the same in each state**
  Urgency category mix unchanged over time

  Performance judgements of comparisons of the median waiting time for all people on the elective surgery waiting list are strongly affected by the mix of urgency categories. To take an extreme example, a median waiting time of 45 days is a good result if everyone on the list is urgency category 3 (treatment within 365 days) but a poor result if everyone is urgency category 1 (treatment within 30 days).

Inferential assumption:

- **Median tells you about an individual**

  Articles where information about an individual's likely waiting time were inferred from the median (or a change in the median) for a population assume that this inference is possible. In practice the median for a population tells you nothing about where an individual might be in the distribution of that population

Question PI-6: What communication assumptions are needed for the comparison to be meaningful on a human level?

In some senses the statistical assumptions covered by Question PI-5 are also communication assumptions but there are also some assumptions with a clearer communication basis. This question is tightly tied to Question PI-8 which deals with value judgements about performance.

Global assumptions

The term 'elective surgery' is not defined in any article in the collection (see Newspaper Analysis Results p 5-3). This means that a communication assumption that holds for every instance of indicator use is that the reader knows what elective surgery is. Closely related to this is the assumption that the reader knows the clinical guidelines for each urgency category. A more subtle global assumption is that the time elapsed since joining the waiting list is the same as the patient's waiting time. As described in Chapter 3.3.2 *Elective surgery process overview*, a patient is only considered to be on the waiting list if they are 'ready for care'. This means that any time a patient spends being too unwell to undergo surgery is not counted as part of the time they have been on the list.

- **Personal experience is more 'true' than statistics**

  This is used when the statistics are described as 'untrue' or 'wrong' because they do not reflect someone's personal experience.

- **Percentages of different sized populations are easy to compare Percentage increases are easy to understand**

  These two assumptions underlie comparisons based on percentages and percentage changes. The communication pitfalls of this are covered on pages 37-38 of *Chapter 2: Literature Review.*

- The following codes capture assumptions about what movement in the indicator means and form part of the basis for the next coding question relating to value judgements.

    Fewer operations is a bad outcome

    This would not be true if the reason for there being fewer operations was that fewer people needed them.

    Guideline is shorter than time comparator

    This is used when the average waiting time for a category is compared to a longer timeframe without the guideline time period being specified.

    Lower proportion of people waiting more than a year is good

    How good an outcome this is depends upon the mix of urgency categories in the two time periods or states being compared.

    More surgeries is a good outcome

    This is would not be true if the reason for there being more surgeries was that more people needed them

    'Shortest' means fewest people on the list

    The description 'shortest waiting list' could either mean the list with the fewest people on it or the list with the shortest waiting time. In the instances where the description is used, the former is a better contextual fit for the intended meaning.

Question PI-7: What value judgement about performance is expressed?

To see how the indicator was being used in as part of communication I added a code for recording any value judgements about the performance being measured e.g. good result, disgraceful figures, not good enough. This code allowed me to categorise any value judgement communicated with the performance indicator.

- **Neutral**

    No comment on the performance being measured.

- **Good**

    Comment on the performance being measured uses words with an element of judgement such as 'dropped', 'best', 'improved', 'good' and 'excellent'.

- **Bad**

    Comment on the performance being measured uses words with an element of judgement such as 'rose', 'decline', 'should have been seen by', 'too long'

'blown out', 'worst' and 'poor'.

If the ACT performance was judged to be bad, but not as bad as another state (e.g. I0528201 "More than 10 per cent of ACT elective surgery patients were forced to wait more than a year to undergo a procedure. Only Tasmania performed worse…") the instance was coded as 'Bad'.

Question PI-8: Does the value judgement reflect the meaning of the statistics?

In the statistical sense, a performance comparison which shows a move towards the performance goals over time or that the jurisdiction of interest is better than some other jurisdiction would be regarded as good performance and *vice versa*. This question captures whether or not the value judgement expressed about performance is consistent with the meaning derived from the statistics.

- **Yes**

  The value judgement is consistent with the meaning of the statistic.

- **Partly/Unable to tell**

  The value judgement is at least partly consistent with the meaning of the statistic.

  It is not possible to tell if the value judgement is consistent with the meaning of the statistic. The most obvious reason for this would be that there are too many statistical and communication assumptions required to be certain that like is being compared with like.

- **No**

  The value judgement is inconsistent with the meaning of the statistic

Worked example of instance coding

"Nationally, the percentage of people waiting longer than a year for elective surgery dropped from 4.8 per cent in 2004-05 to 2.9 per cent in 2008-09." would be coded as

- Gist of indicator use: "Comparison between years of National % waiting over 365 days - drop"
- Type of Statistic: Individual Waiting Time
- Comparator 1: Guideline for least urgent category
- Comparator 2: Percentage of individuals seen within time comparator
- Comparator 3: Result from different time period

- Entity: National System
- Validity Assumption: Urgency Category mix unchanged over time
- Communication Assumption: Fewer waiting more than a year is good
- Value Judgement: Good
- Match between statistics and value judgement: Yes

As with the article-level coding, I did two separate coding passes and then a range of checks for consistency of coding. The internal consistency checks for each pass were:

- Check all codes used exactly match those in the final specification
- Check and fill blanks in the coding
- Find occurrences of the same statistic and comparators to ensure they are coded consistently.
- Compare codes between each pass

Table 4-5 below shows the results of checking the instance-level coding.

Table 4-5: Results of checking instance-level coding

	Question	Matches	Simple Errors	Different Interpretation
PI-1	Are waiting times used as a performance indicator in the article?	96	1	0
PI-2	What type of statistic is used to describe elective surgery waiting time?	115	1	0
PI-3	With what is the waiting time statistic compared? Comparator 1 Comparator 2 Comparator 3	114 115 115	2 1 1	0 0 0
PI-4	Which part of the public health system's performance is being measured?	115	1	0
PI-5	What statistical assumptions are needed for the comparison to be statistically valid?	87	2	27
PI-6	What communication assumptions are needed for the comparison to be meaningful on a human level?	105	11	0
PI-7	What value judgement about performance is expressed?	107	2	7
PI-8	Does the value judgement reflect the meaning of the statistics?	109	7	0

Checking for question PI-1 was done at the article level and checking for the other questions was done at the instance level.

In the second coding pass for question PI-1 (waiting time used as indicator) I found one more article where elective surgery waiting times were used as a performance indicator and two cases where two instances had been coded as one. This took the number of articles up to 43 and the number of instances up to 116. For question PI-2 (what type of statistic), PI-3 9 (what comparators) and PI-4 (what part of the health system) the code mismatches were all simple errors, all but one of these were caused by the earlier error of coding two instances of indicator use as one.

The mismatches for question PI-5 (validity assumption) were interpretation errors, all but three of which related to broadening the circumstances under which I applied assumptions relating to the mix of urgency categories. The three exceptions were because I had used more codes on the second pass to help capture finer detail.

For question PI-6 (communication assumption) the mismatches were all instances that were not coded on the first pass. These were picked up when I checked that each time the same statistic and comparators was used it was coded in the same way.

Question PI-7 had two simple errors and seven interpretation errors. All of the interpretation errors related to the code 'Neutral'. The errors in the coding of question PI-8 were all simple errors caused by the interpretation errors in question PI-7.

Stories relating to individual experiences contained the most difficult instances to code. The description of waiting times was typically less precise and used more emotional language. For some of these stories it was necessary to calculate how long the patient had been waiting based on when they were referred for surgery and the date of the article. Cross-checking between these stories ensured that the coding was a consistent as possible.

As with the article-level coding, the more descriptive codes were far easier to apply than those relating to value judgements and implicit information. Using the same three basic types of coding as for the article level analysis Table 4-6 below shows the approximate time taken to code each question based on the number of passes through the 116 instances of waiting time being used as a performance indicator and

the time taken for each pass. The long times represented the passes through the collection while the code was being developed and the short times represent the final passes used as the basis for cross-checking. Developing and applying this second set of codes was faster than the first, reflecting my increasing skill with this technique.

Table 4-6: Coding times for instance-level thematic coding

Type of question	Long coding time per question	Short coding time per question	Total coding time per question	Number of questions	Total coding time
Explicit content	3 passes @ 5 min/instance = 29 hours	2 passes @ 1 min/instance = 3.9 hours	32.9 hours	3	98.6 hours
Decision-based content	4 passes @ 15 min/instance = 116 hours	2 passes @ 5 min/instance = 9.6 hours	125.6 hours	1	125.6 hours
Implicit content	5 passes @ 20 min/instance = 193.3 hours	2 passes @ 5 min/instance = 9.6 hours	212.6 hours	4	850.6 hours
Total					**1085 hours (135 days**)

4.6 Longer stories and themes

As well as day to day reporting and coverage of policy announcements there were some longer stories and recurring themes which were reported in several articles over the two years of my collection. Each of the articles relating to these longer stories and themes was coded so that the stories could be consolidated.

- **Mayor of Dickson**

  This story focussed on an intellectually disabled individual who had trouble understanding the delays to their elective surgery and the frustrations of their carer as they tried to help.

- **ACT Auditor-general's investigation**

  The ACT Auditor-general investigated claims of mismanagement and

> manipulation of the ACT elective surgery waiting lists. The Mayor of
> Dickson story was part of the impetus behind the investigation (see above and
> Chapter 5.4.1).

- **COAG negotiations and the National Hospital Reform Agenda**
  During 2010 and 2011 negotiations were held between the Commonwealth
  government and the States and Territories over funding responsibilities and
  performance management of Australian public hospitals.

4.7 Complexities and limitations of the material from The Canberra Times

The articles in *The Canberra Times* during my collection period did not fall neatly into
the two categories of those related to elective surgery in the public hospital system
and those not related to elective surgery in the public hospital system. Two other
major issues were often reported in the same article or in closely-related articles. The
first issue was a local one related to Calvary Hospital and the second was the other
main measure of hospital performance, emergency department waiting times.

The relationship between Calvary Hospital, which is operated as both a private
hospital and as a public hospital, and the wider ACT public hospital system was a
continuing issue of public interest during 2010 and 2011. Much of the reporting was
focussed on the possible sale of the hospital to the ACT government and the issue of
who 'really' owned it. The issue impacted on elective surgery waiting times for two
reasons: the issue of operating theatre space, and the potential for the public hospital
system to pay for its patients to be treated in the private system. I picked up the
articles relating to this issue in my collection and included them in my thematic
analysis. In the end I decided against analysing the interaction of the public and
private hospital systems as providers of elective surgery. The issue was too complex to
analyse properly within the scope of this research and there was insufficient publically
available material about waiting times in the private system.

Along with elective surgery waiting times, the other measure of public hospital system
performance that was prominent in the news was emergency department waiting
times. Articles about the public hospital system frequently referred to both as well as
an occasional mention of 'emergency surgery'. While it is likely that there is a
relationship between the three issues, I decided that including all the articles relating

to hospital emergency departments had the potential to add volume rather than clarity to the data set.

One article (A242011) appeared to conflate the two waiting time based measures. It was primarily focussed on elective surgery waiting times except for a key sentence which referred to 'emergency targets'. This made it hard to tell which set of targets, elective surgery or emergency department, the phrase 'waiting longer than clinically recommended' was referring to as both types of treatment use urgency categories each of which has a clinically recommended maximum waiting time. A similar article from the following day (A252011) had the same problem.

This chapter has described the methods used for analysing articles from The Canberra Times for the period 1 January 2010 to 31 December 2011. As well as a broad analysis of the collection of the articles as a whole I did two more detailed analyses using emergent coding techniques within the framework of thematic analysis. The first detailed analysis was at the level of each article in the collection, characterising them according to factors such as what part of the hospital system they related to, whether the stories were about individuals or the system, whether technical terms such as urgency category were used and if they were defined, and whether the concepts of equity, fairness and efficiency were used. The second detailed analysis looked at each instance of elective surgery waiting time being used as a performance indicator and coded for type of statistic, comparators, assumptions, value judgements, and the match between the statistical meaning and the value judgement. The next chapter will contain the results of these analyses.

5 Newspaper analysis results

This chapter contains the results of the newspaper analysis described in the previous chapter. The results are presented in three sections:

- An overview of the collection and how it fits the timeline of events described in Chapter 3.2.2;

- Tables showing the results of analysis of the material at the article level; and

- Tables showing the results of the analysis at the level of each instance of elective surgery waiting time being used as a performance indicator.

5.1 Overview of Collection and matching to timeline

Of the 16,282 articles in the online archive for The Canberra Times over my sample period of 1 January 2010 to 31 December 2011, I identified 96 relating to elective surgery. Of these 58 mentioned elective surgery explicitly while the remaining 38 related to the wider funding and political background to the Australian public hospital system.

The month-by-month distribution of elective surgery related articles as a percentage of total articles published is shown in Figure 5.1 below. A combination of national and local issues which occurred at the same time as a spike in the proportion of elective surgery-related articles are:

- Federal government begins trying to reform the hospital system (National; March 2010)

- Problems with ACT elective surgery waiting times (Local; June, 2010)

- Finalising of agreement between the Federal and State governments and the impact on the ACT (National with local impact; February 2011)

- By 2011, there are clear spikes for the release of the quarterly hospital system performance reports (National with a local emphasis; June, September).

- Negotiations between the ACT Government and the managers of Calvary Hospital the Little Company of Mary. These were relevant because the hospital provides both private and public services (Local, December 2011).

Comparing these events with the timeline relating to the wider background for elective surgery from Chapter 3.3.2 COAG Reform Agenda Timeline, there is no

clear event of either national or local importance associated with the increase in the proportion of elective surgery-related articles in October 2010.
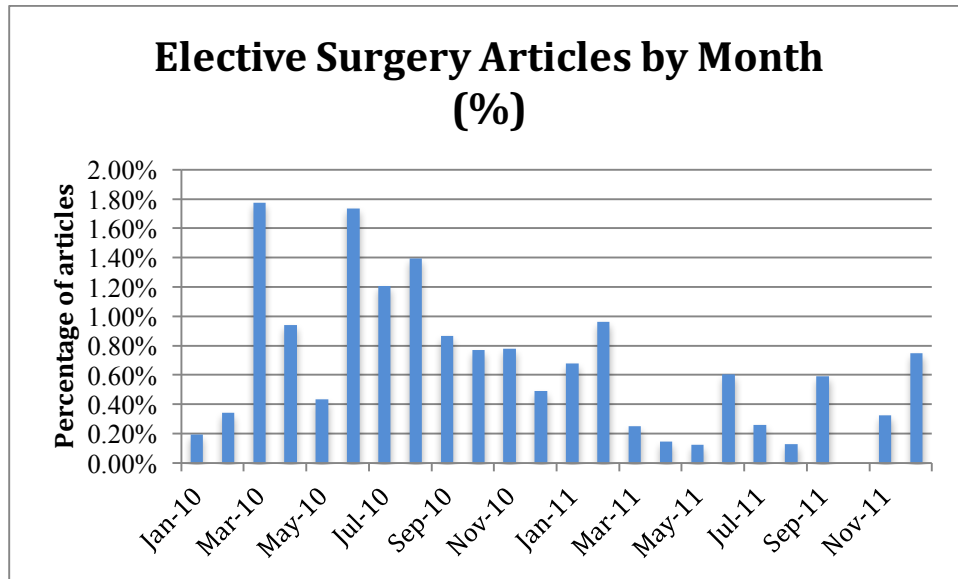
Figure 5.1: Percentage of elective surgery-related articles by month

March 2010

There were 10 articles relating in some way to elective surgery in March 2010. They began with the announcement that the Federal Government was offering $AUS90 Billion to overhaul the public hospital system, in exchange for a transfer to them of responsibility from the States and Territories as well as a cut in GST distribution. All but one of the articles related to the negotiations and opposition responses to the negotiations. The other article was a story about a child who had been waiting for elective surgery far longer than the recommended time. It was published on 15 March, the same day as two articles about the negotiations between the Federal and State governments.

June 2010

An article in June about a disabled patient, nicknamed 'The Mayor of Dickson' who was distressed by his long wait for elective surgery was the first in a series about long elective surgery waiting times in the ACT. During the rest of June and into early July there were nine more articles focussing on the ACT elective surgery waiting lists and allegations that the performance statistics for elective surgery in the ACT were being manipulated. Two of these were follow-up articles about the patient from the first story. These issues are described in more detail in chapter 5.2.8 on page 143. There was a related article about the shortage of doctors in the public health system and the

resulting dependence on fly-in contractors. In late June an article reported on the ACT Legislative Assembly referring the issue of elective surgery waiting list management and data collection to the ACT Auditor-General.

February 2011

In February 2011 there were five articles relating to the final negotiations about public hospital funding between the Federal and State governments. There was also an article relating to negotiations relating to elective surgery between the ACT and Queanbeyan hospital across the border in NSW. Although a smaller number of articles than previous spikes this was still a high proportion of articles for the month.

June 2011

The 5 articles in June 2011 were mostly about the release of the first quarterly report on public hospital performance under the National Hospital Reform Agenda with a focus on how the ACT performed relative to other states and against performance targets. One of the articles concentrated on a patient's objection to being 'treated like a number'.

September 2011

In September 2011 there were four articles relating to the quarterly performance report and the financial penalty incurred by the ACT's failure to meet elective surgery and emergency department waiting time targets. The other article that month was about the death of the Mayor of Dickson whose story began the series of articles about elective surgery delays in the ACT public hospital system. This story is told in more detail in Chapter 5.4.1 on page 143 below.

December 2011

Despite the quarterly performance report scheduled for December 2011 being delayed there were seven articles relating to elective surgery published that month. These were on a range of topics with the resolution of a dispute between the ACT government and The Little Company of Mary over ownership of Calvary Hospital featuring in the greatest number of articles.

5.2 Results of article-level thematic analysis

5.2.1 Question by question results

The results of the basic thematic coding for the Canberra Times articles relating to elective surgery are given below. See Chapter 4.4.1 on page 100. for details of the code development and Appendix B for a listing of the final coding criteria.

Question A-1: Is the article explicitly about elective surgery in the public system or about the (public) hospital system more generally?

Although all the articles collected related to elective surgery in some way, only a subset of them explicitly mentioned it. The others were about the wider public hospital system covering issues such as funding, accountability and performance. Table 5-1below shows the breakdown between the two types of article.

Table 5-1: Number of articles explicitly about elective surgery

	Count
Yes	58
No	38
Total	**96**

Question A-2: Is the article about an individual's experiences or about elective surgery in general?

Some of the articles were solely about issues at the system level while others used personal stories from patients and carers who had experience with elective surgery in the public hospital system. Table 5-2 below shows the breakdown of articles about individuals, the public hospital system or with elements of both.

Table 5-2: Number of articles about individuals, the system or both

	Count
Individual	9
System	80
Both	7
Total	**96**

Question A-2a: Are stories about individuals positive or negative?

The articles containing stories about individuals also expressed value judgements about the public hospital system. Value judgements are explored further in the

coding and analysis for each instance of elective surgery waiting time being used as a performance indicator (See Chapter 4.5 and Chapter 5.3 below). Table 5-3 below shows the breakdown of positive and negative judgements in stories about individuals.

Table 5-3: Types of value judgements in stories about individuals

	Count
Positive	0
Negative	16
Both	0
N/A	80
Total	**96**

Question A-3: Is the article about the Australian hospital system or about the ACT hospital system?

The Australian public hospital system has several levels of funding and governance (See Chapter 3.2 for details). Table 5-4 below shows the breakdown of articles according to which level they discuss.

Table 5-4: Level of hospital system

	Count
Federal	18
ACT	50
ACT Region	10
Federal/ACT	18
Total	**96**

Question A-4: Does the article explicitly refer to either 'fairness' or 'equity' in the provision of elective surgery or other services in the public hospital system?

> *Question A-4a: Does the article use the concepts of fairness or equity in the provision of elective surgery or other services in the public hospital system?*

The principles of fairness and equity underlie key aspects of public health policy development and implementation. For this analysis equity refers to the intent of a policy and fairness to perceptions of that policy when it is implemented. Table 5-5 below shows counts of articles containing either the words fairness and equity or words derived from them such as fairly and equitable. Table 5-6 below shows counts of articles containing the concepts of fairness and equity.

Table 5-5: Fairness and equity – words

	Count
Fairness	0
Equity	2
Both	0
Neither	94
Total	**96**

Table 5-6: Fairness and equity – concepts

	Count
Fairness	15
Equity	13
Both	2
Neither	66
Total	**96**

Question A-5: Is elective surgery defined in the article?

The term 'elective surgery' has a specific technical meaning in the context of the public health system. The meaning of the word elective in this context is not the same as its dictionary and colloquial meanings. Table 5-7 below shows that the term elective surgery was not defined in any of the articles analysed.

Table 5-7: Definition of elective surgery

	Count
Yes	0
No	96
Total	**96**

Question A-5a: Is the idea of 'urgency category' used and either defined or implied in the article?

The urgency category for an elective surgery procedure describes the recommended maximum waiting time for that surgery and waiting longer than recommended is often used as performance indicator. Table 5-8 below shows a breakdown of articles according to whether terms relating to urgency categories were not used, used without being defined, used and defined or used with an implied definition.

Table 5-8: Urgency category definition and use

	Count
No	71
Used with no definition	11
Used and defined	10
Used and implied	4
Total	**96**

Question A-6: Are indicators and their utility discussed?

As well as using elective surgery waiting time as a performance indicator for the public hospital system, some articles reported on discussions about the usefulness and appropriateness of using different waiting time statistics as indicators. These included disputes about how urgency categories were allocated and debates about how well the performance statistics reflected patients' experiences. Table 5-9 below shows how many articles included this type of discussion.

Table 5-9: Discussion of indicators and their utility

	Count
Yes:	28
No:	68
Total	**96**

Question A-7: Does the article refer to the concepts of efficiency and/or equity in the provision of elective surgery or other services in the public hospital system?

In public policy relating to service provision one of the important trade-offs is between efficiency and equity. There is a constant need to balance the need to use public resources efficiently and the need to provide equitable access to healthcare. Table 5-10 below shows a breakdown of articles according to whether one or both of these concepts was used.

Table 5-10: Efficiency and equity – concepts

	Count
Efficiency	9
Equity	12
Both	4
Neither	71
Total	**96**

5.2.2 Combining article-level codes

Looking at the results of the article level coding there are some clear patterns:

- There were articles about individual patients' experiences with elective surgery in the public hospital system combined with or adjacent to more general articles about either elective surgery or the wider public health system. Two of these articles about individuals were explicitly about the statistics and patients' reactions to them.

- Although the terms "elective surgery" and "urgency category" were often used, they were rarely explained or defined.

- A small number of articles dealt explicitly with the meaning, validity and use of performance indicators rather than the system being measured.

- The concepts of fairness, equity and efficiency were referred to in 40 out of the 96 articles.

5.2.3 Individual patients' stories

Questions A-1, A-2 and A-3 relate to the scope of the newspaper article. They cover whether it is directly about elective surgery, whether it is about an individual or the public health system and which part of the public health system it is about. Question A-2a relates to the emotional tone of the commentary about individual stories. Of the 96 articles, nine concerned the experiences of individual patients and seven combined an individual's story with a story about the public hospital system more generally. Of the individual stories, three were published on the same day as more general elective surgery articles.

In every case, the overall tone of the articles about individual's experiences was negative, focussing on how long patients had waited; the pain and disability the patient experienced while waiting; and the frustrations of trying to find out when

their surgery would occur. Even when a patient had been operated on and was feeling better, the commentary focussed on the unreasonable time they had waited and how bad 'the system' was to let them wait so long. The story chosen to accompany an article about improvements in waiting times was that of a patient who had not benefited from the concerted effort to clear surgery backlogs.

The articles about individuals were almost all part of articles relating to the ACT or the ACT Region rather than those relating to Federal matters. On the one occasion that an individual account accompanied articles about issues at the Federal level, the individual account was related to the ACT hospital system and appeared as a separate article (15 March 2010: A082010, A092010, A102010).

One of the hardest aspects of dealing with statistical information is coming to terms with the fact that it can tell you absolutely nothing definite about a particular case and it may appear to be quite unrelated to one's individual experience. This shows up very clearly in two articles from quite different viewpoints.

In article A152011 *I'm Like a Number: Gowrie Mum* there is a description of a patient's reaction to the performance indicators and her strong feeling that being 'treated like a number' is dehumanising and indicative of a system that doesn't care. The patient's concern is not with the accuracy or validity of the statistics but with the social and emotional effects of gathering and reporting statistical information about people. Similar emotions are expressed in A402010 *Family Lost in the Statistics*. The mother of a sick child is critical of a review of hospital services saying that it had left her family feeling as though they were statistics. She goes on to say "It's not about figures and numbers and things … We're talking about people's lives and how they're affected and I think that's the main thing that was really missing in that report." Both articles show a strong sense that statistics do not capture what these patients and carers think of as the important information about their experience.

Article A202011 *My 17-month surgery delay* deals with a patient whose surgery is well overdue and who hasn't benefited from a targeted program to treat overdue patients and reduce waiting times. It included the comment by the patient's wife "I was annoyed when the Chief Minister said the elective surgery waiting list had gone down, because it seems the truth is completely different." In this case there is a clear mismatch between the information provided by the performance indicator and the patient's immediate experience.

An article which clearly shows the effect of a patient's ready for care' status on the timing of their surgery is A292011 *Experience leaves patient sick of the hospital system*. The patient needed surgery for gallstones, he had a history of cardiac problems and was morbidly obese. He was admitted for the surgery on several occasions, only to have it cancelled. On one occasion this was because more urgent cases took priority but on several others it was because the anaesthetist considered the risk of the required anaesthetic to be too high. This lack of agreement between different specialists about whether or not the patient was fit to undergo the surgery meant that according to one specialist he was technically on the waiting list and being scheduled for surgery when, according to another, the anaesthetist, he was not ready for care and therefore should not have been on the list. This confusion will have had an effect on the official time the patient was recorded as having waited for his surgery. Time spent not ready for care is not counted as time spent on the waiting list. What matters to the patient is the elapsed time they have spent with a painful medical condition, not the official time spent on the waiting list.

In the articles about individual experiences, the patients and by extension their carers, are the audience for the performance indicators which tell them about system averages. This is in contrast with what members of this particular audience want to know, which is how long they will have to wait for their surgery. From their perspective, the indicators are, at best, very roughly indicative and at worst misleading. It is natural to conflate 'average' and 'usual' and so to feel poorly treated if you have waited for longer than the published average time. This is reflected in the for Fairness/Equity coding, with 14 of the 17 articles containing the concept of fairness being about individual experiences.

5.2.4 Definitions

Questions A-5 and A-5a are concerned with how formal definitions relating to elective surgery are used in the newspaper articles. Elective surgery is explicitly mentioned in 58 of the 96 newspaper articles coded but is not defined in any of them. The meaning of 'elective' in the phrase 'elective surgery' is a long way from both the colloquial and the dictionary meanings of the word (see Chapter 3.3.1: Data Definitions on page 80 and Chapter 6.1.1 Understanding the language on page 158. This means that, unless they have a strong interest in the public health system or

have been involved with elective surgery in a public hospital, readers of the articles are unlikely to know what the phrase means.

The different urgency categories are used explicitly in 21 of the 96 articles, but only defined in ten of those. The categories are implied in a further four articles. Of the 25 articles, five are about individuals, 14 are about the public hospital system as a whole, while the remaining six are about individuals and the public hospital system.

As with the definition of elective surgery, the names for the urgency categories all have colloquial meanings which do not match the technical meanings. For instance, 'Urgent' elective surgery needs to take place within 30 days, which does not match the colloquial meaning of the word urgent (see Chapter 3.3.1 on page 80 and Chapter 6.1.1 on page 158).

A recurring issue uncovered in thematic coding was that of giving information relating to waiting times without reference to urgency category. In this quote from A182010 "Canberrans have the longest queues for elective surgery in the nation, with more than 10per cent waiting more than 365 days for treatment." there is no mention of urgency categories as part of the comparison to other states. In an aggregate figure like this, the mix of urgency categories for patients on the waiting lists affects whether the number is an indication of 'good' or 'bad' performance. To take an extreme and artificial pair of examples, if everyone on the lists had an urgency category of 1 (surgery within 90 days) a wait of more than 365 days is indicative of very poor performance; if everyone on the lists had an urgency category of 3 (surgery within 12 months) the performance is not nearly as poor.

This appears throughout the 'Mayor of Dickson' story with actual waiting times and the formal urgency category for one type of surgery conflated with an informal waiting time for a different type of surgery and presumable urgency category.

Question PI-3: With what is the waiting time compared? and *PI-5: What statistical assumptions are needed for the comparison to be statistically valid?* were developed as part of the instance-level coding (See Chapters 4.5.3: Instance coding on page 117, 5.3.1: Question by question results on page 145 and 5.3.3: Statistics and comparators on page 152).

5.2.5 *Explicit discussion of performance indicators*
Question A-6 was used to identify those articles which discussed the validity, relevance or usefulness of performance indicators, either on its own or in conjunction

with performance reporting. Within the 28 articles discussing performance indicators as well as stand-alone articles there were several clear groups of articles:

- a debate about the effect of running a targeted campaign to treat people who had been waiting longest;
- the ACT Auditor-general's investigation into the management of ACT elective surgery waiting lists; and
- a dispute between New South Wales (NSW) and the ACT about the costs of treating NSW patients in the ACT public hospital system.

The articles about the effect of targeting people who had been waiting longer than clinically recommended for surgery [A282010, A132011 and A1720111 and A182011] were mostly published in April and June 2011. The core issue was that there are two variations of elective surgery waiting time as a performance indicator. The first is how long a patient has waited as of a reporting date; the second is how long they waited before having surgery and being removed from the waiting list. When the ACT health directorate scheduled additional surgeries targeting people who had been on the waiting list for longer than recommended it had the effect of driving up the median waiting time for people who left the waiting list. The newspaper articles featured the ACT opposition describing the higher median as a bad result and the government claiming it as a good result.

The issue of how ACT elective surgery waiting lists are managed, including data collection and reporting rose to public prominence in June 2010 with allegations that waiting list numbers were being artificially manipulated so that performance targets were met and reward funding paid. This was highlighted in several articles referring to allocation of urgency categories which had two radically different points of view [A302010, A312010, A322010]. In two of the articles the ACT Health Directorate was accused by medical specialists of downgrading patients' urgency categories so that the clinical guideline for how long they should wait would be longer. In the other article a senior person from the directorate accused the specialists of inappropriately inflating urgency categories so that their patients were seen sooner. By late June the matter was being referred to the ACT Auditor-general for a full investigation. The Auditor-general's report was released in early 2011 with a finding that urgency categories were being downgraded without sufficient clinical reasons.

As a major regional hospital for south-east NSW, the main public hospital in the ACT provides elective surgery for patients from the parts of NSW surrounding it. Because funding for hospitals is allocated at the State level, this means that each year the NSW government pays the ACT millions of dollars for this and other hospital services. In December 2010 and December 2011 there were articles about disputes between NSW and the ACT about how had been paid and how much should be paid. The calculation of payment amounts relied heavily on the collection and categorisation of information about what services were provided by the ACT to NSW residents. Each side produced their own data and demanded an audit of the data for the other's.

In a stand-alone article in March 2010 [A072010] the ACT Chief Minister Jon Stanhope commented that he thought people were "…too fixated on waiting times rather than quality measures". This gets to the heart of performance measurement where it is far easier to measure activities than it is to measure quality.

5.2.6 *Fairness, equity and efficiency*

Questions A-4, A-4a and A-7 centre around the concepts of equity, fairness and efficiency. Equity and efficiency are fundamental performance outcomes for the public health system. Public hospitals policy, particularly that relating to the rationing of services, seeks to maximise these two outcomes (see Chapter 3.2). Fairness relates to perception of how the policy is implemented.

Over a third of the articles, 39 out of 96 refer to one or more of the concepts of equity, fairness and efficiency. No article referred to all three.

5.2.7 *Resource constraints*

As discussed in Chapter 3.3.2 *Elective surgery process overview*, a hospital's ability to manage patients' waiting times is dependent on the availability of resources such as operating theatre space, surgeons with appropriate specialties and bed space for recovery. These resource constraints are commented on directly and indirectly in newspaper articles about elective surgery. Article A032010 reports on the opening of new operating theatres which will increase the capacity of hospital to provide surgery. A key point is that the two new operating theatres are designed to allow the faster movement of patients in and out of surgery, meaning that more efficient use is able to be made of the facility. Two more general articles, A152010 and A382010

refers to the lack of bed space in wards causing elective surgery delays. In the first the reason given is the rise in hospital admissions relating to swine flu and the second is about the broader need for more public hospital beds with elective surgery mentioned as one of the services affected by shortages.

Three articles in the two-year period, A182010, A632010 and A362011 refer to holiday periods and staff leave causing delays in elective surgery. In A182010 the spike in staff leave was associated with the ANZAC Day public holiday long weekend and the Autumn school holidays in the ACT. Both A632010 and A362011 related to closures over the Christmas/New Year period which coincides with the long Summer holidays in Australia. The delays were for Category 3 or non-urgent elective surgery.

5.2.8   The Mayor of Dickson

This was a long-running human interest story about a 75 year old intellectually disabled man, Allan McFarlane, with multiple health problems and his, and his carer's, difficulties dealing with the systems and processes relating to elective surgery. The articles (A232010, A252010, A322010, A022011, A042011 and A272011) cover the period from June 2010 to just after Mr McFarlane's death in September 2011.

The first article (A232010) was on 10 June 2010 at a peak of reporting about elective surgery waiting times. It gave the background of Mr McFarlane's case and described his, and his guardian's problems dealing with the intricacies of the public health system. It is difficult to determine the precise series of events from the article but it seems that Mr McFarlane saw a specialist in May 2009 who placed him on the waiting list for the elective surgery to insert kidney stent. This surgery was designated Category 2, i.e. surgery within 90 days. At some time in June 2009 he was told by 'a specialist at the hospital' that he 'probably had prostate cancer and needed surgery in a fortnight'. There is no clear indication of formal referral for investigative surgery on the prostate and the period of two weeks does not correspond to the guidelines for any elective surgery urgency category. In the year between these events and the writing of the article Mr McFarlane was sent several letters by the ACT Health Directorate asking if he still wanted to be on the elective surgery waiting list.

The second article on 11 June 2010 was a short inset in longer article about elective surgery waiting lists in general. It reported that Allan McFarlane had been scheduled

for surgery on his prostate on 28 June 2010. There was no mention of the kidney stent.

The third article combines the waiting time for the kidney stent and the waiting time for the biopsy "…waiting since May last year to have a kidney stent inserted and a biopsy taken as part of investigations into whether he had prostate cancer."

The fourth article (A022011) was primarily about an auditor-general's report into waiting list management which used Mr McFarlane's case as an example of poor management. It describes him as "waiting for more than a year for urgent elective surgery, implying that the prostate biopsy had been classified as urgency category 1.

On 18 January 2011 there was an article reporting on the insertion of a kidney stent and the treatment of the prostate cancer with hormone injections. The case was described as having "ignited a political storm" when it was first reported in June 2010.

The final article in this series was essentially an obituary for Allan McFarlane published on 21 September 2011. This article described the case as leading to an ACT auditor-general's enquiry into the management of elective surgery in the ACT.

All of the articles focussed strongly on the personal aspects of the case. There were details of the origin of Mr McFarlane's nickname "The Mayor of Dickson", references to his friendly nature and descriptions of his pain, suffering and disability.

It was hard to tell from the articles exactly when he was referred for investigative surgery on his prostate and what the surgery eventually entailed. The reference to "needing surgery within a fortnight" has to refer to an informal comment, since the highest urgency category recommends surgery within 30 days. In only one article, the first one of 10 June 2010, was there any mention of a formal urgency category together with an explanation of what it meant.

Overall Mr McFarlane's case was framed in classic David and Goliath terms: the plucky, disadvantaged individual battling a huge, powerful and uncaring adversary, in this instance the ACT public health system.

5.3 Results of instance-level analysis

After coding for themes and concepts within each article in the collection I then looked in more detail at each instance of elective surgery waiting time being used as a

performance indicator for the public health system. For details of instance-level code development see Chapter 4.5.3 on page 117.

5.3.1 Question by question results

The first step in the instance-level analysis was to identify which articles had at least one instance of elective surgery waiting time being used as a performance indicator. The next step was to individually identify each instance.

Question PI-1: Are waiting times used as a performance indicator in the article?

Of the 96 articles from The Canberra Times relating to elective surgery 43 used waiting time as a performance indicator. There were 116 instances of this performance indicator in the collection. Table 5-11 below shows the distribution between articles of how many instances there were of waiting time being used as a performance indicator.

Table 5-11: Number of instances of waiting time being used as a performance indicator per article

Number of instances in an article	Number of articles
0	53
1	15
2	11
3	6
4	3
5	1
6	2
7	2
8	1
9	0
10	1
Total	**116**

Articles with four or more instance were those relating to the release of quarterly performance reports for the public hospital system.

Question PI-2: What type of statistic is used to describe elective surgery waiting time?

A range of different statistical constructs relating to elective surgery waiting time were used as the base statistic in each instance. These ranged from simple counts to how

long an individual had waited to medians for groups of patients. In all cases the form of average used was the median rather than the mean of the group in question. A breakdown is shown in Table 5-12 below.

Table 5-12: Distribution of base statistics

Base Statistic	Count
Count of people added to list	1
Count of people on list	15
Count of surgeries performed	9
Individual waiting time	67
Median for a mix of urgency categories	17
Median for a specific urgency category	7
Total	**116**

Question PI-3: With what is the waiting time statistic compared?

For elective surgery waiting time to be used as a performance indicator it must be compared to something. The types of comparisons used for elective surgery waiting time ranged from simple comparisons of the time waited with the guideline for the elective surgery urgency category to more complicated comparisons of the change in the percentage of people seen within the guideline for different time periods in different states. Table 5-13 below shows a breakdown of the 116 instances according to the number of comparators for the base waiting time statistic.

Table 5-13: Breakdown of instances by number of comparators

Number of Comparators	Count
1	72
2	29
3	15
Total	**116**

Table 5-14 below shows how many times each comparator was used in total and how many times it was used as part of a set of one, two or three comparators.

Table 5-14: Numbers of comparators for waiting time statistics

	Total	Alone	One of two	One of three
Count of people on list	2	0	2	0
Count of people seen within time comparator	4	0	3	1
Doctor's reported recommendation	5	5	0	0
Guideline for appropriate urgency category	46	12	24	10
Guideline for least urgent category	15	5	7	3
Individual waiting time	1	1	0	0
Median for a mix of urgency categories	1	1	0	0
Percentage of people seen within time comparator	31	0	19	12
Perception of 'a long time'	1	1	0	0
Result from previous time period	46	31	9	6
Result in other states	22	15	0	7
Target on time %	1	1	0	0
Total	**175**	**72**	**64**	**39**

Table 5-15, Table 5-16 and Table 5-17 below show the distributions of single comparators, pairs of comparators and trios of comparators.

Table 5-15: Distribution of single comparators for waiting time statistics

Single Comparators	Count
Doctor's reported recommendation	5
Guideline for appropriate urgency category	12
Guideline for least urgent category	5
Individual Waiting Time	1
Median for a mix of urgency categories	1
Perception of 'a long time'	1
Result from previous time period	31
Result in other states	15
Target on time %	1
Total	**72**

Table 5-16: Distribution of comparator pairs for waiting time statistics

Comparator pair	Count
Guideline for appropriate urgency category Count of people seen within time comparator	3
Guideline for appropriate urgency category Percentage of people seen within time comparator	16
Guideline for appropriate urgency category Result from previous time period	5
Guideline for least urgent category Percentage of people seen within time comparator	3
Guideline for least urgent category Result from previous time period	2
Total	**29**

Table 5-17: Distribution of comparator trios for waiting time statistics

Trio of comparators	Count
Guideline for appropriate urgency category Count of people seen within time comparator Result from previous time period	1
Guideline for appropriate urgency category Percentage of people seen within time comparator Result from previous time period	5
Guideline for appropriate urgency category Percentage of people seen within time comparator Result in other states	4
Guideline for least urgent category Count of people on list Result from previous time period	2
Guideline for least urgent category Percentage of people seen within time comparator Result in other states	3
Total	**15**

Question PI-4: Which part of the public health system's performance is being measured?

This question is a variation of question *A-3: Is the article about the Australian hospital system or about the ACT hospital system?* Instead of looking at the topic of the article as a whole it looks at which level of the public health system's performance is being measured in each instance of using waiting times as a performance indicator. Table 5-18 below shows the breakdown of what is being measured.

Table 5-18: Relevant part of the public health system

Entity	Count
ACT	99
National	10
Other State(s)	7
Total	**116**

Question PI-5: What statistical assumptions are needed for the comparison to be statistically valid?

Reporting in a newspaper does not require the same degree of precision as formal statistical reporting. A consequence of can be that the assumptions required for statistical validity are often implicit rather than explicitly described. Table 5-19 below shows the distribution of these statistical assumptions between instances of using elective surgery waiting time as a performance indicator.

Table 5-19: Statistical assumptions needed for comparisons to be valid

Statistical Assumption	Count
All comparators assumed	1
Doctor's recommendation is based on urgency category	6
Each urgency category compared to itself	1
General surgery is all one urgency category	1
Guidelines for the two types of surgery are < the time already waited	1
'Long' means relative to guideline for category	1
Medians tell you something about individuals	1
Rate of people joining list the same for both periods	14
Time to be removed from waiting list is the same as waiting time	1
Urgency category mix the same in each state	13
Urgency category mix unchanged over time	12
Urology patients all the same urgency	1
No assumptions	63
Total	**116**

Question PI-6: What communication assumptions are needed for the comparison to be meaningful on a human level?

As well as statistical assumptions, there are communication assumptions embedded in the commentary surrounding performance reporting. Some of these are very similar to the statistical assumptions while other relate to the meaning of words such as 'longest'. Table 5-20 below shows the distribution of these communication

assumptions between instances of using elective surgery waiting time as a
performance indicator.

Table 5-20: Communication assumptions needed for the comparison to be
meaningful

Communication assumption	No. of Uses
Doctor's recommendation is a firm guideline	6
Experience is more 'true' than statistics	1
Fewer operations is a bad outcome	1
Guideline is shorter than time comparator	2
'Longest' has a clear meaning	4
Lower proportion waiting more than a year is good	2
More surgeries is a good outcome	8
Reader can be bothered working out timings	1
'Real' waiting time is longer than 'claimed waiting time'	1
'Shortest' means fewest people on list	2
'Too long' is relative to guidelines	2
Waiting more than a year is bad	1
No assumptions	85
Total	**116**

Combining information from questions PI-5 (statistical assumptions) and PI-6
(communication assumptions) gives the distribution of the numbers and types of
assumptions associated with each instance. Table 5-21 shows this distribution.

Table 5-21: Types of assumptions per instance of waiting time being used as a
performance indictor

Type of assumptions	Count
Statistical	36
Communication	14
Both	17
Neither	49
Total	**116**

Question PI-7: What value judgement about performance is expressed?

This question is similar to question *A-2a: Are stories about individuals positive or negative?*
but at the level of each instance of using elective surgery waiting time as a

performance indicator rather than at the article level. Table 5-22 below shows the numbers of each type of value judgement.

Table 5-22: Value judgement breakdown

Value Judgement	Count
Bad	65
Good	40
Neutral	11
Total	**116**

Question PI-8: Does the value judgement reflect the meaning of the statistics?

The match between the value judgement expressed for a performance indicator and its strict statistical meaning acts as a proxy for how easy it is to understand the performance indicator. Table 5-23 below shows the number of mismatches, partial matches and complete matches between value judgement and statistical meaning.

Table 5-23: Match between value judgement and statistical meaning

Does value judgement match statistic?	Count
No	1
Partly	51
Yes	64
Total	**116**

5.3.2 Combining instance-level codes

Looking at combinations of question coding shows relationships between the different attributes of each instance of waiting time being used as a performance indicator:

- combining coding for question PI-2 (Type of statistic) with that for question PI-3 (comparators) gives a more detailed picture of the way waiting time is used as a performance indicator;
- combining data about the value judgement for an instance with the degree of match between it and the statistical meaning of the instance gives more detail about these aspects of the data;
- combining data about the number and type of assumptions (PI-5 and PI-6) for an instance with the match between value judgement and statistical

meaning (PI-7) gives more information about the relationship between the complexity of the performance indicator and how easy it is to understand; and

• combining data about the numbers and types of comparator (PI-3) with the match between value judgement and statistical meaning (PI-7) also gives more information about the relationship between the complexity of the performance indicator and how easy it is to understand.

Each of these combinations is expanded on in the following sections.

5.3.3 Statistics and comparators

The results for question PI-2 and PI-3 are combined in

Table 5-24 to give a composite a picture of the elective surgery waiting time statistics used in the newspaper articles. Rows with a count of 0 are omitted from the table.

Table 5-24: Statistics and comparators for waiting time as a performance indicator

Base Statistic	Comparators	No. of uses
Count of people added to list	Result from previous time period	1
Count of people on list	Guideline for least urgent category	1
	Result from previous time period	10
	Result in other states	4
Count of surgeries performed	Result from previous time period	9
Individual waiting time	Doctor's reported recommendation	5
	Guideline for appropriate urgency category	12
	Guideline for least urgent category	2
	Individual waiting time	1
	Median for a mix of urgency categories	1
	Perception of 'a long time'	1
	Result from previous time period	1
	Guideline for appropriate urgency category Count of people seen within time comparator	3
	Guideline for appropriate urgency category Percentage of people seen within time comparator	16
	Guideline for appropriate urgency category Result from previous time period	5

Base Statistic	Comparators	No. of uses
	Guideline for least urgent category Percentage of people seen within time comparator	3
	Guideline for least urgent category Result from previous time period	2
	Guideline for appropriate urgency category Count of people seen within time comparator Result from previous time period	1
	Guideline for appropriate urgency category Percentage of people seen within time comparator Result from previous time period	5
	Guideline for appropriate urgency category Percentage of people seen within time comparator Result in other states	4
	Guideline for least urgent category Count of people on list Result from previous time period	2
	Guideline for least urgent category Percentage of people seen within time comparator Result in other states	3
Median for a mix of urgency categories	Guideline for least urgent category	1
	Result from previous time period	4
	Result in other states	11
	Target on time %	1
Median for a specific urgency category	Guideline for least urgent category	1
	Result from previous time period	6
Total		**116**

5.3.4 *Agreement between statistics and value judgements*

The value judgements expressed in the newspaper articles and how they compare with the statistical meaning give an insight into how the writer of the article, or the writer of the source material they used, interpreted the statistical data. Table 5-25 below shows the breakdown of how well the value judgements and statistics matched for each of the three value judgements.

Table 5-25: Combining value judgement and match with statistic

Value Judgement	Match	Count
Bad	No	1
	Partly	35
	Yes	29
Good	No	0
	Partly	14
	Yes	26
Neutral	No	0
	Partly	2
	Yes	9
Total		**116**

5.3.5 Statistics, comparators and assumptions

Two factors which might influence the communication of performance indicators are the complexity of the statistic and the complexity of the comparators. Table 5-26 below shows a breakdown of coding for PI-8 (match between value judgement and statistic) by PI-2 (type of statistic). Table 5-27 below shows this breakdown by the coding for comparators from PI-3.

Table 5-26: Match between value judgement and statistic by type of statistic

Match	Type of Statistic	Count
No	Individual waiting time	1
Partly	Count of people on list	15
	Count of surgeries performed	1
	Individual waiting time	19
	Median for a mix of urgency categories	16
Yes	Count of people added to list	1
	Count of surgeries performed	8
	Individual waiting time	47
	Median for a mix of urgency categories	1
	Median for a specific urgency category	7
Total		**116**

Table 5-27: Match between value judgement and statistic by comparator

Match	Type of comparator	Count
No	Median for a mix of urgency categories	1
Partly	Count of people on list	2
	Doctor's reported recommendation	5
	Guideline for appropriate urgency category	2
	Guideline for least urgent category	11
	Percentage of people seen within time comparator	4
	Perception of 'a long time'	1
	Result from previous time period	20
	Result in other states	17
	Target on time %	1
Yes	Count of people seen within time comparator	4
	Guideline for appropriate urgency category	44
	Guideline for least urgent category	4
	Individual Waiting time	1
	Percentage of people seen within time comparator	27
	Perception of 'a long time'	0
	Result from previous time period	26
	Result in other states	5
Total		**175**

Another way I looked at the communication of performance indicators using statistics was to look for patterns in the results for question PI-8 (match between value judgement and statistic) when matched with characteristics of the statistical information (type of statistic from question PI-2, number of comparators from PI-3 and number of statistical and communication assumptions from PI-5 and PI-5). The result is shown in Table 5-28 below. Rows with a count of zero have been omitted from the table.

Table 5-28: Match between statistic and value judgement by type of statistic, comparators and assumptions

Match	Type of Statistic	No. Comparators	No. Assumptions	Count
No	Individual waiting time	1	1	1
Partly	Count of people on list	1	1	9
		1	2	6
		2	1	2
	Count of surgeries performed	1	2	1
	Individual waiting time	1	0	1
		1	1	4
		1	2	6
		2	1	1
		2	2	1
		3	1	3
		3	2	1
	Median for a mix of urgency categories	1	1	16
	Count of people added to list	1	0	1
Yes	Count of surgeries performed	1	1	8
	Individual waiting time	1	0	9
		1	1	1
		1	2	1
		2	0	22
		2	1	3
		3	0	10
		3	1	1
	Median for a mix of urgency categories	1	1	1
	Median for a specific urgency category	1	0	6
		1	1	1
Total				**116**

This chapter described the results from several different analyses of newspaper articles in *The Canberra Times* relating to elective surgery. Matching the patterns in the article collection with the timeline of events in the National Health Reform Agenda showed that peaks in the number of newspaper articles about elective surgery coincided with major events in the reform process. The article level analysis showed

that the technical terms relating to elective surgery were used far more often than they were defined, that fairness equity and efficiency were of importance in the reporting and that individual stories were a critical part of the overall reporting on the reform process. The instance-level analysis showed that there were patterns in the match between value judgements about hospital performance, the type of base statistic in the performance indicator and the number and type of assumptions needed for a comparison to be valid and meaningful. This last point will be explored further in the next chapter.

The next chapter will discuss these results, together with material from Chapter 3 relating to the political, funding and contextual background for elective surgery, in the light of relevant theory from the literature review in Chapter 2.

6   Discussion and analytical model

Elective surgery waiting times are used to communicate information about the performance of the Australian public hospital system in a range of ways, from formal reports to the Council of Australian Governments to human interest stories in the daily newspaper. For the purposes of this analysis, these uses will be looked at from two different perspectives which will then be combined. The first perspective is related to the audience's understanding of the performance indicator from the viewpoints of language, statistics and context. The second perspective is related to the stories that use the performance indicators, their characteristics and the value judgements they contain. The two perspectives will then be combined to look at how closely the stories with their value judgements match the technical meaning of the performance indicator and what factors might affect how well they match.

6.1   Understanding the performance indicator

There are three parts to communicating performance information with a numerical indicator: the language used, the statistics used and the context in which the communication takes place.

6.1.1   Understanding the language

The phrase 'elective surgery waiting list' uses well-known English words to describe a highly technical construct. For each part of the phrase: 'elective surgery', 'waiting' and 'list' there is a gap between its colloquial meaning and its technical meaning. As a consequence the whole phrase has a technical meaning far removed from that obtained by combining the colloquial meanings of its parts.

In the context of the Australian public hospital system, elective surgery is defined as any surgery that a patient's doctor or health professional considers to be necessary but which can be delayed by at least 24 hours. In contrast to this technical usage, the Oxford English Dictionary defines elective as follows:

> *Adjective*
>
> *II. Pertaining to choice in general.*
>
> *4. Pertaining to the action of choosing. Of actions: Proceeding from free choice, optional, voluntary. Formerly Obs., but now revived in medical use: optional, not urgent (see quot. 1941).*

1941 Dorland's Med. Dict. (ed. 19) 476/2 Elective, subject to the choice or decision

of the patient or physician · ·, applied to procedures that are only advantageous to the

patient but not necessary to save his life.

(Oxford University Press 2012)

Over time the technical meaning in Australia has moved from the 1941 'optional, not urgent' to the current 'necessary, not urgent'. The dominant colloquial meaning remains as 'optional'. From Table 5.1 on page 133, we can see that elective surgery is explicitly mentioned in 58 of the 98 articles in the collection of articles from *The Canberra Times* while Table 5.7 on page 135 shows that elective surgery is not defined in any of the articles. This means that, unless the reader had direct experience with the way the term is used in the Australian hospital system, they were likely to have a basic misunderstanding of what is meant by 'elective surgery'.

There are similar problems with the term 'waiting'. In technical terms, a patient is waiting for surgery if they have been referred for surgery by a specialist and are classified as 'ready for care' (see Chapter 3.3.3 on page 85 for details of the administrative processes for elective surgery in the ACT). Any time a patient spends not 'ready for care', for example if they are too unwell to undergo surgery, is not counted in official waiting time statistics. From the point of view of a patient or carer, they are waiting for surgery from the time a health professional tells them that they need it. This shows clearly in story A292011 *Experience leaves patient sick of the hospital system* (see page 139 for details of the story) where a patient is talking about how long they have waited and the repeated delays to their surgery arising from their other medical problems. In the case of this patient, not even the medical specialists involved in his case agreed on whether or not he was fit to undergo surgery. The tone of the story is negative about the public health system and clearly shows that to the patient what matters is elapsed time, not time spent officially on the waiting list.

The Australian Medical Association argues that the waiting time should be calculated from the time a patient is first referred to a surgical specialist by a general practitioner (Australian Medical Association 2011; Australian Medical Association 2011) rather than from the time the specialist refers them for surgery. Patients who undergo surgery naturally perceive their waiting time to begin as soon as they are told they need it. A serious impediment to including this time waiting to see the

surgical specialist in the overall elective surgery waiting time is that 30% of patients referred to a surgical specialist are not referred for surgery.

There are two different numbers in the formal reporting which relate to how long someone has waited for elective surgery:

- Definition 1: How long the person has been on the waiting list, without having surgery, up to a given census date.
- Definition 2: How long the person was on the waiting list before they had their surgery and left the waiting list.

The critical difference between these two indicators is clearly illustrated in a series of articles from The Canberra Times in April and June 2011 [A282010, A132011 and A1720111 and A182011]. The ACT Government was criticised by the Opposition because waiting times for elective surgery went up. The response was that this was a side effect of an initiative to target people who had been on the waiting list for too long, so there was a spike in people who had waited a long time leaving the waiting list. In other words an increase in the waiting time according to Definition 2 above was related to an improvement in health outcomes for patients who had waited the longest and would have been accompanied by a decrease in waiting times according to Definition 1. The initiative to target people who were overdue for surgery was part of the National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan (COAG Reform Council 2009) which is described in Chapter 3.2.4 on page 70.

The words associated with elective surgery waiting lists and the people on those lists are all to some extent passive and imply a simple, orderly list of people queued up waiting for surgery. This has the potential to set up the expectation that, if you are a patient, you have a place in a queue of people and you are steadily moving up the queue until it is your turn to be treated. What is not at all apparent from this language of waiting lists and queues is the intensely dynamic nature of the system.

From the time a person is placed on a waiting list for surgery their place on that list is subject to constant change for many reasons:

- They may move relative to other patients because their condition changes so much that they are classified into a different urgency category;

- Other patients may move relative to them according to changes in urgency category and clinical priority; and

- Factors such as the availability of operating theatres, surgeons and hospital beds may speed up or slow down the rate at which the overall queue moves.

The mechanics of these changes are covered in Chapter 3.3.3 on page 85. The final point above means that the queues for different types of surgery can affect each other and that the overall capacity of the public hospital system can affect elective surgery waiting times. In the ACT this is further complicated by each surgeon having their own waiting list, rather than there being a consolidated list for each type of surgery. The effect of resource constraints within the wider hospital system on elective surgery waiting times was clearly shown in an April 2010 newspaper article [A152010] which reported that an increase in hospital admissions due to the outbreak of swine flu had caused non-urgent elective surgery to be delayed due to a lack of beds for pre- and post- operative care.

The contrast between the process implicit in the terms used and the actual processes could account for much of the perceived 'unfairness' and apparently arbitrary nature of the system from the point of view of the patients with 17 out of 96 newspaper articles referring to fairness and all stories relating to individual experiences being negative about the public hospital system.

There are two essential parts to a performance indicator, the measurement and what it is being compared to. A measurement, such as how long the average wait is for elective surgery in the ACT, has little use as an indicator of performance unless it is compared to something: a benchmark, a guideline, what is happening elsewhere or even a personal expectation of what is reasonable. For elective surgery, one of the important sets of comparators is the clinical guidelines for the maximum clinically desirable waiting time for each urgency category. As with the phrase 'elective surgery waiting list', the language used to describe the urgency categories involves words with both technical and colloquial meanings.

As outlined in Chapter 3.3.1 on page 80, there are three urgency categories for elective surgery defined in the National Health Data Dictionary (Australian Institute of Health and Welfare 2010):

- Urgency Category 1: Admission within 30 days desirable for a condition that has the potential to deteriorate quickly to the point where it may become an emergency.

  This is sometimes referred to as "Urgent"

- Urgency Category 2: Admission within 90 days desirable for a condition causing some pain, dysfunction or disability but which is not likely to deteriorate quickly or become an emergency.

  This is sometimes referred to as "Semi-urgent"

- Urgency Category 3: Admission at some time in the future acceptable for a condition causing minimal or no pain, dysfunction or disability, which is unlikely to deteriorate quickly and which does not have the potential to become an emergency.

  This is sometimes referred to as "Non-urgent"

The formal reports to COAG on elective surgery waiting times used both the urgency category numbers and the names while the MyHospitals website used Urgent', Semi-urgent' and 'Non-urgent'. As with the term 'elective', the three verbal descriptions use words in common usage as labels for concepts with precise technical definitions. The primary definition of 'urgent' in the Oxford English Dictionary is:

> *ADJECTIVE*
> *1. Requiring immediate action or attention:*
> *e.g. an urgent demand for more state funding*
> (Oxford University Press 2012)

The modifiers 'semi' and 'non' are defined as:

> *Semi-*
> *Partly; in some degree or particular*
> *e.g. semi-conscious*
> *or*
> *Almost*
> *e.g. semi-darkness*
> (Oxford University Press 2012)

> *Non-*
> *Expressing negation or absence:*

e.g. non-aggression, non-recognition

(Oxford University Press 2012)

Once again there is a disconnect between the technical and the colloquial meanings. This is most apparent with Category 1 or Urgent surgery. Waiting for 30 days to be treated for an urgent condition may be within the clinical guidelines, but in everyday language it sounds like a very long time.

Confusion over the intended meaning of a word or phrase is not confined to those with both technical and colloquial meanings. The way elective surgery waiting times are compared between jurisdictions is a case in point. Some articles [A162010, A182010, A242010, A282010, A502010, A632010, A132011, A242011, A252011] refer to longest waits or longest median waiting times, a few others [A182010, A21010, A302010] refer instead to longest waiting lists. Grammatically speaking, the latter case could mean either the greatest number of people on the list or longest waiting times. Article A182010 refers to both the longest waiting times and the longest queues. Since the ACT has a small population compared to other states/territories it is unlikely that there would be more people on the list than in other places, making the logical meaning that elective surgery patients in the ACT wait the longest.

There is further potential for confusion in phrases such as 'waiting too long' and 'experiencing long waits'. According to the definitions related to elective surgery reporting, a person has waited too long if they have been on the waiting list, i.e. referred for surgery and ready for care, longer than the clinically recommended time for their urgency category. If a patient is experiencing severe discomfort or anxiety a wait that is within the clinical guidelines could easily be perceived as being unreasonably long. This is exacerbated when there is conflict between the appropriate waiting time according to the guidelines and the appropriate waiting time as perceived by patients and their medical practitioners. A clear example of this occurs in the newspaper articles relating to the Mayor of Dickson outlined in Chapter 5.2.8 on page 143. In the articles the patient is referred for Category 2 surgery (surgery required in 90 days) for one condition and also told that they need surgery for another condition within two weeks [A232010]. Understandably the patient was confused about what surgery list they were on and how long they should wait. The two-week period does not correspond to any of the clinical guidelines for

elective surgery. The repetition of the need for surgery within two weeks throughout the series of articles shows that it was the time frame which registered with the patient and his carer providing the anchor for all their perceptions of what constituted waiting too long.

In A532010 the sentence "Waiting lists in the ear, nose and throat area tended to be longer than desirable because most procedures were categorised as non-urgent." at first seems to suggest that non-urgent is an inappropriate categorisation but later in the article it becomes clear that the reason is not the categorisation but that the lowest urgency surgeries are the most likely to be delayed in favour of more urgent surgeries.

Another potential source of misunderstanding is in the differences between how long a surgeon (or other health professional) says someone should wait for surgery, the specific urgency category assigned and the elapsed time the patient waits. The first and the last are the really important ones for individuals; the system-wide performance indicators are based on the second. This shows in the Mayor of Dickson articles and is particularly clear in article A292011 which deals with a patient whose surgery is repeatedly delayed because of a lack of agreement among specialists about whether or not they are ready for care (see page 139 for details of the article).

Looking at Table 5-20 on page 150 the most common communication assumptions are:

- More surgeries is a good outcome (8);
- Doctor's recommendation is a firm guideline (6)
- Longest has a clear meaning (4); and
- 'Too long' is relative to guidelines (2).

Table 6-1 below shows the how well the value judgements and statistical meanings match for each instance with one of those communication assumptions.

Table 6-1: Match between value judgement and statistical meaning for different communication assumptions

	Yes	Partly	No	Total
More surgeries is a good outcome	8	0	0	8
Doctor's recommendation is a firm guideline	1	5	0	6
'Longest' has a clear meaning combined with 'Too long' is relative to guidelines	2	4	0	6

The match for the first assumption, 'More surgeries is a good outcome' can be traced back to the performance benchmarks in the *National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan* (COAG Reform Council 2009). The performance measures for Part One of the plan were increases in the volume of elective surgery in all states and territories. The assumption behind these measures was that increasing the amount of surgery would decrease waiting times and the number of people on the waiting lists for surgery (See Chapter 6.3.1 on page 170 for further discussion of the effects of the agreement).

The language used for describing and reporting on elective surgery waiting times and the performance of the public hospital system uses many common words as labels for technical concepts. In formal documents such as reports to COAG these terms are typically defined and used consistently. In the newspaper articles there is often no definition of formal terms, inconsistency in how they are used and a blurring of the distinction between the technical and the colloquial meaning of words such as 'urgent' and 'elective'.

6.1.2 Understanding the statistics

When elective surgery waiting times were used as performance indicators they were described using a range of statistical constructs for both the performance indicator and up to three comparators. In addition, for many of these comparisons, there were unwritten assumptions needed for them to be statistically valid. All three factors: complexity of statistical constructs, comparator complexity and statistical assumptions affect how easy it is to understand the statistical aspects of the performance indicators.

The statistical constructs used ranged from very simple such as counts of people on or added to the waiting list and individual waiting times, or more complex such as median waiting times for specific urgency categories and median waiting times for a mix of urgency categories. The distribution of the types of statistic used as the basis of the performance indicator is shown in Table 5-12 on page 146 and which comparators are used for each base statistic in Table 5-24 on page 152. Combining the information in these tables and grouping by type of statistical construct gives Table 6-3 below which shows how often each statistical construct is used as a base statistic, as a comparator and in total. For this table, the base statistics and comparators were recoded as shown in Table 6-2.

Table 6-2: Recoding for statistical constructs

Original code	New Code
Count of people added to list	Count
Count of people on list	Count
Count of people seen within time period	Count
Count of surgeries performed	Count
Doctor's reported recommendation	Individual waiting time
Guideline for appropriate urgency category	Individual waiting time
Guideline for least urgent category	Individual waiting time
Individual waiting time	Individual waiting time
Median for a mix of urgency categories	Median
Median for a specific urgency category	Median
Percentage of people seen within time comparator	% in time comparator
Perception of 'a long time'	Individual waiting time
Result from previous time period	Appropriate base statistic or comparator
Result in other states	Appropriate base statistic or comparator
Target on time %	% in time comparator

The second last column in Table 6-3 shows the number of times each statistical construct is used over all instances of waiting time being used as a performance indicator. The final column in Table 6-3 shows the number of instances where each statistical construct was used, either as a base statistic or a comparator.

Table 6-3: Distribution of statistical constructs in base statistics and comparators

Type of statistical construct	Base statistic	Comparator	Total times used	Number of instances using construct
% in time comparator	0	44	44	32
Counts	25	30	55	27
Individual waiting time	67	79	146	70
Median	24	22	46	25

A statistical construct used 44 times in the 116 instances analysed was the percentage of people seen within a particular time comparator. For example, "elective surgery provided on time for 95 per cent of patients" [I01102011]. In all cases the construct was one of the comparators, not the base statistic. Percentages can be a useful way to compare proportions of populations between groups of different sizes although the comparison depends upon the assumption that there are no economies of scale for large groups. In the Australian school system the concept of a percentage is introduced as part of the compulsory Mathematics curriculum in Year 6 and developed as part of problem solving in Year 7 when children are 12-13 years old. Despite this a report from the Productivity Commission (Shomos and Forbes 2014) found that 54.5% of the Australian population had insufficient numeracy skills to deal with any but the simplest data and statistics. Under the definitions in the report, 22% of the adult population did not have the skills to deal comfortably with comparative percentages.

Of the 32 instances where '% in time comparator' was used as one of the comparators, the value judgement expressed about hospital system performance fully matched the statistical meaning in 27 cases and partly matched in the remaining 4. This indicates that the journalists writing for *The Canberra Times* were over all correctly using the statistical material in the reports and press releases underlying their articles.

The potential for communication problems with the simplest statistics, the counts, lies in the complexities of the comparators used and the assumptions underlying them. Complexity in the comparators relates to the statistical constructs of the comparators, the number of comparators and how multiple comparators are combined. Table 5-

24 on page 152 shows which comparators were used for each type of statistic. For the three count-based statistics (Count of people added to the list, Count of people on the list and Count of surgeries performed) the comparators are nearly all simple comparisons of results from previous time periods or results in other states. The assumptions most relevant to the validity of these statistics is that the rate of people joining the elective surgery waiting list is the same for both time periods and/or jurisdictions and that the mix of urgency categories remains unchanged for both time periods and/or jurisdictions. For the 27 instances using one of the three types of counts, the value judgement partly matched the statistical meaning in 18 cases and fully matched in 9. Of the partial matches, 17 were counts of people on the waiting list with either the assumption that the rate of people joining the list was constant or that the mix of urgency categories was unchanged. Both these assumptions are unlikely to be true, meaning that the comparisons are invalid.

Individual waiting time was used 67 times as a base statistic and 79 times as a comparator, making it by far the most common statistical construct in the collection of newspaper articles. On the surface individual waiting time is a simple number but, as discussed earlier in this chapter, in the context of waiting for elective surgery, 'waiting time' has a complex technical meaning which does not readily match the colloquial meaning of the words. In other words, many of the communication problems for this statistical construct are likely to be language-related.

Breaking down the 70 instances which had at least one use of individual waiting time according to how well the value judgement matched the statistical meaning gives one complete mismatch, 20 partial matches and 49 complete matches. In the case of the complete mismatch, the fact that an individual had been waiting far longer than the recommended time for their urgency category was cited as evidence that reports of the median waiting time dropping were untrue (See description of article A202011 on page 138 for details). The article relies on the mistaken assumption that a median for a population gives information directly about an individual within that population. The partial matches use imprecise or inferred comparators such as 'Guideline for the least urgent category' and 'Doctor's recommended waiting time'. The vast majority, 44 out of 49, cases of a full match between statistical meaning and value judgement compare individual waiting time to the guideline for the appropriate urgency category.

In almost 21% of instances of waiting time being used as a performance indicator the base statistic was a median, either for a single urgency category or for an unspecified mix of urgency categories. Table 5-24 on page 152 shows that the comparators for these statistics were: Guideline for least urgent category, Result from previous time periods, Result in other states and the percentage target for on-time surgeries. Two of these, Result in pervious time period (10 instances) and result in other state (11 instances) are themselves medians.

The concept of a median is introduced in Year 7 of the Australian school system, when students are approximately 13 years old. The Productivity Commission report (Shomos and Forbes 2014) mentioned earlier, found that 32.5% of Australian adults did not have the numeracy skills to use or understand the concept of a median.

The 25 instances which had at least one use of a median can be broken down according to how well the value judgement matched the statistical meaning. This gave one complete mismatch, 16 partial matches and 8 complete matches. The complete mismatch is between an individual waiting time and a median and is discussed above. All but two of the partial matches were comparing the medians for a mix of urgency categories and hence had the underlying assumption that the mix of urgency categories was the same in the two jurisdictions or time periods being compared. The complete matches either compared individual waiting time to the median for appropriate urgency category or compared medians for the same urgency category.

Under the constructivist approach to communication, each audience member constructs the meaning of what they read or hear through the interaction between what they already know and the communication materials (Yager 1991; Stocklmayer 2013). This means that just over half the Australian adult population does not have the underlying knowledge necessary to correctly interpret a statistic expressed as a median and almost quarter would have problems comparing percentages. Even if the journalists writing about elective surgery waiting times and public hospital performance had the skills to use and understand the statistical information in hospital performance reports, it is likely that many of their readers did not and that the readers lacked the knowledge to critique or question the way statistics were used in the newspaper articles.

Comparing like with like is part of the bedrock of statistical description and analysis. For each type of statistical construct, many of the mismatches between statistical meaning and value judgements can be traced back to breaching this principle whether it be by assuming that a mix of urgency categories is constant, assuming that the rates of people joining the list are constant or that a doctor's reported recommendation is the same as assigning a formal urgency category.

6.1.3   Understanding the context and usage

The over-arching context within which elective surgery waiting times were used as performance indicators in *The Canberra Times* was that of the overhaul of Australian public hospital funding and performance measurement. This context is described in detail in Chapter 3.2 starting on page 61.

As described in Chapter 3-2, elective surgery waiting list lengths and elective surgery waiting times, along with emergency department waiting times, were highlighted as key indicators of public hospital performance. The outcome of the first series of agreements between the Australian State governments and the Commonwealth government set out in paragraph 3 of the overarching *National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan* (COAG Reform Council 2009) was to 'reduce the number of Australians waiting longer than clinically recommended times for elective surgery by improving efficiency and capacity in public hospitals'. This single statement encompasses three separate outcomes: reducing the number of people waiting longer than clinically recommended; improving the efficiency of the public hospital system; and improving the capacity of the public hospital system. The implication is that the first outcome will be achieved if the second two are achieved. The agreement provided some upfront funding divided among the states and territories, $150 million to bring about an immediate reduction in people waiting longer than recommended; $150 million for system and infrastructure improvements; and funding of up to $300 million based on how well jurisdictions met a series of performance targets.

The performance targets related to numbers of elective surgeries performed and the number of patients waiting longer than clinically recommended (see Chapter 3.2.4 starting on page 70 for details of the agreement, targets and performance against targets). The interaction of the two targets attracting reward funding was complex and it would have been difficult for each jurisdiction to work out the trade-offs

between performing as many surgeries as possible and reducing the number of overdue surgeries. The performance reports for this agreement show that each jurisdiction met and exceeded its elective surgery volume targets (see Table 3-5 on page 72) and so received the reward funding. Only three of the eight jurisdictions met their target for reducing the number of patients with overdue surgery (see Table 3-7 on page 74) while the national median waiting time for elective surgery increased from 34 to 35 days. Table 6-4 below shows the changes in median waiting times for individual states between the two time periods.

Table 6-4: Changes in median waiting time for elective surgery between 2007-2008 and 2009-2010 by state/territory

State	2007-2008	2009-2010
NSW	38	44
VIC	32	36
QLD	27	27
WA	31	32
SA	42	36
TAS	36	36
ACT	73	73
NT	43	44
Australia	**34**	**35**

The fact that most states met overall surgery volume targets but did not meet the targets for primary goal of the agreement shows that the links between them are far more complex that they appear on the surface the two targets were. The rise in the median waiting times was also an indicator that increasing surgery volumes was not necessarily having the desired effect on waiting lists.

A further complication it that the rate of people joining the elective surgery waiting list is not constant and seems to be affected by how long the list is at a given time. This issue was raised by the ACT Chief Minister and Health Minister in December 2011 when she said there was an 'uncanny' surge in additions to the waiting list every time significant inroads were made [A352011].

The pressure to meet targets and receive reward funding was intense and, in the ACT, there were accusations that the Health Directorate was manipulating patients' urgency categories to improve the waiting time data (see Chapter 5.2.5 on page 140).

By tying reward funding to the achievement of a particular activity measure, the COAG reform process encouraged jurisdiction to work to maximise performance according to the measure rather than according to the overall outcome. As observed by Donella Meadows back in 1998 (Meadows 1998) the feedback loop between measuring what we value and valuing what we measure is "common, inevitable, useful and full of pitfalls" (p2, Meadows 1998).

A continuing problem in the use of reward-based funding was the lack of uniform reporting standards coupled with a lack of strong penalties for manipulating data. A paper written during the reform process (Nocera 2010) looks at the early effects of using waiting list data to allocate performance funding, asking whether the process works as a tool for reform or an incentive for fraud. As well as documenting cases where data was manipulated he gave an extreme example where work practices appeared to have been altered so much in the name of improving performance that patients were being injured and in some cases dying. One of the paper's conclusions was that public sector data fraud should be a criminal offence.

The *National Health Reform Agreement – National Partnership Agreement on Improving Public Hospital Services* (COAG Reform Council 2011) followed on from the *National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan* (COAG Reform Council 2009) and its performance targets included additional constraints in an effort to counteract the perceived 'gaming' of the original system. The reward funding was based on the proportional reduction in the number of patients who have waited longer than the recommended time with two constraints:

- Urgency Category 1 cases were all to be seen within clinically recommended times by the end of 2012; and
- The 10% of patients within each urgency category who have waited the longest must have their procedures in the reporting year (COAG Reform Council 2011).

In their work on public health performance indicators Van Peursem, Pratt et al. gave three guidelines for applying performance measurement:

> *A balance of ordinal, nominal and ratio indicators should be produced to avoid the impression that precision has been achieved, as well as to provide a more balanced view.*

> *The way in which they are measured, as well as the measures themselves, needs to be an open, communicated process. It may be advisable to disclose that process and the participants who engage in it.*
>
> *It should be made clear that measures are an indication of a situation which may call for further enquiry. Indicators do not provide answers, they inspire questions, and this should be made clear (p 60, Van Peursem, Pratt et al. 1995).*

Looking at the use of elective surgery waiting time data in the light of these guidelines, they partially meet the first guideline in that there are ordinal indicators (ranking the performance of states and territories) and ratio indicators (median waiting times, number of surgeries etc. They also do not fully meet the second guideline. While it was possible to find public information about how the data was collected, it took weeks of work to trace all the details and synthesise them into the narrative found in Chapter 3.3. Guideline three, arguably the most important, is not met in the available public information about elective surgery waiting times.

6.2 How is the performance indicator used in newspaper stories?

In the collection of stories from *The Canberra Times* elective surgery waiting time is used as a performance indicator in a range of contexts. Many of the concepts from the literature covered in Chapter 2 are reflected within the collection. Three which stand out are health services rationing, the importance of stories and the limitations of statistics. The choice and framing of stories relating to elective surgery reflect many of the concepts covered in Section 2.3.1.2 which discussed the values underlying the selection of news stories.

6.2.1 Health services rationing

In Chapter 2.3 I reviewed literature relating to public health policy and practice. A key social principle underlying the provision of public services in general and public health care in particular is that of providing help to those who are seen to need it, regardless of resource constraints. In some of the literature (Hadorn 1991; Hadorn 1996; Nord, Richardson et al. 1995; Nord, Richardson et al. 1995; Stewart 2009) this is referred to as the Rescue Principle or the Rule of Rescue. Stewart refines this by distinguishing between 'equity' which relates to formal decisions about resource

allocation and 'fairness' which is concerned with public perceptions about resource allocation.

One or both of the concepts of fairness and equity appear in 30 out of the 96 newspaper articles analysed. Of these, nine were about an individual patient's experience, fourteen were about the hospital system and seven were about both. The majority of the articles, 26 out of 30 dealt with the hospital system in the ACT and surrounding region. All the articles containing an individual patient's experience related to the hospital system in the ACT and surrounding region. Discussions relating to fairness concentrated on patients who had not been treated within the appropriate urgency category timeframe or who had not received treatment within the most recently released median timeframe.

There were also six articles dealing explicitly with resource constraints within the public hospital system (See Chapter 5.2.7 on page 142), drawing attention to the effect of finite resources on the system's ability to provide timely access to public elective surgery patients. The resources were so tight that all except the most urgent surgery was being cancelled during the Christmas/New year Holiday period.

6.2.2 The importance of stories

If a newspaper is to be read and sell, it must publish articles that appeal to a sufficiently large audience. Drawing on the work of Pratchett, Stewart et al; Dunbar and McRae (Dunbar 1996; Dunbar 1998; Pratchett, Stewart et al. 1999, McRae 2011), it is clear that humans and human relationships are at the heart of how we tell stories. In order to understand and care about a problem, humans respond more strongly to descriptions based on stories about individuals than those using facts and statistics. This shows up in research relating to the Rule of Rescue and its counterparts (Hadorn 1991; Hadorn 1996; Nord, Richardson et al. 1995; Nord, Richardson et al. 1995; Stewart 2009). This can be seen in the way human stories were used to illustrate some of the newspaper articles about elective surgery. As well as seven articles being about both individual experiences and the wider hospital system, a further four out of the nine articles about individual experiences appeared on the same day as a more abstract article about elective surgery waiting times.

All the stories about individual patients were negative in tone about their experience. Every person selected to be written about had faced, or was still facing, a long delay

for their surgery. They had also found the system hard to understand and the communication from the hospital confusing. Even when additional funding had increased the number of surgeries and reduced the number of people waiting, the patient featured was one who had not benefited from the funding, was still waiting for their surgery and expressed disbelief that the overall situation was improving. The predisposition for new stories to refer to bad rather than good news can be traced back to the core news values found in journalism textbooks (Conley 2002; Harrower 2007). These values are impact, immediacy, proximity, prominence, novelty, conflict, and emotional content. The negative stories emphasise conflict and the emotions of sadness, pity, outrage and anger. A paper about the Australian health system in 2010 opens by stating that elective surgery waiting lists serve as a newsworthy focus for broader discontent with the public hospital system (Curtis, Russell et al. 2010).

In the wider political landscape, public healthcare accounts for a huge amount of public expenditure, $140.2 billion AUS in 2011-12 (AIHW, 2013), the period studied in this research. The case study period was one of rapid change and public debate about the public health system making it a key political news story. In line with the principles of newsworthiness outlined in Chapter 2.3.1.2, the news stories relating to negotiations between different levels of government about public hospital governance and funding were framed in terms of conflict between the parties and the effects of any shortcomings in the system on clearly identified individuals. Looking back at Table 6-1 on page 165 which analyses the match between statistical meaning and value judgement shows that, even with the framing constraints of writing a newsworthy story, any errors and ambiguities in the way performance statistics were reported are broadly traceable to factors such as statistical complexity and the problems of comparing like with like.

One of the longest-running stories in the collection of articles was that of a disabled man nicknamed The Mayor of Dickson (see Chapter 5.2.8 on page 143). The story was told in six articles from June 2010 to September 2011. The articles were all written with a strong emphasis on the vulnerability of the protagonist and the unfairness of his treatment by a faceless, uncaring state health system. Many of the core news values were evident in this series of stories:

- Immediacy: after the first one, each article was presented as an update in a continuing story;

- Proximity: the articles made frequent mention of the protagonist's role in his local community and close relationship with the people in his part of Canberra;

- Novelty: one of the articles in the series focussed on the protagonist's cheerful, quirky personality;

- Conflict: the protagonist's experience was framed as a David and Goliath conflict between a powerless but righteous individual and a large, powerful opponent;

- Emotional content: the protagonist was described as vulnerable and deserving of pity while the people within the health system were described as uncaring. The protagonist's carer described her anger and frustration resulting from their experience with the hospital system.

The power of the personal story and its emotional content completely overshadowed any narrative relating to performance data.

6.2.3 The limitations of statistics

A recurring theme in the collection was that the statistics don't measure what is important and don't provide patients with useful information. This was apparent both in stories relating to individual experiences and in more general discussions about hospital system performance.

In his early work on organisational performance management, W. Edwards Deming's noted that "The most important things are unknown or unknowable" (p97-98 Deming 1982 (reprinted 2000)). This insight into the limitations of performance measurement is particularly relevant to the use of waiting time data in the context of public hospital performance.

Although he was probably referring to emergency department waiting times rather than elective surgery waiting time, the then Chief Minister of the ACT commented "An undue focus on waiting times and waiting lists detracts attention from the quality of care, the clinical outcomes and the overall health and well-being, and the overall functioning of a system." [A072010].

Two articles in the collection deal with the emotional reaction of patients to having their highly personal experiences form part of the basis for statistical reporting. A152011 *I'm Like a Number: Gowrie Mum* and A402010 *Family Lost in the Statistics* both

express the opinion that treating people like numbers is somehow dehumanising and disrespectful. The latter also goes on to say that the numbers do not capture the effects of delays on people's lives and that the lack of acknowledgement of this is a serious omission from the formal reporting.

A paper from the Netherlands (Stoop, Vrangbæk et al. 2005) looks at the limitations of performance-oriented waiting time data as a tool for informing users of the health system. They conclude that the amount of system knowledge and interpretation required to understand waiting time data means that it is of little or no use to people waiting for surgery. What the patients and their carers want to know is when they will be treated. This shows in some of the individual stories where the published median waiting times were either dismissed as being untrue because they didn't reflect the patient's experience [A202011] or deemed irrelevant to the patient because they didn't give a clear idea of when the patient might have their surgery [A092010].

6.3 Analytical Model

In Section 6.1 I discussed the effects of language, statistical complexity, and context and usage on how easily a lay person could understand elective surgery waiting times as performance indicators for the Australian public hospital system. These factors can also be arranged as two themes:

- the conceptual distance between the definition of a performance indicator and how it is used; and
- the effect of the complexity of the relationship between the indicator and what is being measured.

The first of these centres around the clarity and appropriateness of the comparators used as well as the conclusions draw from the comparison. The second relates to how clear the relationship is between the measurement, the entity being measured and the value judgements made about performance.

Figure 6-1 below shows a potential model for predicting how the combination of indicator use and indicator complexity will affect the ease with which information is communicated using statistical performance indicators. It could also be used to analyse communication problems relating to this type of performance indicator.
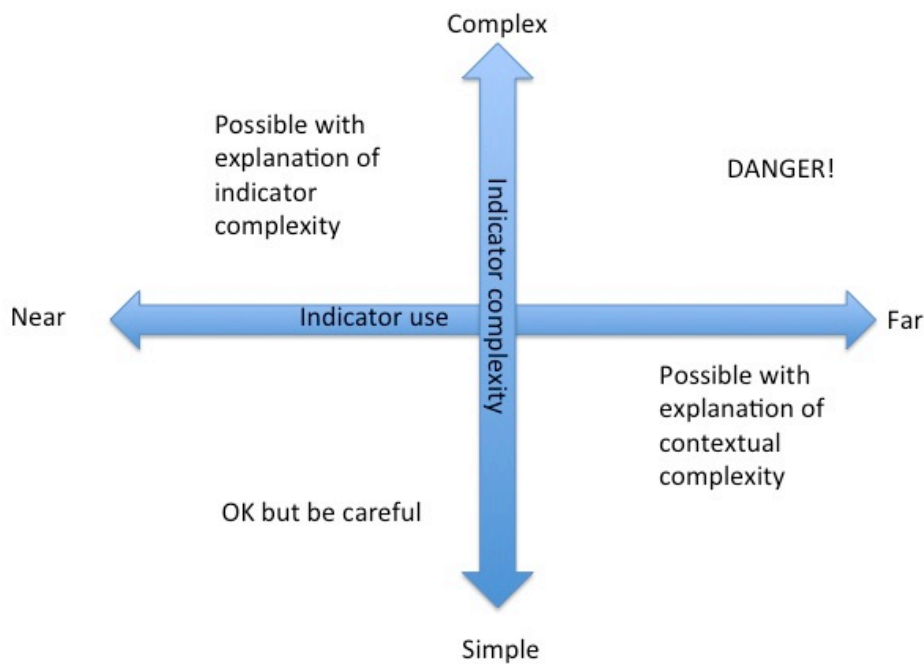
Figure 6-1: Interaction of indicator use and indicator complexity

Quadrant One in the upper left represents those instances of indicator use where the indicator is complex either because of the type of statistical construct or the complexity of the comparators but it is being used close to the way it was designed and like is clearly being compared with like. Indicators with these characteristics can be used to communicate clearly provided the indicator complexity is explained.

Quadrant Two in the upper right represents those instances of indicator use where the indicator is complex either because of the type of statistical construct or the complexity of the comparators and it is being used in way other than how it was designed with like not being clearly compared with like. Indicators with these characteristics should not be used.

Quadrant Three in the lower left represents those instances of indicator use where the indicator is simple both in statistical construct and comparators and it is being used close to the way is was designed with like clearly being compared with like. Indicators with these characteristics can be used to communicate clearly.

Quadrant Four in the lower right represents those instances of indicator use where the indicator is simple both in statistical construct and comparators but it is being

used in way other than how it was designed with like not being clearly compared with like. Indicators with these characteristics can be used to communicate clearly provided the use is explained.

In the code sets developed for describing each instance of performance indicator use in the collection of newspaper articles some, such as those relating to the type of statistic and its comparators, relate to how the performance indicator is being used. Others, such as those relating to communication and statistical assumptions relate to the complexity of the indicator itself.

6.3.1 Testing the model against the data

As a test of this model I developed two new coding questions relating to the axes of the model and applied them to the instances of performance indicator use in the articles from The Canberra Times. This final code set was based on two questions, one relating to the way the indicator is used (Q-1) and one relating to the complexity of the indicator (Q-2). The answers to the two new questions were formatted as binary pairs to prevent the results clustering at the intersection of the model's axes. The codes for each new question were assigned using the coding already applied to each instance of indicator use: type of statistic (Question PI-2); the number of comparators (Question PI-3); the validity assumptions (Question PI-5); and the communication assumptions (Question PI-6).

Question Q1: How closely to its defined purpose is the indicator being used?

This question addresses whether or not the indicator is being used in accordance with the constraints of its technical definition. The two answers are specified with detailed examples to maximise the consistency of how they are chosen. The main bases for this question are the codes relating to the type of statistic and the comparators.

- Near
    - The indicator is being used exactly as specified E.g. the waiting time for a particular surgery urgency and speciality is being used only in reference to that category, "90% of Category 1 Cardiac surgery patients were seen within the recommended waiting time". or
    - The indicator is being used broadly within its specifications E.g. aggregated elective surgery waiting times are being used as a measure

of the performance of an individual hospital or state-based hospital system.)

- o An individual patient's actual waiting time is compared to the guidelines for their urgency category.

- Far:
  - o The indicator is being used for something similar to its specified purpose but without enough information to tell if it is being used validly. E.g. the percentage of people waiting more than six months is given without stating the urgency category or mix of urgency categories. More than six months is too long for Urgency Categories 1 and 2 but well within the clinical guidelines for Urgency Category 3.
  - o The indicator is being used so far from its specified purpose that it is effectively useless. E.g. a statistical average is being used in relation to an individual's case.

Detailed examples for descriptive categories:

- Near because (implicit) comparison between states
- Near because (implicit) comparison between time periods
- Near because (implicit) comparison with guidelines
- Near because related to a specific target
- Far because comparator unclear or made up of several comparators
- Far because input rather than outcome or activity

Question Q-2: How complex is the relationship between the indicator and the organisation being measured?

This question addresses the complexity of the relationship between what is measured and the organisation being measured. The two answers are specified with detailed examples to maximise the consistency of how they are chosen. The main bases for this question are the codes relating to statistical and communication assumptions.

- Complex
  - o There is a complex chain of reasoning and assumption linking what is being measured with the aspect of organisational performance of interest. E.g. using state-level aggregates of elective surgery waiting times as a measure of overall hospital performance in that state.

- o There is missing information about comparators.
- o Aggregate averages across several categories are used without giving a breakdown of the underlying categories. E.g. State-level median waiting times for all urgency categories.
- o There are undocumented assumptions needed for statistical validity
- o Waiting times have vague comparators E.g. 'long wait', 'waiting too long'. Impossible to tell if it is longer than the clinical guideline or seems a long time to the patient.

- Simple:
  - o There is a straightforward relationship between what is being measured and the aspect of organisational performance of interest. E.g. using rates of avoidable infections as a measure of the effectiveness of infection control procedures.
  - o The comparison is of exactly the same indicator for two different time periods.

Detailed examples for descriptive categories:

- Complex because comparator so vague or assumed
- Complex because urgency category mix not given or assumed
- Complex because urgency category not given or assumed
- Simple because actual time compared to guideline
- Simple because same indicator, no assumptions, with comparator, between states
- Simple because time-based comparison of same indicator, no assumptions

As with earlier coding, I checked for:

- blank fields;
- combinations of codes and descriptions (e.g. all instances of indicator use have codes for questions Q-1 and Q-2; instances coded as 'simple' had descriptors relating to the choice of simple); and
- using the 'Gist of indicator use' descriptor to find instances of the same indicator used in the same way and checking that the coding for questions Q-1 and Q-2 were the same.

The results of coding the instances of indicator use against Questions Q-1 and Q-2 are shown in Table 6-5 below.

Table 6-5: Coding indicator use for complexity and closeness

	Total
Quadrant 1: Near-Complex	**34**
Quadrant 2: Far-Complex	**22**
Quadrant 3: Near-Simple	**58**
Quadrant 4: Far-Simple	**2**
Total	**116**

I used the codes relating to question PI-8 (match between value judgement and statistic) as a proxy for how easy it was, at least for the journalists writing the articles, to understand each instance of performance indicator use. Table 6-6 below shows the breakdown of codes for PI-8 against the four quadrants for complexity and closeness.

Table 6-6: Match between value judgement and statistic by complexity/closeness quadrant

	Total	Match between Value Judgement and Statistic		
		No	**Partly**	**Yes**
Quadrant 1: Near-Complex	34	0	23	11
Quadrant 2: Far-Complex	22	1	21	0
Quadrant 3: Near-Simple	58	0	7	51
Quadrant 4: Far-Simple	2	0	0	2
Total		**1**	**51**	**64**

Figure 6-2 below shows a visual breakdown of the codes for question PI-8 (How well does the value judgement match the statistical meaning) for each quadrant in the proposed model.
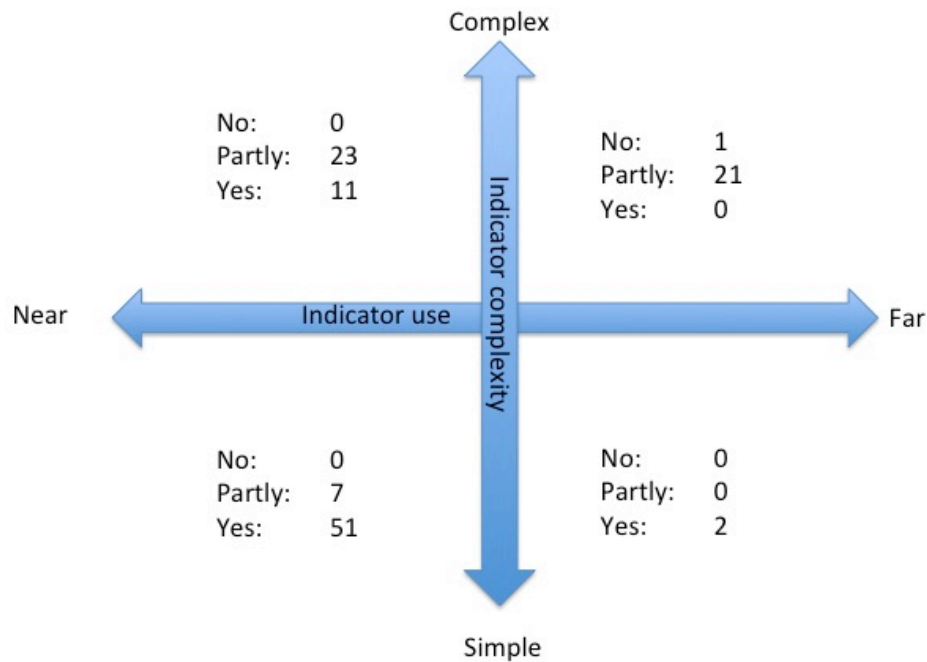
Figure 6-2: Matches between value judgement and statistic (No, Partly, Yes) for each quadrant

The coding for instances of indicator use in Figure 6-2 shows similar patterns to the model proposed in Figure 6-1. The match is best for simple indicators used in a way that is near to their definition (bottom left) and worst for complex indicators used in a way that is far from their definition (top right). As in the model, the top right quadrant is between the two extremes for match between statistic and value judgement. There are too few cases of a simple indicator being used in a way that is far from its definition to draw any solid conclusions.

The analytical model can be used in two ways: as a design tool to prevent problems with indicator use and communication; and as a post-hoc tool for understanding why there are problems with the communication of a performance measure. It has the capacity to improve the design and communication of performance measure in any arena where statistical constructs are used as part of the reporting.

6.4 Key findings

As described in Chapter 3.2, the funding and governance of the Australian public hospital system is described by a complex combination of the Australian constitution, legislation and agreements between the Federal government and the state and territory government. Statistics about elective surgery waiting times are used in several different ways within this context:

1. to communicate the performance of the public hospital system to the Australian public;
2. as performance indicators to allocate reward funding to the states and territories for improving the public hospital system; and
3. as a tool for managing and prioritising access to the public hospital system.

These uses involve communication with very different audiences which have different needs and starting points for understanding both the public hospital system itself and the elective surgery statistics used to describe its performance. The wider Australian public was an audience for points one and two while the audience for point three was focussed on professionals within the public health system and its governance.

Typically members of the wider public have little technical knowledge of how the hospital system works; little understanding of the intricacies of public health funding, and a limited understanding of statistical concepts. As detailed in Chapter 3.2 the funding and governance of the Australian public hospital system was a complicated system of agreements, which were constantly being renegotiated, between the Federal government and the State/Territory governments. A high level of detailed technical knowledge of the details of these agreements was necessary to understand the way elective surgery waiting times were used as performance indicators and to allocate reward funding.

The detailed thematic analysis of relevant newspaper articles combined with detailed analysis of each instance of elective surgery waiting time being used as a performance indicator for the public hospital system yielded insights into understanding the performance indicator and how it was used in the newspaper articles.

The three main factors affecting how elective surgery waiting times are understood as performance indicators are: ambiguities in the language used, limitations in the understanding of statistics, and the complexity of the context in which they are used.

These factors do not apply in isolation but interact to shape the clarity of communication about public hospital performance.

At the article level it was clear that as well as reports using waiting times as performance measures there were reports of discussion about what particular results meant, how valid they were and which particular measure was the most relevant in specific contexts (see Chapter 5.2.5 on page 140). The importance of the human side of storytelling showed strongly in the way personal stories were used to illustrate the effect health system performance had on individuals. The people in articles were given names and backgrounds, making them real enough to provoke Rule of Rescue type reactions in the reader.

Close analysis of each instance of using elective surgery waiting time for the match between any value judgement made about hospital system performance and the strict technical meaning of the statistic, its comparator and context showed clear patterns. When arranged into the quadrant model shown in Figure 6-1 on page 178 and Figure 6-2 183 it was possible to see that the interaction of statistical complexity and how closely the use matched the design of the performance indicator gave a clear insight into how readily the indicator could be used to form accurate value judgements.

This chapter has explored the results of my research in the light of literature relating to performance indicators, public health and communicating statistics. In it I have also described the development and testing of a tool to determine how appropriate a statistical measure is to use as a performance indicator in a particular context. The next chapter will summarise my research so far and make recommendations for further research.

7   Conclusions

This research has come full circle, from an initial enquiry based on the tension between complex data and the frequently-expressed need for simple answer, via a complex case study, to an analytical model that can be used to work out when a simple answer, or performance indicator, is and is not possible. The model has, of course its limitations but serves the purpose of breaking down the relationships between indicator, organisation and context into a manageable form. It provides at least some of the answer to my original research questions:

1. What makes performance indicators hard to understand? and
2. Is there anything we can do to make them easier to understand?

7.1   What this means for communicating performance using numerical indicators

Performance indicators are based on measurements that need to be both reliable/accurate (measuring the same thing in the same way gives the same result) and valid/unbiased (measuring they are intended to measure). To get from a measurement to a performance indicator we need evidence and an argument that relates what we want to know to the parameter we are measuring. If the audience for a performance indicator is to understand and accept what it means, the following conditions must be satisfied:

- the assumptions behind the indicator must be explained in a way that enables the audience to meaningfully integrate the new information with information it already has;
- the construction of any statistics in the indicator must be transparent enough to avoid setting up unhelpful heuristic shortcuts; and
- the rationale for using and valuing the indicator must be expressed in both logical terms and in terms of human relationships and values.

The analytical model described in Chapter 6.3 and shown in Figure 6-1 on page 178 can be used both as a design tool to test if the first two of these conditions are likely to be met and as an analytical tool for understanding why there are problems with the communication of a performance measure.

When it comes to measuring the performance of publically funded services such as the hospital system a key outcome is the meaningful communication of that performance to the public by whom the service is funded and for whom the service is delivered (Bird, Cox et al. 2005). Designing public sector performance indicators with their communicability in mind as well as their validity and accuracy means that there is a greater chance of their informing a wider audience in a way their value judgements about performance to align with the technical meaning behind the indicators.

7.2 Coding time tool

The thematic coding for both newspaper articles and instances of elective surgery waiting time was very time consuming and I realised that it would have been extremely useful to have some way of estimating how long it would take to develop and apply different types of codes. I divided my coding questions into three categories, those where I was looking for explicit words or phrases; those where I had to make simple decision about the code; and those where I was coding for abstract ideas and implicit content. For each type of code there were at least two slow passes through the material being coded while the codes were being developed and a further two, faster, passes through the material to test the robustness of the final coding definitions.

For a newspaper article of 300-400 words the number of passes and coding time for each type of question are shown in Table 7-1 below. The times are taken from Table 4-4 on page 115 and adjusted to account for my greater skill at developing and applying this type of coding.

Table 7-1: Coding time estimates

Question type	Number of long passes	Article coding time per long pass	Number of short passes	Article coding time per short pass
Explicit content	2	3 min	2	1 min
Decision-based content	3	10 min	2	5 min
Implicit content	5	20 min	2	5 min

This tool allows research using this style of thematic analysis to be clearly planned by giving a researcher the ability to work out the number and type of codes which can be used within their time and personnel constraints.

7.3 Limitations of this research

This research looked in detail at a small case study within a small public sector environment. This means that any generalisations about the usefulness of the analytical model have yet to be widely tested. It is also framed within the context of an English-speaking Western country with a high overall level of education and a readily available public health system. Differences in the emphasis of different education systems, health systems and language assumptions mean that any communication conclusions need to be viewed in the light of these limitations.

7.4 Further research

There are many possible extensions and refinements of the research described in this thesis. Two which relate to the analytical model for indicators are to:

- Repeat the research as closely as possible within a different public policy setting to see which findings are more broadly applicable and if the basic quadrant model is generalizable. In Australia a prime candidate would be the national literacy and numeracy testing system NAPLAN which is used to rate the performance of student cohorts according to the school they attend; and
- Look more closely within the quadrants of the model for more detailed and nuanced descriptions of how performance indicators can be applied and understood.

A variation on the second refinement would be to experiment with reworking the axes in terms of a question and its answer. Instead of "indicator use" and "indicator complexity" one axis would be the difference between the perceived and actual complexity of the question and the other the difference between the perceived and actual complexity of the answer. This difference is not necessarily static, as both the question and the form of the answer are likely to change during the process of analysing them.

A complementary piece of research would be to re-examine the newspaper articles with a closer focus on journalists' decisions about what to report and how. This could

be integrated with analysis of the specific communication methods used by hospitals and health departments and testing to see if the model developed in this research could be used to improve the clarity of those communications.

As long as simple numbers are used to report on the performance of complex systems there will be a need to understand how this form of communication is used and understood. There are many audiences for this type of information and many contexts within which it is used. The model developed as part of this research provides a tools for analysing how useful a numerical performance indicator is likely to be in a particular context.

References

ACT Health (2011). ACT Elective Surgery Access Policy 2011. ACT Directorate of Health. Canberra, ACT Government: 32.

ACT Health Services Directorate (2010). ACT Public Health Services Quarterly Performance Report December 2010. Canberra, ACT Department of Health.

ACT Health Services Directorate (2010). ACT Public Health Services Quarterly Performance Report March 2010. Canberra, ACT Department of Health.

ACT Health Services Directorate (2010). ACT Public Health Services Quarterly Performance Report September 2010. Canberra, ACT Department of Health.

ACT Health Services Directorate (2011). ACT Public Health Services Quarterly Performance Report June 2011. Canberra, ACT Department of Health.

ACT Health Services Directorate (2011). ACT Public Health Services Quarterly Performance Report March 2011. Canberra, ACT Department of Health.

ACT Health Services Directorate (2011). ACT Public Health Services Quarterly Performance Report September 2011. Canberra, ACT Department of Health.

Aristotle (350BCE). Rhetoric.

Australian Bureau of Statistics (2016) 6202.0 Labour Force Australia September 2016. Retrieved 30 October 2016, 2015, from http://www.abs.gov.au/ausstats/abs@.nsf/mf/6202.0

Australian Commission on Safety and Quality in Health Care (2015). National Consensus Statement: essential elements for safe and high-quality end-of-life-care. Sydney, ACSQHC.

Australian Curriculum Assessment and Reporting Authority (ACARA) (2010, 2010). "Australian National Curriculum - Mathematics." Retrieved 24 February 2015, 2015, from http://www.australiancurriculum.edu.au/mathematics/rationale.

Australian Federal Police (AFP) Annual Report 2014-15 Retrieved 27 October 2016, from https://www.afp.gov.au/afp-annual-report-2014-15#chapter_3-2

Australian Government (n.d.). "How government works." Retrieved 15 November 2015, from http://www.australia.gov.au/about-government/how-government-works/.

Australian Government Department of Health (2014). <u>Medicare Benefits Schedule Book</u>. Canberra, Commonwealth of Australia.

Australian Healthcare & Hospitals Association (2009). "Australia - Better data the key to hospital performance." <u>International Journal of HealthCare Quality Assurance</u> **22**(5).

Australian Institute of Health and Welfare (2010, August 2010). "National Health Data Dictionary." Version 15. Retrieved 13 July 2012, 2012, from http://www.aihw.gov.au/publication-detail/?id=6442468385.

Australian Institute of Health and Welfare and Royal Australasian College of Surgeons (2012). National Definitions for Elective Surgery Urgency Categories. Can berra, Australia.

Australian Institute of Health and Welfare (2013) , Health expenditure Australia 2011-1, 2 September 2013

Australian Medical Association (2011). Public Hospital Report Card 2010, Australian Medical Association.

Australian Medical Association (2011). Public Hospital Report Card 2011, Australian Medical Association.

Australian Taxation Office (2012). "GST." Retrieved 16 November 2015, 2015, from https://http://www.ato.gov.au/Business/GST/.

Bartlett, C., J. Stern, et al. (2002). "What is newsworthy? Longitudinal study of the reporting of medical research in British newspapers." <u>British Medical Journal</u> **325**: 4.

Bird, S. M., D. Cox, et al. (2005). "Performance indicators: good, bad, and ugly." <u>Journal of the Royal Statistical Society A</u> **168**(1): 27.

Boyatzis, R. E. (1998). <u>Transforming Qualitative Data: Thematic Analysis and Code Development</u>. THousand Oaks, California, Sage publications.

Brimelow, A. (2014, 6 November 2014). "UK doctors have 'ethical duty' to prevent waste." Retrieved 28 January 2015, 2015, from http://www.bbc.com/news/health-29920025.

Broom, A., J. Broom, et al. (2014). "Cultures of resistance? A Bourdieusian analysis of doctors' antibiotic prescribing." <u>Social Science & Medicine</u> **110**: 8.

Buchwald, J. and M. Feingold (2013). <u>Newton and the Origin of Civilisation</u>. Princeton and Oxford, Princeton University Press.

Cameron, C. D. and B. K. Payne (2011). "Escaping Affect: How Motivated Emotional Regulation Creates insensitivity to Mass Suffering." <u>Journal of Personality and Social Psychology</u> **100**(1): 1-15.

Chen, P., R. Gibson, et al. (2007). Australian governments and online communication. <u>Government Communication in Australia</u>. S. Young. Melbourne, Cambridge University Press.

COAG (2010). Council of Australian Governments Meeting, 19 and 20 April 2010: Communiqué. Canberra, COAG.

COAG (2011). Attachment A to Communiqué, 13 February 2011 - Heads of Agreement - National Health Reform. COAG. Canberra, COAG.

COAG (2011). Intergovernmental Agreement on Federal Financial Relations. COAG. Canberra, COAG.

COAG Reform Council (2009). National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan. Canberra.

COAG Reform Council (2011). National Health Reform Agreement - National Partnership Agreement on Improving Public Hospital Services. Canberra

COAG Reform Council (2011). National Healthcare Agreement 2011.

COAG Reform Council (2011). National Healthcare Agreement: Performance Report for 2009-10. Sydney, COAG Reform Council.

COAG Reform Council (2011). National Partnership Agreement on the Elective Surgery Waiting List Reduction Plan: Period 3 assessment report, Sydney, COAG Reform Council.

Commonwealth Grants Commission (2015). "Commonwealth Grants Commission." Retrieved 16 November 2015, 2015, from https://http://www.cgc.gov.au/.

Commonwealth of Australia Bureau of Meteorology (2015). "Weather Words." Retrieved 27 February 2015, 2015, from http://www.bom.gov.au/info/wwords/.

Conley, D. (2002). <u>The Daily Miracle</u>. Oxford, New York, Oxford University Press.

Cosmides, L., H. C. Barrett, et al. (2010). "Adaptive specializations, social exchange, and the evolution of human intelligence." Proceedings of the National Academy of Sciences **107**(suppl. 2): 9007-9014.

Cox, N. (1988). Public Records. Contemporary History: Practice and Method. A. Seldon. Oxford, Basil Blackwell.

Cronin, D. (2010). Hospital website gives wrong diagnosis. The Canberra Times. Canberra, Fairfax Media.

Curtis, A. J., C. O. H. Russell, et al. (2010). "Waiting lists and elective surgery: ordering the queue." Medical Journal of Australia **192**(4): 217-220.

Deming, W. E. (1954). On the Teaching of Statistical Principles and Techniques to People in Industry. Session of the International Statistical Institute. Rome.

Deming, W. E. (1982 (reprinted 2000)). Out of the crisis. Cambridge, MA, MIT Press.

Department of Human Services (2015, 16 September 2015). "Australian Government." Retrieved 17 November 2015, 2015, from http://www.humanservices.gov.au/customer/subjects/medicare-services.

Department of the Parliamentary Library (2002, 20 August 2002). "Research Note No. 1: Horizontal Fiscal Equalisation." Retrieved 16 November 2015, 2015, from http://www.aph.gov.au/binaries/library/pubs/rn/2002-03/03rn01.pdf.

DiPrete, B. and D. Coffman (2007). A Brief History of Health Services Prioritization in Oregon. Salem, Oregon, Oregon Health Services Commission: 7.

Dunbar, R. (1996). Grooming, Gossip and the Evolution of Language. London, Faber and Faber.

Dunbar, R. (1998). "The Social Brain Hypothesis." Evolutionary Anthropology **6**(5): 13.

Fairfax Media (2013). "Fairfax Regional Media." Retrieved 11 OCtober 2013, from http://www.fairfaxregionalmedia.com.au/.

Gigerenzer, G. and A. Edwards (2003). "Simple tools for understanding risks: from innumeracy to insight." BMJ **327**: 4.

Gigerenzer, G., W. Gaissmaier, et al. (2008). "Helping Doctors and Patients Make Sense of Health Statistics." Psychological Science in the Public Interest **8**(2): 43.

Gilbert J. K. and Stocklmayer S. <u>A Reader in Science Communication</u>. New York, Taylor & Francis**:** 320.

Grigoroudis, E., E. Orfanoudaki, et al. (2011). "Strategic Performance Measurement in a healthcare organisation: A multiple criteria approach based on balanced scorecard." <u>Omega</u> **40**: 15.

Hadorn, D. (1991). "The Oregon Priority-Setting Exercise: Quality of Life and Public Policy." <u>The Hastings Center Report</u> **21**(3): 6.

Hadorn, D. (1996). "The Oregon Priority-setting Exercise: Cost-effectiveness and the Rule of Rescue Revisited." <u>Medical Decision Making</u> **16**: 3.

Hardin, G. (1968). "The Tragedy of the Commons." <u>Science</u> **162**(3859): 6.

Harrower, T. (2007). <u>Inside Reporting</u>. New York, McGraw-Hill.

Head, B. (2007). The Public Service and government communication: Pressures and Dilemmas. <u>Government Communication in Australia</u>. S. Young. Melbourne, Cambridge University Press.

International Standards Organisation (2004). ISO/IEC 11179 Information technology -- Metadata registries (MDR) I. E. Commission. Geneva, Switzerland.

Jecker, N. S. (2013). "The Problem with Rescue Medicine." <u>Journal of Medicine and Philosophy</u> **38**: 18.

Jordanova, L. (2006). <u>History in Practice</u>. London, Hodder Education.

Juster, N. (1961). <u>The Phantom Tollbooth</u>. London, HarperCollins Publishers Ltd.

Kahan, D. M., E. Peters, et al. (2012). "The polarising impact of science literacy and numeracy on perceived climate change risks." <u>Nature Climate Change</u> (Advance online publication).

Kahnemann, D. (2011). <u>Thinking, Fast and Slow</u>. New York, Penguin Books.

Kahnemann, D. and A. Tversky (1979). "Prospect theory: an analysis of decision under risk." <u>Econometrica</u> **42**(2): 263-291.

Kahnemann, D. and A. Tversky (1984). "Choices, values and frames." <u>American Psychologist</u> **34**.

Kaplan, R. M. (1992). A Quality-of-Life Approach to Health Resource Allocation. Rationing America's Medical Care: The Oregon Plan and Beyond. M. A. Strosberg, J. M. Wiener, R. Baker and I. A. Fein. Washington D.C., The Brookings Institution.

Katelaris, A. (2011). "Time to rethink end-of-life care." Medical Journal of Australia **194**(11).

Kazandjian, V. A., N. Matthes, et al. (2003). "Are performance Indicators generic? The international experience of the Quality Indicator Project®." Journal of Evaluation in Clinical Practice **9**(2): 11.

Kelvin, W. T., Baron (1883). Popular Lectures: Electrical Units of Measurement.

Klassen, S. (2010). "The Relation of Story Structure to a Model of Conceptual Change in Science Learning." Science and Education (19): 305-317.

Leathey, C. M., R. Gilbert, et al. (2008). "Measuring Hospital Performance -- 2008 forum summary." Medical Journal of Australia **193**(8): S95-S96.

Lemieux-Charles, L., W. McGuire, et al. (2003). "The use of multilevel performance indicators in managing performance in health care organizations." Management Decision **41**(8): 10.

Lichtenstein, S. and P. Slovic (1971). "Reversals of preference between bids and choices in gambling decisions." Journal of Experimental Psychology **89**(1): 9.

Lowenstein, G. and D. A. Small (2007). "The Scarecrow and the Tin Man: The Vicissitudes of Human Sympathy and Caring." Review of General Psychology **11**(2): 112-126.

Manchikanti, L. and J. A. Hirsch (2012). "Regulatory Burdens of the Affordable Care Act." Harvard Health Policy Review **13**(2): 4.

Massola, J. (2010). PM Plans overhaul of health system. The Canberra Times. Canberra, Fairfax Media.

Maughan, D. and J. Ansell (2014). Protecting Resources, promoting value: a doctor's guide to cutting waste in clinical care. London, UK, Academy of Medical Royal Colleges.

Mcintyre, D. and G. Mooney (2007). Why this book? The Economics of Health Equity. D. Mcintyre and G. Mooney. Cambridge, Cambridge University Press.

References

McKie, J. and J. Richardson (2003). "The Rule of Rescue." <u>Social Science & Medicine</u> **56**: 13.

McRae, M. (2011). <u>Tribal Science: Brains, Beliefs and Bad Ideas</u>. Brisbane, University of Queensland Press.

Meadows, D. (1998). Indicators and Information Systems for Sustainable Development. Hartland, The Sustainability Institute**:** 78.

Media Spy (2013). "Newspaper Circulation, March 2011." from http://www.mediaspy.org/report/wp-content/uploads/2011/05/March-2011-circulation.pdf.

Mirelman, A., E. Mentzakis, et al. (2012). "Decision-Making Criteria among National Policymakers in Five Countries: A Discrete Choice Experiment Eliciting Relative Preferences for Equity and Efficiency." <u>Value in Health</u> **15**(3): 534-539.

National Health Performance Authority (2016). "MyHospitals." Retrieved 8 February 2016, 2016, from http://www.myhospitals.gov.au/.

National Health Service (2015, 7 January 2015). "The NHS in England." Retrieved 20 January 2015, 2015, from http://www.nhs.uk/NHSEngland/thenhs/about/Pages/overview.aspx.

Newton, I. (1687) <u>Philosophiæ Naturalis Pricipia Mathematica,</u> S. Peyps, Royal Society Press, London

Nocera, A. (2010). "Performance-based hospital funding: a reform tool or an incentive for fraud?" <u>Medical Journal of Australia</u> **192**(4).

Nord, E., J. Richardson, et al. (1995). "Maximising Health Benefits vs Egalitarianism: An Australian Survey of Health Issues." <u>Social Science & Medicine</u> **41**(10): 1429-1437.

Nord, E., J. Richardson, et al. (1995). "Who cares about cost? Does economic analysis impose or reflect social values?" <u>Health Policy</u> **34**: 15.

OECD (2015). OECD Reviews of Health Care Quality: Australia 2015: Raising Standards. Paris.

Oxford University Press (2012). "Oxford English Dictionary." Online edition. Retrieved 19 July 2012, from http://www.oed.com/.

Parliament of New South Wales (n.d.). "The Roles and Responsibility of Federal, State and Local Governments." Retrieved 12 November 2015, 2015, from http://www.parliament.nsw.gov.au/prod/web/common.nsf/key/TheRolesandResponsibilitiesofFederalStateandLocalGovernments.

Peters, E., D. Västfjäll, et al. (2006). "Numeracy and Decision Making." Psychological Science **17**(5): 407-413.

Peters, E., J. Hibbard, et al. (2010). Numeracy Skill and the Communication, Comprehension and Use of Risk-Benefit Information. The Feeling of Risk: New Perspectives on Risk Perception. P. Slovic. Abingdon, New York, Earthscan/Routledge.

Pfeffer, J. and Sutton, R.I. (2000). The Knowing-Doing Gap: How Smart Companies Turn Knowledge Into Action, Boston, MASS USA, Harvard Business School Press.

Porter, T. M. (1995). Trust in Numbers: The Pursuit of Objectivity in Science and Public Life. Princeton, NJ USA, Princeton University Press.

Pratchett, T., I. Stewart, et al. (1999). The Science of Discworld. London, Random House.

Reyna, V. F. (2008). "A Theory of Medical Decision Making and Health: fuzzy Tract Theory." Medical Decision Making.

Reyna, V. F. and C. J. Brainero (1995). "Fuzzy-trace theory: an interim synthesis." Learning and Individual Differences **7**(1): 75.

Romzek, B. S. (2000). "Dynamics of Public Sector Accountability in an Era of Reform." International Review of Administrative Sciences **66**(21): 21-44.

Rowland, W. (2009). "Corporate Social Responsibility and Garrett Hardin's 'Tragedy of the Commons' as Myth and Reality." The Journal of Corporate Citizenship **35**.

Royal Statistical Society (2016). "Statistical Excellence in Journalism." Retrieved 17 February 2016, 2016, from http://www.rss.org.uk/RSS/About/Recognising_Statistical_Excellence/Journalism_awards/RSS/About_the_RSS/Recognising_statistical_excellence_sub/Statistical_Excellence_in_Journalism_Awards.aspx

Scutchfield, E. D., M. W. Bhandari, et al. (2009). "Public Health Performance." American Journal of Preventative Medicine **36**(3): 7.

Seldon, A., Ed. (1988). Contemporary History: Practice and Method. Oxford, Basil Blackwell.

Shomos, A. and M. Forbes (2014). Literacy and Numeracy Skills and Labour Market Outcomes. Productivity Commission Staff Working Paper. Canberra.

Simpson, I., K. Beninger, et al. (2015). Public Confidence in Official statistics. London, NatCen Social Research**:** 44.

Slovic, P. (2007). " 'If I look at the mass I will never act' : Psychic numbing and genocide." Judgement and Decision Making **2**(2): 79-95.

Stewart, J. (2009). Public Policy Values. London, Palgrave Macmillan.

Stewart, T.A. (1997). *Intellectual Capital: the new wealth of organisations*. New York, NY USA, Doubleday/Currency.

Stewart, T.A. (2001). The Wealth of Knowledge: Intellectual Capital and the Twenty-first Century Organization. New York, NY USA, Currency.

Stocklmayer, S. (2013). Engagement with Science: Models of Science Communication. Communication and Engagement with Science and Technology: Issues and Dilemmas

Stoop, A. P., K. Vrangbæk, et al. (2005). "Theory and practice of waiting time data as a performance indicator in health care: A case study from the Netherlands." Health Policy **73**: 10.

Tufte, E. R. (2001). The Visual Display of Quantitative Information. Cheshire, Connecticut, Graphics Press LLC.

Tversky, A. and D. Kahnemann (1974). "Judgement Under Uncertainty: Heuristics and Biases." Science **185**.

Ubel, P. A., G. Lowenstein, et al. (1996). "Individual Utilities Are Inconsistent with Rationing Choices: A partial explanation of Why Oregon's Cost-Effectiveness List Failed." Medical Decision Making (16): 9.

Van Peursem, K. A., M. J. Pratt, et al. (1995). "Health management performance: A review of measures and indicators." Accounting, Auditing & Accountability Journal **8**(5): 36.

Weinex, B. J., J. A. Alexander, et al. (2006). "Quality Improvement Implementation and Hospital Performance on Quality Indicators." HSR: Health Services Research **41**(2): 27.

Yager, R. E. (1991). "The constructivist learning model: towards real reform in science education." The Science Teacher **58**(6): 6.

Young, S. (2007). Introduction: The theory and practice of government communication. Government Communication in Australia. S. Young. Melbourne, Cambridge University Press.

Young, S., Ed. (2007). Government Communication in Australia. Melbourne, Cambridge University Press

Appendix A – Listing of Articles from *The Canberra Times*

Article ID	Date	Title
A012010	25-Jan-10	PM plans overhaul of health system
A022010	15-Feb-10	Abbott to hand over hospitals
A032010	19-Feb-10	Theatres up and running at last
A042010	03-Mar-10	Rudd unveils national health plan
A052010	04-Mar-10	Rudd's $90b plan to overhaul hospitals
A062010	05-Mar-10	Tax rises flagged on health reforms
A072010	06-Mar-10	Reforms give $240m to territory
A082010	15-Mar-10	Premiers hit Rudd for more health cash
A092010	15-Mar-10	Waiting for day he can breathe easy
A102010	15-Mar-10	Rudd's $632m medico-training plan
A112010	16-Mar-10	Rudd's regional hospital diagnosis
A122010	18-Mar-10	Stanhope warms to Rudd on health
A132010	19-Mar-10	Election battle opens with health debate
A142010	27-Mar-10	Mixed feelings after health funding talks
A152010	01-Apr-10	Swine flu blamed for waiting lists
A162010	13-Apr-10	PM puts up $3b to win over states
A172010	15-Apr-10	NSW holds out on $30m owed
A182010	18-Apr-10	Calvary closes ward for holidays
A192010	21-Apr-10	Rudd pays for health victory
A202010	22-Apr-10	Better ACT health 'by year-end'
A212010	03-May-10	$18m for elective surgery, cancer services

Article ID	Date	Title
A222010	18-May-10	ACT to hand over 47pc of GST for hospitals deal
A232010	10-Jun-10	Mayor of Dickson's long wait
A242010	11-Jun-10	ACT Government moves to reduce surgery backlog
A252010	11-Jun-10	HELP ARRIVES FOR MAYOR OF DICKSON
A262010	12-Jun-10	Claim ACT Health doctoring waiting lists
A272010	15-Jun-10	Call to fund new path for patients
A282010	17-Jun-10	Battle to reduce our hospitals' wait lists
A292010	19-Jun-10	Flying doctors keeping hospitals afloat
A302010	24-Jun-10	Govt blasted over surgery downgrades
A312010	29-Jun-10	Some patients 'inappropriately' listed for urgent elective surgery
A322010	01-Jul-10	Mayor gets his day in surgery, at last
A332010	02-Jul-10	Better bill of health for ACT hospitals
A342010	08-Jul-10	Timetable set for Govt health overhaul
A352010	22-Jul-10	Sick spell to stretch emergency wait times
A372010	26-Jul-10	Gillard stumbles over pre-allocated health pledge
A382010	05-Aug-10	Abbott's $3.1b for beds in hospitals
A402010	06-Aug-10	Family feels lost among the stats
A432010	14-Aug-10	No deal: Calvary buy-out cancelled
A442010	16-Aug-10	Gillard announces $400m in health promises
A452010	20-Aug-10	Govt rethinks delivery of hospital services
A462010	11-Sep-10	Adjudicator called in for health dispute
A472010	12-Sep-10	Doctors 'quitting over shortages'
A482010	23-Sep-10	Yass man saved, but fears for others

Article ID	Date	Title
A492010	24-Sep-10	ACT Libs, Greens vote for Calvary bid inquiry
A502010	29-Sep-10	ACT patients play a waiting game
A512010	02-Oct-10	Surgical centre takes on hospitals
A522010	14-Oct-10	Dedicated emergency surgical unit open
A532010	18-Oct-10	Public patients to get private surgery
A542010	19-Oct-10	Northside hospital decision in sight
A552010	04-Nov-10	Patients to have input on health under plan
A572010	18-Nov-10	11 new hospital beds not enough: AMA
A582010	19-Nov-10	Shock to capital's hospital system
A592010	22-Nov-10	Hospital wait deaths on par with road fatalities
A602010	23-Nov-10	No improvement in hospital 'blockages'
A612010	11-Dec-10	Hospital website gives wrong diagnosis: AMA
A622010	23-Dec-10	Doctors call for Gallagher's dismissal
A632010	27-Dec-10	Holiday delays minor surgery
A642010	31-Dec-10	ACT sends NSW $10m hospital bill
A012011	02-Jan-11	Hospital plan a farce: Podger
A022011	17-Jan-11	Auditor-General slams waiting lists
A032011	18-Jan-11	Auditor finds surgery patients downgraded without explanation
A042011	18-Jan-11	Relief at last: 12-month wait for prostate surgery
A052011	10-Feb-11	Canberra in Queanbeyan Hospital deal
A062011	11-Feb-11	Territory set to lose out in health changes
A072011	12-Feb-11	Gillard rolls for new deal on health
A082011	14-Feb-11	ACT a winner in health agreement

Article ID	Date	Title
A092011	15-Feb-11	ACT $10m better off under new agreement
A102011	15-Feb-11	Leaders need to 'get off and racing' to improve care
A112011	17-Mar-11	Elective surgery up 14pc but wait times stay high
A122011	18-Mar-11	Casualty increases may result in upgrades
A132011	30-Apr-11	New data to reveal health burden
A142011	04-May-11	Extra funds for more operations at hospitals
A152011	06-Jun-11	I'm like a number: Gowrie Mum
A162011	08-Jun-11	New hospital watchdog gets green light
A172011	19-Jun-11	ACT's improving prognosis
A182011	22-Jun-11	Gallagher announces 'challenging' health goals
A192011	28-Jun-11	Health reforms on track: Gallagher
A202011	03-Jul-11	'My 17-month surgery delay'
A212011	07-Jul-11	Government sets new targets for elective surgery
A222011	03-Aug-11	One year to prepare for new surgery deadlines
A232011	04-Sep-11	Repeated delays add to pain of waiting for surgery
A242011	06-Sep-11	Long hospital waits cost ACT $900k
A252011	07-Sep-11	ACT hospitals left waiting for federal reward funding
A262011	19-Sep-11	Extra surgery slashes elective queues
A272011	21-Sep-11	Dickson's 'mayor' remembered
A282011	04-Nov-11	Hospitals not up to standard, AMA says
A292011	05-Nov-11	Experience leaves patient sick of hospital system
A302011	30-Nov-11	Mixed results for hospital wait times
A312011	02-Dec-11	Govt denies cancellation of surgery
A322011	07-Dec-11	ACT, Calvary hospital deal

Article ID	Date	Title
A332011	08-Dec-11	ACT, Calvary all smiles on $130m hospital deal
A342011	12-Dec-11	ACT medical firm wins award with healthy growth
A352011	20-Dec-11	Patients going public
A362011	22-Dec-11	Medical services wind down over Christmas
A372011	29-Dec-11	NSW owes ACT millions

Article ID   **Date**   **Title**

Appendix B – Listing of Instances of waiting time as a performance indicator

Instance ID	Article ID	Gist of indicator use
I01032010	A032010	Current average waiting times for urgent surgery compared to longest guideline
I02032010	A032010	Current average waiting time for 'less urgent' surgery unfavourably compared to longest guideline
I01092010	A092010	Individual waiting time compared unfavourably to guidelines for urgency category, 'should have been seen'
I01152010	A152010	Comparison between time periods of numbers of people on waiting list - up
I02152010	A152010	Comparison between time periods of numbers of people receiving surgery
I01162010	A162010	Reference to 'waiting too long' but times/numbers not given
I01182010	A182010	ACT has the longest elective surgery waiting list in Australia
I02182010	A182010	Percentage of people on ACT lists waiting over 365 days
I01192010	A192010	Vague reference to spending less time waiting in emergency and for surgery
I01212010	A212010	ACT has the longest elective surgery waiting list in Australia
I02212010	A212010	Percentage of people on ACT lists waiting over 365 days
I01232010	A232010	Actual waiting time compared to what sounds like an informal timeframe given by a doctor.

Instance ID	Article ID	Gist of indicator use
I02232010	A232010	Waiting time and urgency category for kidney stent
I03232010	A232010	Waiting time and Dr's comment about how soon diagnostic surgery needed for possible prostate cancer.
I01242010	A242010	ACT has the longest median elective surgery waiting times in Australia
I02242010	A242010	Number of people waiting greater than 12 months
I01262010	A262010	January Category 1 (30 days) no surgery by June.
I02262010	A262010	January Category 2a (60 days) no surgery by June.
I03262010	A262010	95% of Category 1 had surgery on time
I01272010	A272010	Waited over a year when told needed surgery within two weeks.
I01282010	A282010	The hospitals admitted 10,104 elective surgery patients, up from 9577 in 2007-08.
I02282010	A282010	Comparison of ACT median waiting time with other states - longest
I03282010	A282010	Queensland median waiting time
I04282010	A282010	Percentage of people on ACT lists waiting over 365 days, no mention of guidelines, comment that this is one of the longest in the country.
I05282010	A282010	Percentage in ACT waiting over 365 days compared to percentage in Tasmania
I06282010	A282010	Comparison between time periods of National % waiting over 365 days - drop
I07282010	A282010	Comparison between time periods of numbers of operations nationally - improvement
I01302010	A302010	Claim' 95% of urgent patients operated on within

Instance ID	Article ID	Gist of indicator use
		30 days
I02302010	A302010	ACT has the longest elective surgery waiting list in Australia
I01322010	A322010	Waited since last May for kidney stent and biopsy
I02322010	A322010	Waited over a year when told needed surgery within two weeks.
I03322010	A322010	Category 1 but implication that actual waiting time longer than 30 days, downgrade to category 2
I01332010	A332010	Comparison between time periods of waiting times for urgent - dropped
I02332010	A332010	Comparison between time periods of waiting times for less urgent - extended delays
I03332010	A332010	Comparison between time periods of aggregated median waiting time - down
I04332010	A332010	Comparison between time periods of waiting times for urgent - down
I05332010	A332010	Comparison between time periods of waiting times for semi-urgent - up
I06332010	A332010	Comparison between time periods of waiting times for non-urgent- rose
I01502010	A502010	ACT has the longest elective surgery waiting list in Australia
I01532010	A532010	Connection between urgency category, waiting times 'longer than desirable' and priority of non-urgent surgery
I01622010	A622010	Percentage of 'general surgery' patients waiting more that 1 year

Instance ID	Article ID	Gist of indicator use
I02622010	A622010	% of urology patients waiting more than 1 year - one of the worst results in Australia
I01632010	A632010	ACT has the longest elective surgery waiting times in Australia
I01022011	A022011	Waiting for more than a year for urgent elective surgery
I01042011	A042011	Waiting for more than a year for urgent elective surgery
I01052011	A052011	Escalating waiting lists for non-urgent surgery
I01102011	A102011	Elective surgery provided on time to 95% of patients
I01112011	A112011	Comparison between time periods of number of operation in the ACT-up
I02112011	A112011	Comparison between time periods of number of patients waiting ;longer than 1 year - up
I03112011	A112011	Percentage of Category 1 surgery done on time
I01122011	A122011	Comparison between time periods of number of people waiting > 1 year - rose
I02122011	A122011	Comparison between time periods of number of procedures (% change) - increase
I03122011	A122011	Comparison between time periods of number of procedures (raw numbers) - increase
I04122011	A122011	Comparison between time periods of median waiting times for category 1 - deteriorated
I05122011	A122011	Comparison between time periods of median waiting times for category 2 - deteriorated
I06122011	A122011	Comparison between time periods of median

Instance ID	Article ID	Gist of indicator use
		waiting times for category 3 - deteriorated
I01132011	A132011	Comparison of ACT and NSW average waiting time for elective surgery - ACT longer than ACT, NSW second-longest
I01142011	A142011	Comparison between time periods of number of surgeries - record numbers
I02142011	A142011	Vague reference to 'experiencing long waits'
I01152011	A152011	Category 2 patient, waiting time to date and forecast waiting time.
I02152011	A152011	Comparison between time periods of people waiting longer than guidelines - improved
I03152011	A152011	Comparison between time periods of median waiting time across all categories - increase
I04152011	A152011	Comparison between time periods of % of category 1 seen on time - dropped
I01172011	A172011	Comparison between time periods of number of people on the waiting list - decrease
I02172011	A172011	Comparison between time periods of number of people waiting > 1 year - lowest
I03172011	A172011	ACT has worst elective waiting time in Australia
I01182011	A182011	Comparison between time periods of number of people on the waiting list - down
I02182011	A182011	Comparison between time periods of number of people waiting > 1 year - lowest
I01192011	A192011	Elective surgery waiting lists at lowest levels since 2004
I02192011	A192011	Number of people waiting too long for surgery was

Instance ID	Article ID	Gist of indicator use
		at its lowest point since 2003
I03192011	A192011	Comparison between time periods of number of people on the waiting list - down
I04192011	A192011	ACT has the longest elective surgery waiting times in Australia
I01202011	A202011	Category 2 patient waited 90 days, forecast to wait for a further 420 days
I02202011	A202011	Comparison between time periods of number of procedures - extra
I03202011	A202011	Contrast between personal experience and official statistics
I01212011	A212011	Change in target % patients on time
I01222011	A222011	% of category 1 patients seen on time
I02222011	A222011	% of category 2 patients seen on time
I03222011	A222011	% of category 3 patients seen on time
I04222011	A222011	Comparison of % ACT category 2 seen on time with other states - lowest
I05222011	A222011	Comparison of % ACT category 2 seen on time with other states - lowest
I06222011	A222011	Comparison of % ACT category 1 seen on time with other states - third highest
I07222011	A222011	Comparison of % ACT category 1 seen on time with other states - third highest
I01232011	A232011	Comparison between time periods of aggregated median waiting time - blown out
I02232011	A232011	Comparison between time periods of number of people on the waiting list - down

Instance ID	Article ID	Gist of indicator use
I01242011	A242011	ACT has the longest elective surgery waiting times in Australia
I02242011	A242011	Change over time of people waiting longer than recommended - good
I03242011	A242011	Seen on time most urgent category - met target
I042422011	A242011	Seen on time less urgent categories - did not meet targets
I01252011	A252011	ACT has the longest elective surgery waiting times in Australia
I02252011	A252011	Change over time of people waiting longer than recommended - reducing
I03252011	A252011	Seen on time most urgent category - met target
I04252011	A252011	Seen on time less urgent categories - did not meet targets
I05252011	A252011	Median time to be removed from waiting list - high
I01262011	A262011	Elective surgery waiting lists at their shortest since 2003 - slashed
I02262011	A262011	Comparison between time periods of numbers of people on waiting list - fell
I03262011	A262011	Change over time in number of people waiting longer than recommended - fell/cut
I04262011	A262011	% of patients waiting longer than clinically recommended - still waiting longer
I05253011	A262011	Elective surgery waiting lists at their shortest since 2003.
I06262011	A262011	Change over time of people waiting longer than recommended - halved

Instance ID	Article ID	Gist of indicator use
I07262011	A262011	Comparison between time periods of aggregated median waiting time - up
I08262011	A262011	% of category 1 patients seen on time - met target
I09262011	A262011	% of category 2 patients seen on time - only 50%
I10262011	A262011	% of category 3 patients seen on time - fewer than 80%
I01272011	A272011	Waited over a year when told needed surgery within two weeks.
I01282011	A282011	Failure to meet elective surgery targets
I02282011	A282011	Record levels of elective surgery being performed
I01302011	A302011	ACT has worst elective waiting time in Australia
I01352011	A352011	Fewer public patients facing long waits
I02352011	A352011	Comparison between time periods of number of people added to the waiting lists
I03352011	A352011	Comparison between time periods of number of people on the waiting list - fell
I04352011	A352011	Comparison between time periods of number of people waiting longer than clinically recommended - came down
I05352011	A352011	% of category 1 patients seen on time compared to target - more
I06352011	A352011	% of category 2 patients seen on time compared to target - fewer than half
I07352011	A352011	% of category 3 patients seen on time compared to target - just under target
I08352011	A352011	Median waiting time for elective surgery in days

Appendix C - Final coding criteria for thematic analyses

Articles from The Canberra Times

Question A-1: Is the article explicitly about elective surgery in the public system or about the (public) hospital system more generally?

Purpose of question
To identify all articles in the collection that refer explicitly to elective surgery.

Possible answers
Yes
The article refers to elective surgery funding; elective surgery waiting times; elective surgery administration; elective surgery facilities; as well as stories relating to individuals and their elective surgery experiences. Articles relating to procedures performed privately (i.e. privately funded) are not included. Articles referring to treating public patients in private facilities are included.

No
The article is about the broader political, administrative and funding context surrounding elective surgery in the public health system. There is no direct mention of elective surgery.

Question A-2: Is the article about an individual's experiences or about elective surgery in general?

Purpose of question
To determine the mix of articles relating to individuals and the articles about elective surgery and public hospitals in general.

Possible answers
Individual
The article is centred on an individual's story and any mention of the hospital system is directly related to them and their experiences. Any generalisations about the hospital system in an article about an individual trigger the answer "Both".

System

The article does not mention any individual or their experiences relating to elective surgery.

Both

The article deals with both elective surgery in general and an individual's experiences. This category includes stories using an individual's experience to illustrate a more general story.

Question A-2a: Are stories about individuals positive or negative?

Purpose of question

To find out if there was a dominant way of framing personal stories with respect to the public hospital system. This question only applies to articles where the answer to Question 2 is either "Individual" or "Both".

Possible answers

Positive

The article is about a positive outcome for a patient and is positive about elective surgery in the public hospital system.

Negative

The article is either about a negative outcome for the patient and/or is negative about elective surgery in the public hospital system.

Both

The article meets the criteria for both 'positive' and 'negative'.

N/A

The answer to Question 2 was "System"

Question A-3: Is the article about the Australian hospital system or about the ACT hospital system?

Purpose of question

To determine the which parts of the health system were covered by the article.

Much performance measurement relies on comparisons between different jurisdictions or levels of governance. There are punishment/reward systems for different jurisdictions based on elective surgery performance measures.

Possible answers

Federal

Federal-level policy and funding, relationship between the States and the Federal Government

ACT

Specifically relating to the ACT, its budget and governance

ACT Region

Relating to services and policy that covers the ACT and the surrounding region, including the treatment of NSW patients in ACT hospitals

Federal/ACT

Relations between the ACT and the federal governments, includes Federal-level initiatives covering the ACT region.

Question A-4: Does the article explicitly refer to either 'fairness' or 'equity' in the provision of elective surgery or other services in the public hospital system?

Purpose of question

To differentiate between a stated intent of policy initiatives (equity) and how the implementation is perceived (fairness). This distinction comes from Jenny Stewart's *Public Policy Values*.

Possible answers

Fairness

The word 'fair' or words with 'fair' as a root are used in the article

Equity

The word 'equity' or words with 'equity' as a root are used in the article

Both

The word 'fair' or words with 'fair' as a root are used in the article and the word 'equity' or words with 'equity' as a root are used in the article.

Neither

Neither the word 'fair' or words with 'fair' as a root are used in the article nor the word 'equity' or words with 'equity' as a root are used in the article.

Question A-4a: Does the article use the concepts of fairness or equity in the provision of elective surgery or other services in the public hospital system?

Purpose of question

To differentiate between a stated intent of policy initiatives (equity) and how the implementation is perceived (fairness). This distinction comes from Jenny Stewart's *Public Policy Values*. The two words are being used in a tight, technical sense for this coding.

This question was added as a refinement to Question 4.

Commenting on a change in waiting times by itself does not trigger this question. There must be at least an implied judgement about the fairness or equity represented by the change.

Possible answers

Fairness

The article concerns the perception of the fairness of implementing policies/procedures. Typically the article relates to individual cases and specific events. Code as 'fairness' if the implementation of the policy/process is perceived as (in)equitable.

Equity

The article concerns the process of calculating or setting policy guidelines for equitable distribution of resources, typically done at a high level and in the abstract. In this case 'equity' includes equity of access to resources. Code as 'equity' if the *intent* of the policy/process is perceived as (un)fair.

Both

Both the concept of 'fair' as defined above, and the concept of 'equity' as defined above are used in the article

Neither

Neither the concept of 'fair' as defined above, nor the concept of equity as defined above is used in the article

Question A-5: Is elective surgery defined in the article?

Purpose of question

To find out if the key term 'elective surgery' is defined in the article and hence whether there appears to be an assumption about the readership's understanding of the term.

The way the word 'elective' is used in the term 'elective surgery' is very different to the way it is usually used.

Possible answers

Yes

The term 'elective surgery' is used and defined in the article'

No

The tem 'elective surgery' is not defined in the article.

Question A-5a: Is the idea of 'urgency category' used and either defined or implied in the article?

Purpose of question

To find out if and how much the formal urgency categories for elective surgery are being used in describing how long people are waiting for surgery.

The urgency category determines the longest the patient should wait before having their surgery. It sometimes can be contrasted with how long the patient or their carer believed they would have to wait. Under the COAG Hospital Reform Program the categories are:

- Urgency Category 1: Admission within 30 days desirable for a condition that has the potential to deteriorate quickly to the point where it may become an emergency.
  This is sometimes referred to as "Urgent"
- Urgency Category 2: Admission within 90 days desirable for a condition causing some pain, dysfunction or disability but which is not likely to

deteriorate quickly or become an emergency.

This is sometimes referred to as "Semi-urgent"

- Urgency Category 3: Admission within 365 days for a condition causing minimal or no pain, dysfunction or disability, which is unlikely to deteriorate quickly and which does not have the potential to become an emergency.

  This is sometimes referred to as "Non-urgent"

  (in the National Health Data Dictionary the time scale for this category is 'at some time in the future'.)

Possible answers

No

No mention of waiting times for elective surgery.

Used

Either Category 1, Category 2, Category 3 or Urgent, semi-urgent, non-urgent mentioned. Includes using the label with a vague timeframe.

Defined

One of the labels from the list for 'Used' modified by a statement of the maximum recommended number of days waiting is mentioned.

Implied

A (possibly vague) timeframe for when surgery should have happened is mentioned by itself, without a label from the list for 'Used'.

Question A-6: Are indicators and their utility discussed

Purpose of question

To find out if the indicators themselves as well as the performance being measured, is discussed.

Possible answers

Yes

Article mentions disagreement about what indicators mean; discussion of/disagreement about data collection, quality and coding; discussion of which indicator to use; discussion of validity of indicators; discussion of utility of indicators; need for and establishment of National Health Performance Authority.

No

Article reports on discussion of performance; using indicators to describe performance; changes in value of indicators.

Article does not mention performance indicators.

Question A-7: Does the article refer to the concepts of efficiency and/or equity in the provision of elective surgery or other services in the public hospital system?

Purpose of question

To determine if the competing policy values of equity and efficiency are discussed, either explicitly or indirectly.

Commenting on a change in waiting times by itself does not trigger this question. There must be at least an implied judgement about the efficiency or equity represented by the change.

Possible answers

Efficiency

The article mentions efficiency; cost/benefit; value for money or similar values.

Equity

If Question 4 or question 4a is coded as 'equity' or 'both', this question must also be coded as 'yes' for 'equity'. The criteria for coding as 'equity' are identical to those used for questions 4 and 4a.

Both

The concept of efficiency is discussed and the answer to question 4a is either "equity' or 'Both'.

Neither

Neither efficiency not equity are referred to in the article.

If the answer to question 4a is 'Neither' then the only possible answers to this question are 'efficiency' and 'neither'.

Instances of waiting time used as a performance indicator

Question PI-1: Are waiting times used as a performance indicator in the article?

Purpose of Question

To identify each article in which elective surgery waiting time is used as a performance indicator.

Possible answers

- **Yes**

  A waiting time, either a system average or an individual's is compared to something else.

  E.g. using the waiting time in connection with an urgency category or other clinical guidelines; comparing a waiting time in one place to that in another; commenting on whether the waiting time seems short or long compared to the patient's expectations.

- **No**

  A waiting time is used in isolation without reference to a comparator. Waiting times are not mentioned.

Question PI-2: What type of statistic is used to describe elective surgery waiting time?

Purpose of Question

To categorise how elective surgery waiting time is expressed. Statistically speaking the simplest expressions are counts, followed by individual waiting times. Next in complexity is the median for a single urgency category followed by the median across all urgency categories.

Possible answers

- **Counts:**

  Performance indicators relating to elective surgery that are expressed as counts include: people added to waiting list; people on the waiting list; and

surgeries performed. While the latter is not explicitly a waiting time, is it used as part of the suite of performance measures for elective surgery.

- **Individual waiting time**

  This is usually expressed relative to the clinical guidelines and is implied within phrases such as 'waiting too long for urgent surgery'.

- **Median waiting times**

  Medians rather than means are used to describe state and national average waiting times. The two medians used are: the median for a single urgency category and the median for all people waiting for surgery in the jurisdiction of interest. The latter is statistically difficult as if refers to performance against an unknown mix of urgency categories, each of which has a different recommended waiting time. This difficulty is flagged in the coding for 'Statistical Assumptions' below.

Question PI-3: With what is the waiting time statistic compared?

Purpose of Question

To count and classify the comparators for each statistic identified in Question PI-2. The most complex cases were a combination of three comparators.

Possible answers

- **Count of people seen within time comparator**

  Used as a second or third comparator when an earlier comparator is time-based e.g. a guideline for an urgency category. Statistically difficult if comparison is between two populations of widely different sizes. This difficulty is dealt with in the coding for 'Statistical Assumptions' below.

- **Doctor's reported recommendation**

  Optimal waiting time according to treating doctor, as reported by patient. No clear reference to guidelines or urgency categories. E.g. 'Told he needed surgery within a fortnight' [I01232010].

- **Guideline for appropriate urgency category**

  There is enough information to determine that the waiting time statistic is being compared to the guideline for a specific, and by inference, appropriate urgency category. This is not limited to instances where an urgency category

is specified. The phrase 'waiting too long' is taken to mean waiting longer than the guidelines for the patient's urgency category.

- **Guideline for least urgent category (1 year)**

  The waiting time statistic, either 'Individual Waiting Time' or a median is compared to one year, which is also the guideline for Category 3 (non-urgent) surgery.

- **Median for the appropriate urgency category**

  This is used when the waiting time statistic is 'Individual waiting time' and an urgency category is given or able to be inferred.

- **Median for a mix of urgency categories**

  A median for elective surgery waiting times across all urgency categories. Since the proportion of each urgency category in the statistical population is not given, this is statistically difficult. This difficulty is dealt with in the coding for 'Statistical Assumptions' below.

- **Percentage of people seen within time comparator**

  Used as a second or third comparator when an earlier comparator is time-based e.g. a guideline for an urgency category. Comparing percentages of different sized populations can be difficult to understand. This difficulty is dealt with in the coding for 'Communication Assumptions' below.

- **Result from previous time period**

  The same statistic, and other comparators if any, is compared between two different time periods.

- **Result in other state(s)**

  The same statistic, and other comparators if any, is compared between two different state(s) or against national results. Used when the ACT's rank (e.g. worst in Australia) is mentioned, even if no state jurisdiction is referred to.

Question PI-4: Which part of the public health system's performance is being measured?

Purpose of Question

To identify the part of the public health system for which performance is being measured.

Possible answers

- **ACT**

  The waiting list statistic is for patients using the ACT public hospital. system for elective surgery.

- **National**

  The waiting list statistic is for all patients in Australia using the public hospital system for elective surgery.

- **Other State(s) − excludes ACT**

  The waiting list statistic is for patients using a state public hospital system other that the ACT's.

Question PI-5: What statistical assumptions are needed for the comparison to be statistically valid?

Purpose of Question

To codify any unwritten assumptions needed for a comparison to be statistically valid. The assumptions fall into two broad categories: those required to ensure that like is compared with like and those relating to what inferences can be drawn from the statistic.

Possible answers

Comparison assumptions:

- **All comparators assumed**

  This barely qualifies as a performance indicator, a median is reported with no explicit comparators.

- **Doctor's recommendation is a firm guideline**

  When a doctor gives a recommended time within which surgery should occur but this is not related to a formal urgency category, the assumption is that the recommendation takes the place of the category guideline.

- **Each urgency category compared to itself**

  If a combination of urgency categories or an imprecise description of urgency category such as 'less urgent' is compared between time periods or state(s) the assumption is that each urgency category is compered to itself.

- **Guidelines less than time already waited**

  Used for comparisons of 'patients/people waiting too long'. The assumption is that the time people have waited is greater than the clinical guideline for their urgency category.

- **Rate of people joining list is constant**

  Comparisons of the number of people on the waiting list, either by time period or state(s) is only a performance indicator if the rate of people joining the list is the same for both parts of the comparison.

- **Time to be removed from the waiting list is the same as waiting time**

  There are two periods of time that are referred to as 'waiting time': the total time from joining the list until surgery occurs and the time from joining the list until the reporting date. These are sometimes confused or conflated.

- **Urgency category mix the same in each state**
  Urgency category mix unchanged over time

  Performance judgements of comparisons of the median waiting time for all people on the elective surgery waiting list are strongly affected by the mix of urgency categories. To take an extreme example, a median waiting time of 45 days is a good result if everyone on the list is urgency category 3 (treatment within 365 days) but a poor result if everyone is urgency category 1 (treatment within 30 days).

Inferential assumption:

- **Median tells you about an individual**

  Articles where information about an individual's likely waiting time were inferred from the median (or a change in the median) for a population assume that this inference is possible. In practice the median for a population tells you nothing about where an individual might be in the distribution of that population

Question PI-6: What communication assumptions are needed for the comparison to be meaningful on a human level?

Purpose of Question

To codify any communication assumptions relating to the instance of waiting time being used as a performance indicator. In some senses the statistical assumptions covered by Question PI-5 are also communication assumptions but there are also some assumptions with a clearer communication basis. This question is tightly tied to Question PI-8 which deals with value judgements about performance.

Global assumptions

Since the term 'elective surgery' is not defined in any article in the collection a communication assumption that holds for every instance of indicator use is that the reader knows what elective surgery is. Closely related to this is the assumption that the reader knows the clinical guidelines for each urgency category. A more subtle global assumption is that the time elapsed since joining the waiting list is the same as the patient's waiting time.

Possible answers

- **Personal experience is more 'true' than statistics**

   This is used when the statistics are described as 'untrue' or 'wrong' because they do not reflect someone's personal experience.

- **Percentages of different sized populations are easy to compare**
   Percentage increases are easy to understand

   These two assumptions underlie comparisons based on percentages and percentage changes..

- The following codes capture assumptions about what movement in the indicator means and form part of the basis for the next coding question relating to value judgements.

   Fewer operations is a bad outcome

   This would not be true if the reason for there being fewer operations was that fewer people needed them.

   Guideline is shorter than time comparator

   This is used when the average waiting time for a category is compared to a longer timeframe without the guideline time period being specified.

Lower proportion of people waiting more than a year is good

How good an outcome this is depends upon the mix of urgency categories in the two time periods or states being compared.

More surgeries is a good outcome

This is would not be true if the reason for there being more surgeries was that more people needed them

'Shortest' means fewest people on the list

The description 'shortest waiting list' could either mean the list with the fewest people on it or the list with the shortest waiting time. In the instances where the description is used, the former is a better contextual fit for the intended meaning.

Question PI-7: What value judgement about performance is expressed?

Purpose of Question

To categorise any value judgements about the performance being measured e.g. good result, disgraceful figures, not good enough.

Possible answers

- **Neutral**

  No comment on the performance being measured.

- **Good**

  Comment on the performance being measured uses words with an element of judgement such as 'dropped', 'best', 'improved', 'good' and 'excellent'.

- **Bad**

  Comment on the performance being measured uses words with an element of judgement such as 'rose', 'decline', 'should have been seen by', 'too long' 'blown out', 'worst' and 'poor'.

  If the ACT performance was judged to be bad, but not as bad as another state (e.g. I0528201 "More than 10 per cent of ACT elective surgery patients were forced to wait more than a year to undergo a procedure. Only Tasmania performed worse…") the instance was coded as 'Bad'.

Question PI-8: Does the value judgement reflect the meaning of the statistics?

Purpose of Question

To capture whether or not a value judgement expressed about performance is consistent with the meaning derived from the statistics.

Possible answers

- **Yes**

  The value judgement is consistent with the meaning of the statistic.

- **Partly/Unable to tell**

  The value judgement is at least partly consistent with the meaning of the statistic.

  It is not possible to tell if the value judgement is consistent with the meaning of the statistic. The most obvious reason for this would be that there are too many statistical and communication assumptions required to be certain that like is being compared with like.

- **No**

  The value judgement is inconsistent with the meaning of the statistic