



Variation and improving services: analysing and interpreting variation

Authors: Tom Love and Nieves Ehrenberg, Sapere Research Group

Date: 5 March 2014

Contents

Executive summary	3
1. Observing variation	5
1.1 What does variation say about clinical practice?	5
2. Analysing variation	9
2.1 Data completeness.....	9
2.2 Data coding	9
2.3 Small numbers and statistics	11
2.4 Appropriate standardisation	13
2.5 Control charts	15
2.6 Populations and time periods.....	16
2.7 Presenting data	17
3. Addressing the key questions	22
4. What to do about variation	27
4.1 Need to manage variation.....	27
4.2 Approaches for managing variation	27
5. Case study: polypharmacy in the elderly	32
5.1 Background	32
5.2 Review	36
5.3 Analysis.....	39
5.4 Actions	43
5.5 Commentary.....	45
6. Case study: ventilation tubes	47
6.1 Background	47
6.2 Review	49
6.3 Analysis.....	52
6.4 Actions	55
6.5 Commentary.....	56
7. References	57

*Published May 2014 by the Health Quality & Safety Commission,
PO Box 25496, Wellington 6146, New Zealand
ISBN: 978-0-478-38582-3 (print)
ISBN: 978-0-478-38583-0 (online)
Available online at: www.hqsc.govt.nz*

Executive summary

Addressing variation in health services is one of the major challenges for clinicians and managers, both in New Zealand and internationally. Observing variation raises difficult questions about the equity, quality and consistency of care, and is an important starting point for quality improvement and service development. Better management of variation is one of the core elements of performance improvement identified by an Expert Advisory Group to the Ministry of Health in developing a new Integrated Performance and Incentive Framework (Expert Advisory Group 2014).

The Health Quality & Safety Commission's Atlas of Healthcare Variation (the HQSC Atlas) is a starting point for analysing variation in a local area, and for developing service improvement activities to address variation. This guidance provides advice and examples, based on the HQSC Atlas, of analysing variation and working towards reducing variability and improving care.

The following questions form the basis for a systematic approach to analysing the causes and consequences of variation, and for making the decision about whether variation matters enough to engage in quality improvement activity:

- Are the data complete and accurate?
- Is there uncertainty or ambiguity in the clinical evidence?
- Are there quality issues?
- Is there inequity for patients?
- Is there inefficient use of resources?
- What environmental and population factors are relevant?
- How much variation should you expect?

If variation does matter, and there is an adverse impact which justifies working to improve care, there are generically five areas of activity which can contribute to managing variation and improving quality of care for the population:

- analysis
- patient expectations
- work with practitioners at the extremes of variation
- using and disseminating clinical information
- clinical consensus and service design.

Two examples of variation derived from the HQSC Atlas are examined with this generic framework, and supplementary analysis is provided. These illustrate

different aspects of variation, and some of the various approaches which can be used to analyse information about variation, and to support improved practice.

Ultimately, managing variation is about the core functions of clinical governance. Both in New Zealand and internationally, there is increasing interest in using observations of variation as the starting point for identifying and prioritising quality improvement and clinical governance activities. The HQSC Atlas approach can support careful analysis, judicious use of clinical judgement, and challenging work with clinical professionals. Effectively utilised, there is enormous potential for positive effect in the form of improved health services for patients and communities.

1. Observing variation

1.1 What does variation say about clinical practice?

Different categorisations of variation have been used by researchers and commentators which reflect different emphases upon what variation means in terms of clinical practice. Categorisations include:

- *small area variation*: variation in population rates observed across geographical areas, typically hospital districts or local government areas. The results presented in the Health Quality & Safety Commission's Atlas of Healthcare Variation (the HQSC Atlas) explicitly base variation analysis on district health board (DHB) areas, and in this sense is an example of small area variation.
- *medical practice variation*: variation in clinical decision-making between different health professionals. Analysis of, for example, referral rates of different general practitioners within a primary health organisation (PHO) is implicitly presenting information as a form of medical practice variation in which the denominator population is not necessarily defined geographically, and there is a strong correspondence between each unit of variation and a clinical decision-maker.

It is important to be clear that area variation and medical practice variation may be related, but that this isn't necessarily the case. Observing variation at one level does not necessarily imply that there will be variation at the other level. For example:

- Every DHB in the country might have exactly the same average level of referral for some condition, but there might be significant variation among the referrers within all DHBs, representing a wide range of practice in primary care. One example is medication for patients with cardiovascular disease (CVD), which shows relatively little variation across New Zealand but consistent difference in practice for male and female patients (see Figures 1 and 2).

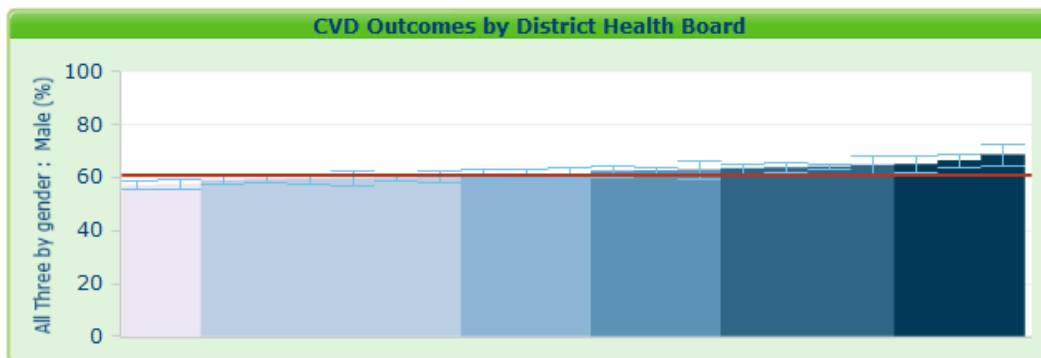


Figure 1: Three medicines for CVD: male

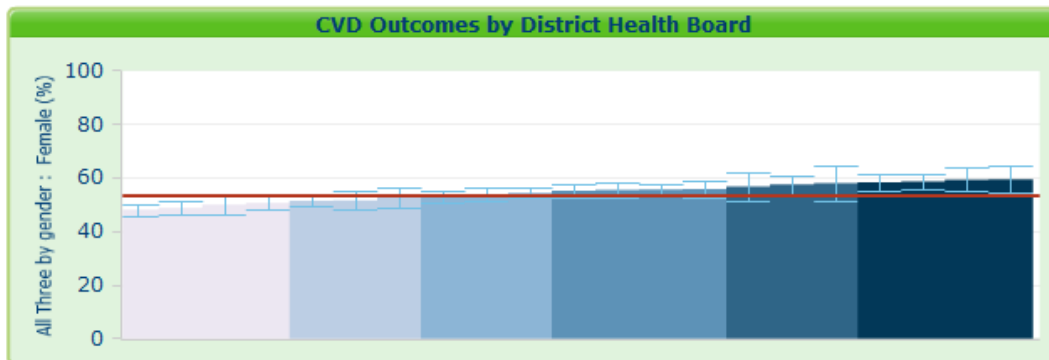


Figure 2: Three medicines for CVD: female

- Conversely, all the referrers in a single area might practise very consistently with each other, but the referrers in a different area might have a different consensus, and show a different area level of intervention, even though in both cases there is very little variation in practice among clinical colleagues. This would reflect environmental influences upon practice which vary from area to area (such as referral protocols, resource availability, or practice among particular hospital specialists). This is likely to be the case in use of tonsillectomy (see Figure 3).

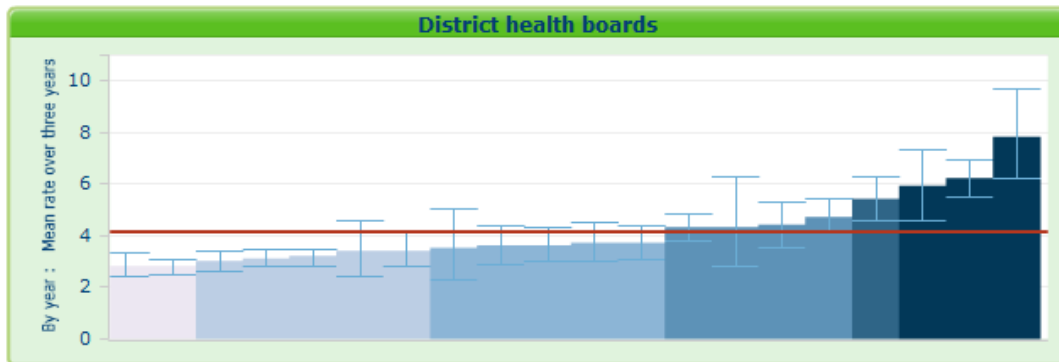


Figure 3: Tonsillectomy per 1000 children aged 0–14

These are quite different situations, and would suggest quite different approaches to thinking about why variation is happening and what, if anything, should be done to improve the situation. Reasoning from an observation of area variation to a conclusion that there must be variation in medical practice can sometimes be misleading.

Another approach to thinking about the impact of variation is to consider the kind of care and how susceptible it might be to different influences upon clinical decision-making. Wennberg (2011) has developed the following approach.

- *Effective care*: defined as interventions for which the benefits far outweigh the risks; in this case the “right” rate of treatment is 100 percent of patients defined by evidence-based guidelines to be in need. Unwarranted variation is generally a matter of underuse.
- *Preference-sensitive care*: when more than one generally accepted treatment option is available, such as elective surgery. The right rate should depend on informed patient choice, but treatment rates can vary extensively because of differences in professional opinion.
- *Supply-sensitive care*: clinical activities such as doctor visits, diagnostic tests, and hospital admissions for which the frequency of use relates to the capacity of the local health care system. The key issue with this one is that, at least in the United States, those living in regions with a high-intensity pattern of care have worse or no better survival than those living in low-intensity regions. This means that greater intensity of care does not necessarily equate to improved outcomes.

This taxonomy explicitly addresses the issue of how clinicians and patients make decisions about care, and makes useful distinctions between the kinds of influences which might generate variation. It should be noted that in all cases, the definition of a single right rate of care at a population level is very difficult, although in the case of effective care a good understanding of the underlying epidemiology of a condition might provide an estimate of how much treatment

should be expected. It is therefore important to be clear about what measure of variation is meaningful to clinicians, and will be credible at a local level to those involved in providing care.

2. Analysing variation

Simply describing that there exists a range of activity on some measure does not necessarily indicate that there is variation, or that clinical judgement is being applied differently. Robust analysis of variation, which draws useful conclusions for clinicians about their practice, can be complicated to perform, and may sometimes require considerable technical skills in data analysis and interpretation.

2.1 Data completeness

Sometimes there are systematic reasons why data are more complete in some areas than in others. For example, information from privately funded hospital care is not included in national datasets of public hospital admissions, so differing levels of private hospital use can make a difference to the apparent volumes of hospital events, even though the true total utilisation across the population might not be different from anywhere else. Similarly, where some medicines are available both over the counter and on prescription, national datasets will only capture the medicines dispensed on prescription, rather than those purchased directly over the counter by the patient. Incompleteness can therefore have the effect of either masking true variation, or of producing artefactual variation where none actually exists. Only careful consideration of the specific example of variation in question, and what the possible kinds of incompleteness are, will help in understanding the impact of data incompleteness upon your conclusions about variation. Incompleteness in data need not necessarily stop the analysis of variation, but it is important to be clear about what datasets cover, and therefore how resulting analyses can or should be interpreted.

Checking data completeness and whether there are systematic effects which determine whether data is recorded in the first place is a very important first step in being able to distinguish real from apparent variability. Where data are incomplete, the challenge to the analyst and clinician researching variation is to decide whether this is a serious issue for the analysis, to be clear about the most legitimate interpretation of the information, or to find a suitable piece of proxy information which is less affected by problems with completeness.

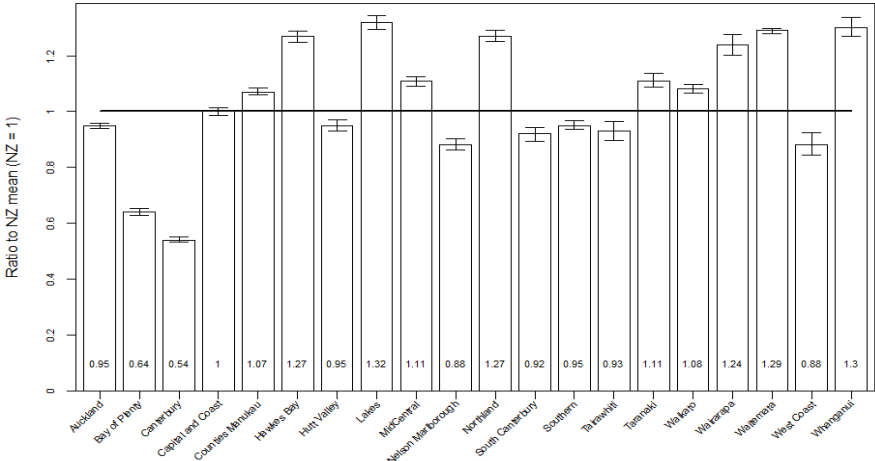
2.2 Data coding

Coding of diagnoses in hospitals is done by trained coders who read a patient's clinical notes and assign standardised codes from the International Classification of Diseases. By contrast, diagnosis coding in primary care is usually done by the general practitioner or nurse, if they wish to do so in addition to writing clinical

notes, and therefore can be incomplete and sometimes inconsistent across different clinicians within a practice or between practices.

Coding accuracy may be less important if you can be sure that, even if it's not accurate, coding has at least been done consistently. Much depends upon the particular question or issue for analysis, but a common approach is to aggregate codes to a higher, less detailed level, which can lose detailed information but improve the comparability of results.

Coding consistency																																											
Issue	<p>Different hospitals have different practice in coding acute medical admissions as being managed by different specialties. In particular, some hospitals tend to code acute medical admissions as general medicine, while others record the specialty managing the admission as emergency medicine. This can make it difficult to compare acute medical admission rates across hospitals or DHBs.</p> <p>The graph below shows the observed variation in age/sex standardised emergency medicine discharges across DHBs (source: National Minimum Dataset).</p> <div style="text-align: center;"> <p>Standardised discharge ratio for DHBs 2011/12: acute or arranged emergency medicine discharges</p> <table border="1"> <caption>Data for Standardised discharge ratio for DHBs 2011/12</caption> <thead> <tr> <th>DHB</th> <th>Ratio to NZ mean (NZ = 1)</th> </tr> </thead> <tbody> <tr><td>Auckland</td><td>1.08</td></tr> <tr><td>Bay of Plenty</td><td>0.09</td></tr> <tr><td>Canterbury</td><td>0.29</td></tr> <tr><td>Capital Land Coast</td><td>1.13</td></tr> <tr><td>Counties Manukau</td><td>1.03</td></tr> <tr><td>Hawkes Bay</td><td>0.98</td></tr> <tr><td>Hutt Valley</td><td>0.87</td></tr> <tr><td>Lakes</td><td>1.38</td></tr> <tr><td>MidCentral</td><td>1.49</td></tr> <tr><td>Nelson Marlborough</td><td>0.81</td></tr> <tr><td>Northland</td><td>1.05</td></tr> <tr><td>South Canterbury</td><td>0.04</td></tr> <tr><td>Southern</td><td>1.16</td></tr> <tr><td>Tairāwhiti</td><td>0.42</td></tr> <tr><td>Tairānaki</td><td>1.56</td></tr> <tr><td>Waikato</td><td>1.45</td></tr> <tr><td>Waikaregata</td><td>1.28</td></tr> <tr><td>Wairarapa</td><td>1.28</td></tr> <tr><td>West Coast</td><td>0.04</td></tr> <tr><td>Wairarapa</td><td>2.0</td></tr> </tbody> </table> </div>	DHB	Ratio to NZ mean (NZ = 1)	Auckland	1.08	Bay of Plenty	0.09	Canterbury	0.29	Capital Land Coast	1.13	Counties Manukau	1.03	Hawkes Bay	0.98	Hutt Valley	0.87	Lakes	1.38	MidCentral	1.49	Nelson Marlborough	0.81	Northland	1.05	South Canterbury	0.04	Southern	1.16	Tairāwhiti	0.42	Tairānaki	1.56	Waikato	1.45	Waikaregata	1.28	Wairarapa	1.28	West Coast	0.04	Wairarapa	2.0
DHB	Ratio to NZ mean (NZ = 1)																																										
Auckland	1.08																																										
Bay of Plenty	0.09																																										
Canterbury	0.29																																										
Capital Land Coast	1.13																																										
Counties Manukau	1.03																																										
Hawkes Bay	0.98																																										
Hutt Valley	0.87																																										
Lakes	1.38																																										
MidCentral	1.49																																										
Nelson Marlborough	0.81																																										
Northland	1.05																																										
South Canterbury	0.04																																										
Southern	1.16																																										
Tairāwhiti	0.42																																										
Tairānaki	1.56																																										
Waikato	1.45																																										
Waikaregata	1.28																																										
Wairarapa	1.28																																										
West Coast	0.04																																										
Wairarapa	2.0																																										
Approach	<p>In this case, one strategy for improving the comparison is to combine the coded specialties of emergency medicine and general medicine, to give a higher-level, but more consistent, picture of acute discharges. The resulting pattern of acute and arranged discharges across DHBs looks</p>																																										

	Coding consistency																																										
	<p>quite different to the pattern for emergency medicine only (source: National Minimum Dataset). This is a good example of a trade-off between analysing data at a more aggregate level, which can provide better consistency, versus analysing data at a lower level, which will give more detail but is less robust for at least some of the DHBs.</p> <p style="text-align: center;">Standardised discharge ratio for DHBs 2011/12: acute or arranged general and emergency medicine discharges</p>  <table border="1" data-bbox="422 694 1300 1153"> <thead> <tr> <th>DHB</th> <th>Ratio to NZ mean (NZ = 1)</th> </tr> </thead> <tbody> <tr><td>Auckland</td><td>0.95</td></tr> <tr><td>Bay of Plenty</td><td>0.64</td></tr> <tr><td>Canterbury</td><td>0.54</td></tr> <tr><td>Capital and Coast</td><td>1</td></tr> <tr><td>Counties Manukau</td><td>1.07</td></tr> <tr><td>Hawke's Bay</td><td>1.27</td></tr> <tr><td>Hutt Valley</td><td>0.95</td></tr> <tr><td>Lakes</td><td>1.32</td></tr> <tr><td>McCombs</td><td>1.11</td></tr> <tr><td>Nelson Marlborough</td><td>0.88</td></tr> <tr><td>Northland</td><td>1.27</td></tr> <tr><td>South Canterbury</td><td>0.92</td></tr> <tr><td>Southern</td><td>0.95</td></tr> <tr><td>Tairāwhiti</td><td>0.93</td></tr> <tr><td>Tairāhiki</td><td>1.11</td></tr> <tr><td>Waikato</td><td>1.08</td></tr> <tr><td>Waikarepa</td><td>1.24</td></tr> <tr><td>Waitemata</td><td>1.29</td></tr> <tr><td>West Coast</td><td>0.88</td></tr> <tr><td>Wairarapa</td><td>1.3</td></tr> </tbody> </table>	DHB	Ratio to NZ mean (NZ = 1)	Auckland	0.95	Bay of Plenty	0.64	Canterbury	0.54	Capital and Coast	1	Counties Manukau	1.07	Hawke's Bay	1.27	Hutt Valley	0.95	Lakes	1.32	McCombs	1.11	Nelson Marlborough	0.88	Northland	1.27	South Canterbury	0.92	Southern	0.95	Tairāwhiti	0.93	Tairāhiki	1.11	Waikato	1.08	Waikarepa	1.24	Waitemata	1.29	West Coast	0.88	Wairarapa	1.3
DHB	Ratio to NZ mean (NZ = 1)																																										
Auckland	0.95																																										
Bay of Plenty	0.64																																										
Canterbury	0.54																																										
Capital and Coast	1																																										
Counties Manukau	1.07																																										
Hawke's Bay	1.27																																										
Hutt Valley	0.95																																										
Lakes	1.32																																										
McCombs	1.11																																										
Nelson Marlborough	0.88																																										
Northland	1.27																																										
South Canterbury	0.92																																										
Southern	0.95																																										
Tairāwhiti	0.93																																										
Tairāhiki	1.11																																										
Waikato	1.08																																										
Waikarepa	1.24																																										
Waitemata	1.29																																										
West Coast	0.88																																										
Wairarapa	1.3																																										

2.3 Small numbers and statistics

Small numbers of patients or relatively rare events can magnify random effects, and give the appearance of substantial variation where there is none (Diehr 1990). It is therefore important to be careful about the formal statistics of variation, and to take a robust approach to the analysis. Otherwise there is a high risk of responding to variation which is not real, or at least not robustly demonstrable with the available information. Where population rates of a procedure are relatively low, or the denominator population is relatively restricted (eg, children under 5, or Pacific people aged 65–75), then technical issues arising from small numbers can be an analytical challenge.

It is important to make sure that confidence intervals and appropriate statistical tests are conducted when looking at variation, but sometimes small numbers can mean that variability is inherently very difficult, or even impossible, to measure. This is particularly a problem in primary care settings where, because the service is generalist, many individual conditions or events are only seen in relatively small numbers. For example, the incidence of new myocardial infarctions (MIs) is

usually quite low for the enrolled population of a single general practitioner, so measuring secondary prevention prescribing for new MI patients might be quite difficult to do robustly at practice level.

A classic example in primary care might be back pain, where one general practitioner can see large numbers of patients, but another might see only one or two cases over a year. If you see two patients per annum, your referral rate (eg, for radiology) must be, by definition, 0, 50 or 100 percent. A tiny random variation in the referral process can make the difference between observing a rate of 0 and 100 percent. It may be possible to aggregate data over multiple years in order to achieve more reliable statistical results, if data are available, but this kind of approach potentially reduces the ability to monitor trends. Where numbers are very small, analysis should not be attempted, since there is a risk of discrediting the analysis entirely if invalid results are presented to clinicians. As a rule of thumb, there should be thirty or more events per clinician (or practice, or whatever unit of analysis is being used) before analysis is attempted. Where there are less than 100 events, close attention should be paid to the statistics of small numbers.

It is important to bring some level of statistical rigour to the analysis of variation. At a minimum, it is good practice to calculate confidence intervals for the areas or practitioners in the dataset, exemplified in the HQSC Atlas. Calculating confidence intervals is not usually complex, and can serve as a good initial test of whether apparent outliers in the dataset really are different from the population mean. Figure 4 illustrates this in the HQSC Atlas “Trauma” chapter. The majority of all DHBs are significantly different from the national mean, indicating that there is meaningful variation across DHBs.

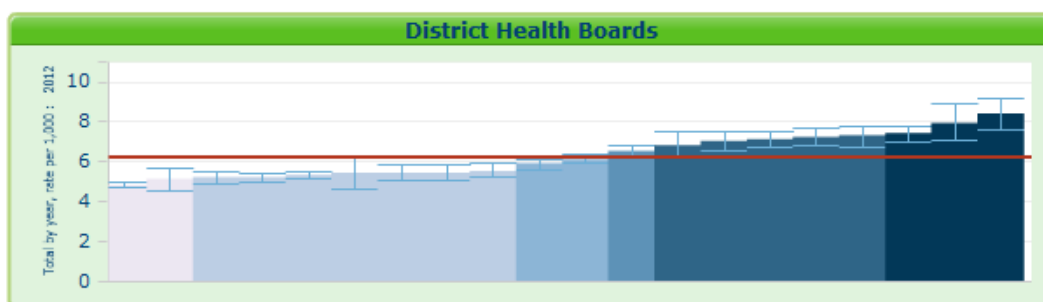


Figure 4: Admissions due to injury by DHB per 1000 population

By contrast, Figure 5 shows the variation across DHBs in number of women with CVD receiving three or more preventative medicines (statins, blood pressure lowering medication, and antiplatelets). The confidence intervals indicate that, on the basis of these data, few DHBs are significantly different from the national rate. This example illustrates an important variation in practice for different

populations (there is a substantial difference in prescribing rates for men and women), which doesn't show up at all as variation when you take the perspective of looking across districts.

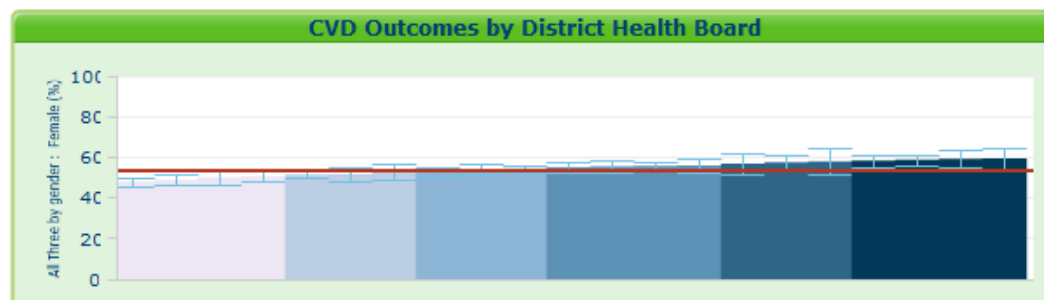


Figure 5: Women receiving three medicines for CVD by DHB

Other more sophisticated techniques can be used to explore the reality of variation, and the impact of small numbers on the apparent variation across practitioners or areas. These include multilevel modelling, in which unexplained variation is attributed explicitly to a different level of aggregation such as to a prescriber or a district, and simulation techniques in which the expected level of variation expected from the underlying statistics of the distribution across practitioners or districts can be calculated, and compared with the observed degree of variation. This can be a useful approach for estimating how much variation there should be from first principles, and then deciding whether there is more variation than you would expect when you look at the real data.

Resource
1. Web-based calculator from Vassar, including confidence intervals for a mean, and confidence interval for a proportion: http://vassarstats.net/
2. Formulae for confidence intervals in a variety of situations commonly found in health data: http://www.apho.org.uk/resource/item.aspx?RID=48457
3. Calculating Byars' approximation for confidence intervals of a ratio: http://www.cihi.ca/cihi-ext-portal/pdf/internet/hsmr_tech_notes_en

2.4 Appropriate standardisation

Comparing measures across different areas or enrolled populations typically involves either restricting the comparison to specific comparable populations (eg, a narrow age range) or adjusting for demographic characteristics so that the underlying rate of intervention can be compared meaningfully. The HQSC Atlas

does not standardise its comparative data, but provides the ability for the user to drill into specific age, sex or ethnicity groups in order to make direct comparisons.

Standardisation may also take into account ethnicity, deprivation and other factors. Age adjustment is particularly important, since health need varies so significantly by age. Comparison of intervention rates that do not have some sort of age adjustment are always highly suspect. Even if adjusting for age of the population isn't straightforward, calculating rates for different age bands can be a simple way of making sure that age is taken into account when presenting data.

Taking population characteristics into account in this way can be important in making sure that results are comparable, but it is important to think carefully about the particular case, and to make sure that standardising does not hide important differences, or build them into expectations for future care. For example, across New Zealand, HQSC found that the percentage of men with CVD prescribed three preventive medicines ranged from 57.0 to 68.6 percent, whereas for women the range was 47.9 to 59.3 percent. Applying age/sex standardisation to this data would mask an important issue about management of heart disease for women compared with men.

Similarly, standardising data can build inequity into the standard rate of intervention. If, for example, standardising access to elective surgery takes into account the ethnicity and deprivation structure of the population, then current inequities in access to care for different ethnic and deprivation populations will form part of the benchmark against which variation is assessed. So, unduly enthusiastic standardisation has the risk of both masking important differences in care, and of building such differences into the expected pattern of care.

In some cases it can be possible to use more sophisticated techniques to standardise or control for differences, based upon multiple regression. These techniques can be powerful, but must be used with care. There can be a risk of overfitting regression models, and producing results which are not robust and repeatable, even where they appear to have worked well in the baseline case.

Resource
Simple guide to direct and indirect standardisation of health statistics, with worked examples: Borman B. 1995. <i>Standardising rates of disease</i> . Wellington: Ministry of Health. http://www.health.govt.nz/publication/standardising-rates-disease

2.5 Control charts

Control charts are increasingly being used in the analysis of New Zealand health services, and tools for generating them are becoming more widely available. Control charts have been widely used in some circumstances, particularly in the United Kingdom National Health Service, where this approach to presenting variation data has become commonplace, and the ability to interpret control charts has become widespread.

Control charts emerge from industrial process engineering, where people monitoring the quality of a product seek to determine whether variation in quality is a result of common cause ('normal' variation, usually associated with statistical randomness) and special cause (an inherent change in the underlying process which is increasing or decreasing the measure being monitored). Where several observations of the quality measure are sequentially above or below marker levels (known as control limits), the process is deemed to be 'out of control', and it is time to investigate whether there is some special cause which has changed the process (a machine has malfunctioned, or a setting has changed).

Control charts were originally developed to apply to processes being monitored over time, to see whether a production process was changing or whether it was stable. Applying control charts to a cross section of data, rather than to a longitudinal measure, is controversial and should be done very carefully. An alternative for cross sections of data is to use a funnel plot, which shows how areas or practices of differing sizes are distributed across the mean and standard deviations of the whole distribution.

Another important limitation of control charts is that the underlying concept for setting control limits relies upon the base data being normally distributed, which is often not the case in health care measures.

Resource
Introduction to control charts in a health context, with worked examples: Amin SG. 2001. Control charts 101: A guide to health care applications. <i>Quality Management in Health Care</i> 9(3): 1–27. http://www.npaihb.org/images/training_docs/NARCH/2010/Amin%20Control%20Charts%20101%20Quality%20Management%20in%20Health%20Care%202001.pdf

2.6 Populations and time periods

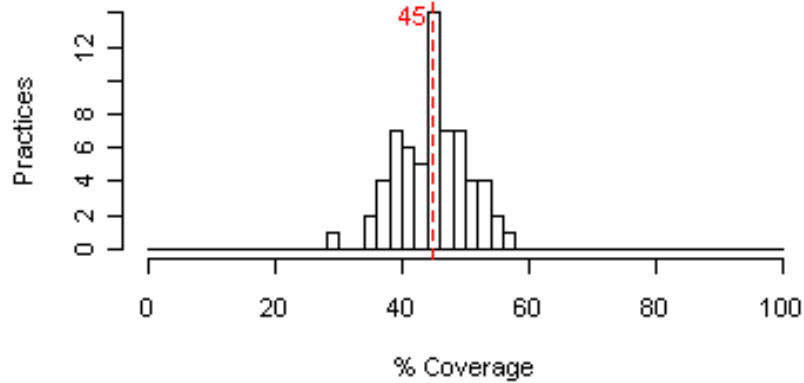
The easiest definition of a population is a simple cross section: the number of people who live in a particular place, or who are enrolled on register at a given point in time; for example, the population enrolled in a PHO or general practice on today's date. This is technically straightforward, but can sometimes mask changes in clinical practice over time, making comparison difficult. For example, if the number of people in the population with a diagnosis of angina includes both people diagnosed a decade ago who have been stable ever since, and people diagnosed last week, then practice might have changed in the interim. Referral to a cardiologist might be recommended for a newly diagnosed patient, but might not have been the standard a decade previously, and might not be considered appropriate for someone who has been stable over a long period.

Cross-sectional analysis can therefore sometimes hide differences in clinical practice over time. In some cases it can be more robust to analyse a cohort of patients who present in a given period, rather than a whole population. It is important to consider the perspective of the clinician, and how people present to them, as well as the overall population perspective.

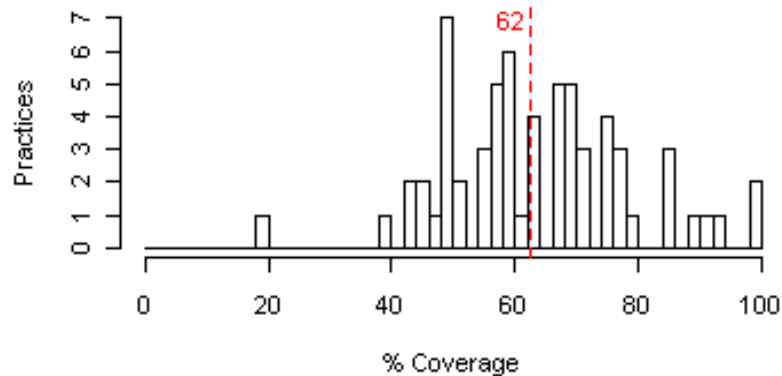
Sometimes cohort effects are not significant, or they may serve as a marker for clinical practice which people are interested in – much depends upon the particular case and the nature of clinical practice in the field. But it is important to consider cross-section effects, and whether they might be affecting the observed variation in a patient group.

	Example: Cross sections and cohorts
Issue	In trying to understand how much variation exists across a group of practices in secondary prevention for heart disease, one measure was the number of patients who had an MI and had been prescribed beta blockers. But guidelines and incentives for prescribing beta blockers had changed considerably over the preceding decade.
Result	The histogram of percentage of MI patients receiving a beta blocker in a cross section of MI patients enrolled in practices looked like this, with a relatively low level, and a relatively narrow spread (the average value is indicated with a dotted line and a number above the histogram):

Example: Cross sections and cohorts



The same measure applied to a cohort of newly diagnosed MI patients (18 months, finishing at the time that the cross section was taken) has a higher average, but considerable range across practices.



(Source: data from Tayside Health Board)

The cohort picture is quite different, although the variability in this diagram is also affected by relatively small numbers of patients in the cohort compared to the cross section.

2.7 Presenting data

Once data have been analysed and interpreted, it is important to present the information in a way which is clear and meaningful to a clinical audience. There are a number of different approaches which have various strengths and weaknesses.

Ranked bar graph

A common method for presenting variation is to rank each different area, practice or clinician on a measure, and present the result as a bar graph or a league table in visual form. The HQSC Atlas uses this approach, among other methods of data presentation. The example in Figure 6 shows hospital admissions where the primary diagnosis is gout by DHB.

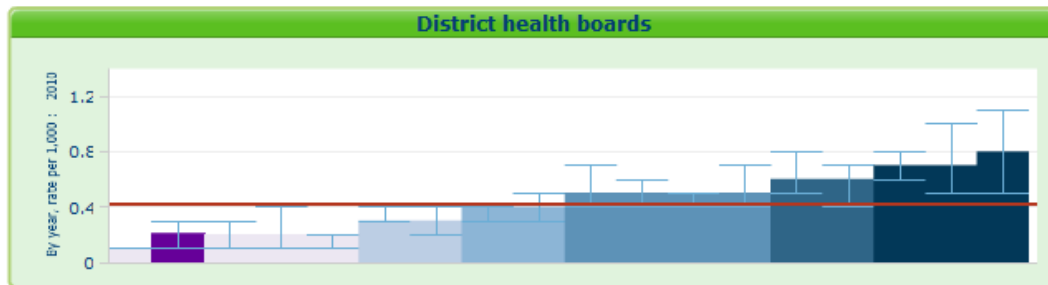


Figure 6: Hospital admissions for gout/1000 population

Advantages

Ranked bar graphs:

- are easy to calculate
- make it easy to show a reference rate, such as a national or local average
- can show confidence intervals for an individual DHB or practice
- can show how far an individual is from the reference rate, or from another individual.

Disadvantages

Ranked bar graphs:

- visually focus attention upon the extreme outliers
- make it difficult to understand the shape of the whole distribution
- can be difficult to compare visually to other bar graphs, particularly where the mean is different.

Histogram

A more formally correct way to look at variation is to plot a histogram, and examine the bell curve which results. This can show the shape of the distribution, and allow for easier visual judgement about the range of variation.

For example, the histogram in Figure 7 shows the same information as in the bar chart above. The curve is very flat, with no clear central tendency.

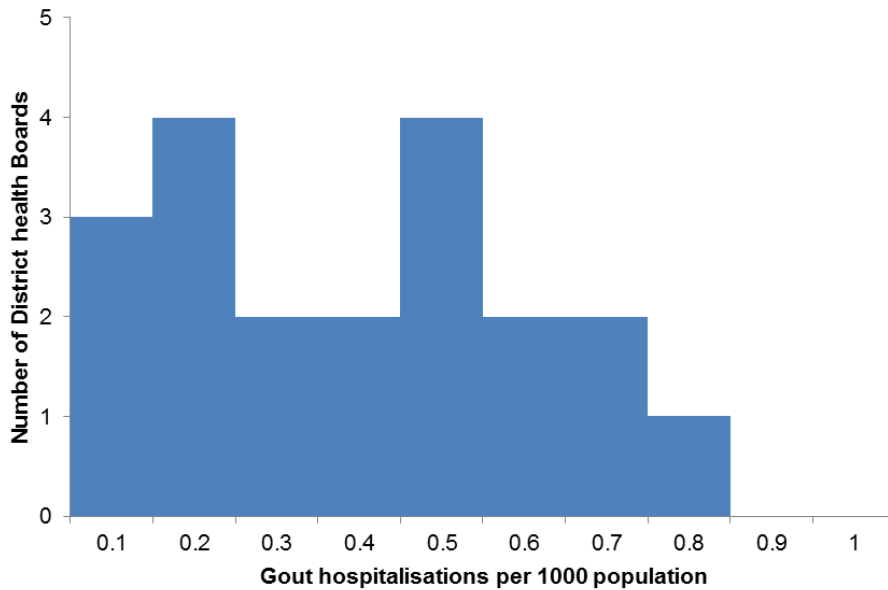


Figure 7: Hospital admissions for gout/1000 population

Advantages

Histograms:

- show a whole distribution, and make outliers obvious
- allow easy visual inspection of the narrowness or width of the distribution
- allow easy comparison between distributions, even where the mean changes and the curve shifts up and down
- allow a clear distinction between working to narrow the distribution, or to shift the curve.

Disadvantages

Histograms:

- do not allow individual organisations (DHBs, practices, etc) to see easily where they are in the distribution
- do not show how an individual relates to a reference or average
- can take more effort to calculate and present.

Funnel plots

Funnel plots explicitly look at how much variation is expected from a given size of population. Effectively, a funnel plot looks at the rate of something, and works out how much variation at 2 or 3 standard deviations is to be expected, given the size of the underlying populations. The example in Figure 8 shows the same

information as in the previous two graphs, plotting the rate of gout hospitalisations against the size of DHB populations (aged 20–79). It places two lines above and below the national average, at two and three standard deviations respectively, showing how much variation you might expect from DHBs with given population size. In effect this is another representation of the bar chart, showing confidence intervals around the national mean.

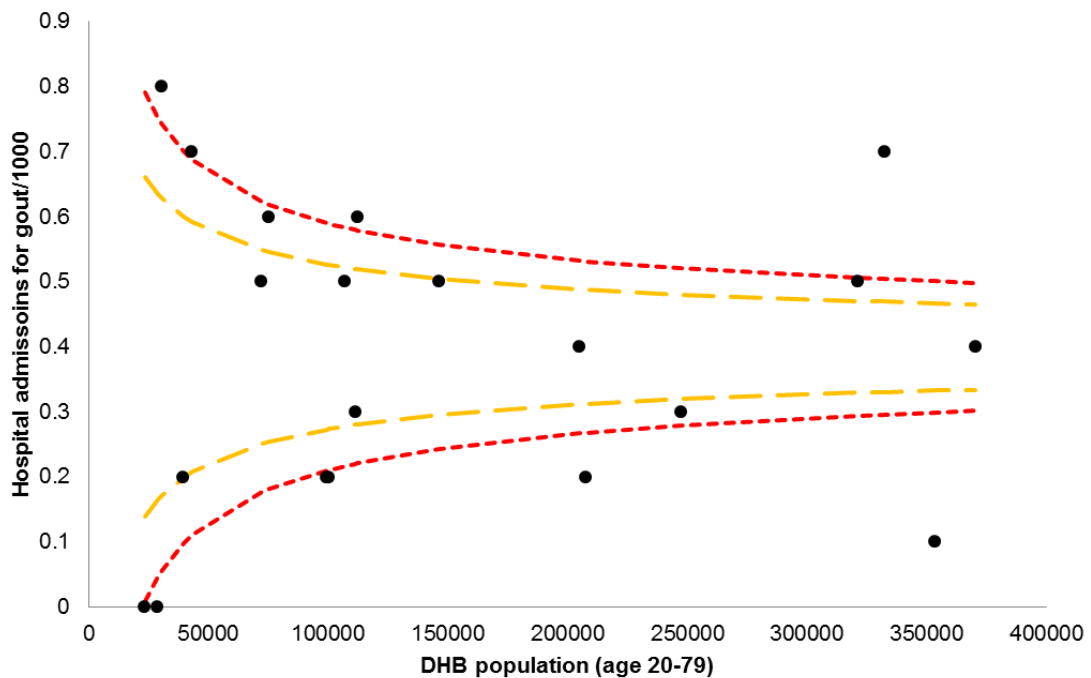


Figure 8: Hospital admissions for gout/1000 population

Advantages

Funnel plots:

- take account of the statistics of small numbers
- allow a quick assessment of the magnitude of variation
- can allow easy comparison between individual data points.

Disadvantages

Funnel plots:

- can sometimes be fiddly and time-consuming to prepare
- don't entirely answer the question of how much variation should be expected
- aren't always straightforward to interpret for a general audience.

Broadly, each of these main ways of presenting variation data has a place, but it is important to bear in mind that they each have limitations for different purposes. In particular, the visual properties of a ranked bar graph can tend to focus attention upon outliers in the range, rather than upon the shape of the whole curve.

Resource

General principles for good data presentation, from the University of Reading Statistical Services Centre:
<http://www.reading.ac.uk/ssc/n/resources/Docs/PresentationOfTablesGraphsAndStatistics.pdf>

3. Addressing the key questions

It is important to focus quality improvement efforts where they will make the most difference, and to make the best judgement about the nature of a given example of variation possible with the available data. Before trying to address variation, it is therefore important to think about what it means and why it matters in any one particular case. Systematically considering these questions can help in deciding how to interpret observations of variation.

Is there uncertainty or ambiguity in the clinical evidence?

Wide variation can be an indicator that clinical evidence is ambiguous or difficult to apply, or that there is a change in practice which is spreading throughout a group of professionals. This raises one of the fundamental issues in interpreting variation: across an observed range of practice, it can sometimes be very difficult to identify the preferred rate of intervention or service delivery at a population level.

This issue is also connected to Wennberg's approach of defining preference-sensitive care. Where there is ambiguity about the cost and benefit of a procedure for a given patient (eg, in some screening applications), then it might be expected that a wide variation in patient preferences could result in a wide variation in delivery of the service. In this case, variation in service delivery might suggest that the underlying cause is variable patient information and support for patient decision-making, and that developing effective patient decision aids could be an appropriate response to managing the variation (Wennberg 2010).

Are there quality issues?

Observing variation can raise questions about the underlying quality of care, particularly if there is clear evidence about where an intervention should be used, and what benefits and side effects exist. Where evidence is less clear, then issues of quality associated with variation are likely to be more difficult to pin down, with different points of view among professionals about what constitutes best practice. For example, different clinicians may have quite different thresholds for prescribing selective serotonin reuptake inhibitors, but in the absence of guidance, it is hard to argue that variation indicates a problem with quality, unless it is very extreme. But where there is clear clinical evidence (eg, for prescribing beta blockers and statins in secondary prevention of CVD), then it is much easier to interpret variation as a marker of issues with quality of care.

Where variation is suspected to be a sign of a quality problem, the challenge is typically to seek to narrow the curve, bringing practice within a closer range of consensus (Figure 9).

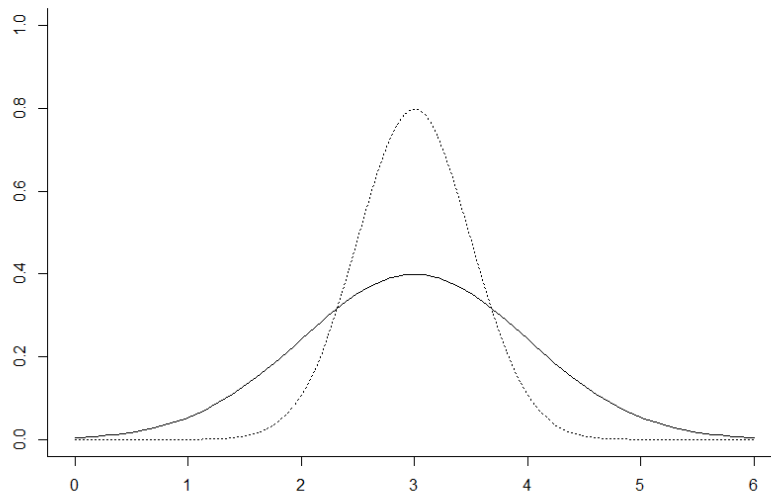


Figure 9: Narrowing a wide distribution

Is there inequity for patients?

Variation can be a sign of inequity in the delivery of health services. This can operate in two different ways:

- *inequity in resource allocation.* Variation can reflect differences in the resource devoted to different populations, and can potentially reveal differences in fairness or equity of resource allocation, or underlying differences in which resources are used across different services. For example, a high intervention rate might reflect supplier-induced demand, where a high rate of supply means a high rate of delivery (eg, more private hospitals means more private surgery, even where need across populations is equal).
- *inequity of access.* Inequity of access can be related to inequity of resources, but can also be a wider issue of access on the basis of timing of service delivery, financial barriers, workforce mix and availability, and other factors.

In a New Zealand context this can be complex, since different DHBs can, and should, make different resource allocation decisions for their populations as need and priorities vary around the country. But if there are differences in the rate at which patients are receiving services, it is important to ask whether this indicates some form of unfairness or inequity in health care from a patient or community point of view.

Inequity can manifest as a difference in distributions for different demographic groups (eg, in men and women for CVD treatment, as seen in the HQSC Atlas),

in which case the challenge is to merge two distributions into a single curve (Figure 10).

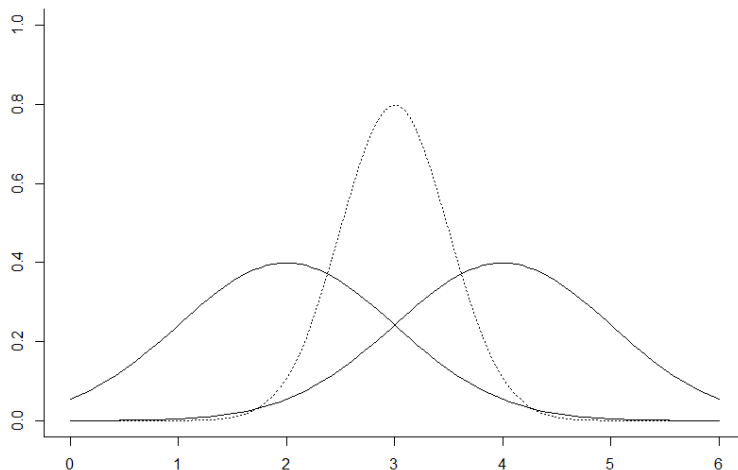


Figure 10: Bringing two different distributions together

Is there inefficient use of resources?

A common question raised by observing variation is whether resources are being used efficiently and effectively. Typically, if two areas have very different rates of intervention, then either under-servicing or over-servicing might be suspected in one of the areas. This is often a difficult question to disentangle from observations of variation, since greater resource devoted to a service on a population basis (eg, higher prescribing of analgesics) might actually reflect a lack of resource elsewhere in the health system (eg, waiting times for surgery). Observing variation may genuinely indicate an inefficient use of resources, but establishing that this is the case, as opposed to just a different mix of resource use, can be complex.

In this case, the goal is likely to be to shift the whole curve, as well as to narrow it, either by ensuring that at least a minimum level of resource is delivered (if the problem is under-servicing), or setting limits to the maximum amount of service which the system can resource.

For example, under-servicing might mean that the goal is to shift the curve to the right, narrow the variation and set a minimum level of intervention (Figure 11).

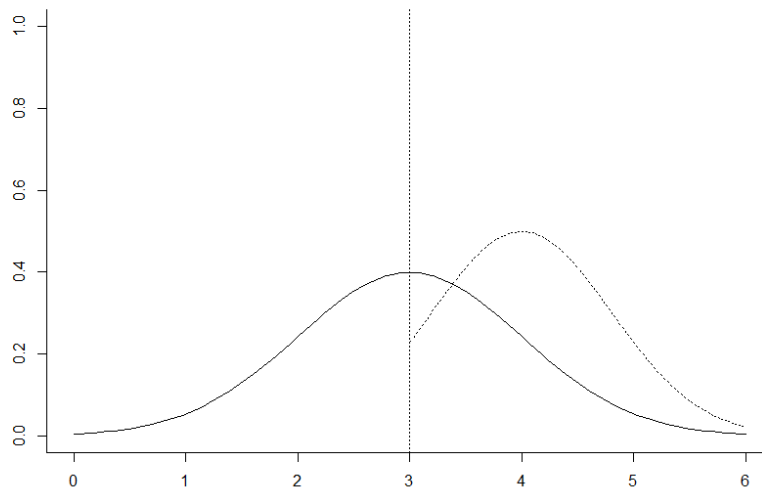


Figure 11: Addressing under-servicing

If there is an issue of over-servicing, the challenge will be to move the curve to the left, and to narrow the variation. If over-servicing and under-servicing both exist at once, then the issue reduces to the general problem of improving quality, and narrowing the distribution.

More broadly, efficient use of resources is an issue of understanding the cost-effectiveness of an intervention, and combining cost-effectiveness analysis, local epidemiological data and clinical judgement to achieve an understanding of what constitutes an appropriate level of intervention. Funders and planners have a role within the wider health system of directing resources to the activities which will produce the greatest benefit for their populations. Working with clinical leaders to establish consensus on whether there is under-servicing or over-servicing for specific interventions can play a key role in addressing issues of variation.

What environmental and population factors are relevant?

Variation, especially across areas, can reflect a range of environmental and population factors, as well as variation in individual clinical decision-making. Such factors include population age, sex, ethnicity and deprivation structure, but might also include aspects such as differing levels of hospital service provision, which in turn can impact upon the way that primary care delivers services. Factors such as rurality, or particular health-related risks (such as a dangerous road, or presence of particular industries), could reasonably result in different ways of delivering care and different approaches to practice at the individual clinical level.

New Zealand has a relatively localised health system, with a high level of discretion for DHBs to make decisions in response to different needs and

priorities. So some degree of variation is to be expected, and is built into the structure of the health system. Understanding the role of environmental influences upon clinical practice and the delivery of care can be an important starting point for developing service and quality improvement programmes. This aspect of variation can be an important catalyst for thinking about what influences clinical practice in a given area, and whether those influences are appropriate for the needs of the population.

How much variation should you expect?

Variation is often seen as a surprise, but many things can drive variability, including the random statistics of small numbers, differences in training, and differences in patient expectation. Before deciding whether variation is a problem it can be helpful to step back and assess how much variation you should actually expect to see.

Considering how much variation to expect can involve:

- reviewing the clarity of evidence for clinical decision-making
- assessing whether small numbers and statistical effects generate variation
- considering environmental drivers of variation across different areas.

Some researchers have tried to understand how much variation is typical, and how much exceptional, at the level of general practices, with the aim of improving power calculations for studies based in primary care (Adams et al 2004; Turner et al 2001). Although their purpose is not to understand variation per se, their results cast useful light upon typical levels of variability. One interpretation of their results is that if more than 1 percent of variation in some measure is associated with the general practice a patient is enrolled in (as opposed to patient factors such as age or ethnicity), then there should be a suspicion of practice level variation. If more than 5 percent of variation is attributable to the general practice, then there is probably substantial practice level variation, and this is worth further analysis.

In the end, the question of how much variation is too much is a judgement, and pragmatically should depend upon what is believed about the impact of variation, and whether there are ways of addressing it. If there is substantial variation in some measure, but it does not provoke concerns about quality, equity or efficiency, then it may not be worth devoting precious quality-improvement resources to addressing the issue. By contrast, if even a relatively small level of variation raises serious issues about quality of care, then it is worthy of attention and follow-up.

4. What to do about variation

4.1 Need to manage variation

The challenges of quality, equity and efficiency are all, in themselves, good reasons for health service organisations to use the tools available to them for addressing variation. The direction set by the Ministry of Health Expert Advisory Group explicitly envisages a focus on variation at a local level, with PHOs and DHBs taking responsibility for managing resources more effectively, and for using clinical governance mechanisms to reduce variability. This is likely to take place at two levels: variation on national system level indicators will be monitored centrally, while local alliances will be likely to have a high degree of freedom to identify and address quality and resource issues within each district across New Zealand. Reducing variation in clinical outcomes and in effective resource use is likely to be an increasingly explicit element of performance monitoring in New Zealand health care. There will be increasing expectation that planners and funders will work collaboratively with service delivery organisations to monitor and manage variation (Expert Advisory Group 2014).

4.2 Approaches for managing variation

At a high level, Wennberg (2010) recommends adopting three key strategies for addressing the challenges of variation:

- promoting organised systems of health care delivery to prevent underuse of effective care. Team medicine seems to lead to less unwarranted variation.
- establishing informed patient choice as the ethical and legal standard for decisions surrounding elective surgeries, drugs, tests, and procedures, and care at the end of life. In terms of preference-sensitive care, treating patients according to their preferences – and not giving them treatments they do not want – requires a clinical environment that supports shared decision-making and encourages the active engagement of patients in the choice of treatment.
- improving the science of health care delivery. In terms of supply-sensitive care, the most important challenge to the clinical and research communities is to rationalise the clinical pathways for managing chronic disease: to undertake the clinical research required to convert supply-sensitive care into evidence-based care that is effective or preference-sensitive.

The generic approach is therefore to engage in effective service improvement and clinical governance as the mechanism for addressing the fundamental issues raised by variation. Achieving this requires effective, expert clinical leadership

and high quality analysis, conducted by clinicians and analysts working closely together.

The toolbox for addressing variation includes the following five key elements.

Good analysis (but don't get hung up on perfection)

Good quantitative analysis, informed by strong clinical expertise, is essential. It is important for analysts to bring technical expertise about datasets, data validity and statistical techniques, and to work closely with clinicians who can refine questions about what is going on in the data, and understand the significance of different observations. Ideally analysts and clinicians should work together, interrogating data and cutting it in different ways, coming to a joint interpretation of the observed variation which is both statistically and clinically robust.

The HQSC Atlas provides data at DHB level on a number of quality indicators. This information is presented in context with a summary of relevant evidence for each indicator, and with the definitions and methodology used in the underlying analysis (circled in the Atlas page example in Figure 12). This resource represents a flying start for analysis on these indicators, since the complex job of assessing the data and developing workable definitions that are comparable across areas has already been completed. For example, working definitions of polypharmacy are provided in the background material linked to HQSC Atlas pages, and these can serve as a guide for analysts to extract and present data in a standardised and comparable format. Building upon the HQSC infrastructure of definition, and the detailed work done in developing the Atlas, which has involved careful thought about how best to define and measure meaningful aspects of variation, will give the recipients of variation data more confidence in the information they are given.

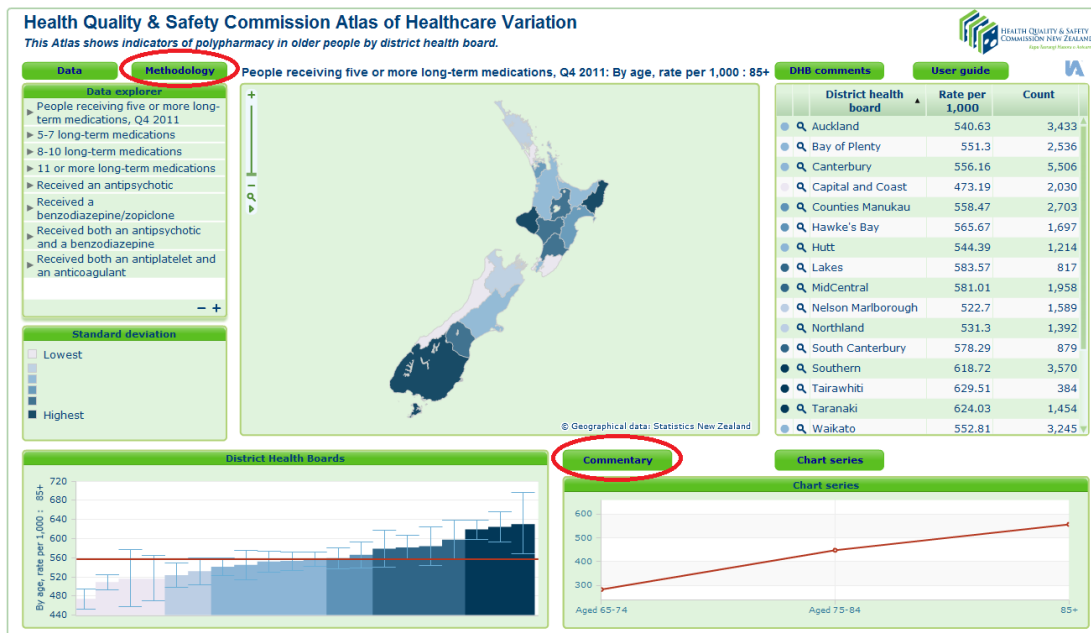


Figure 12: HQSC Atlas of Healthcare Variation

While it is important to ensure that analysis is robust, and to avoid being trapped into misleading conclusions about variability, it is also important to be realistic about what information can be collected and analysed, and how much can be interpreted from that information. It is nearly always possible to conclude something robust and defensible, so long as data interpretation is performed carefully, is informed by both statistical and clinical expertise, and the limitations of interpreting data are clearly acknowledged. Typically, the HQSC Atlas definitions of patient populations are the result of extensive clinical and expert debate on the best way to define and measure the populations and interventions involved.

Patient expectations and information

Where evidence is unclear, or where there are choices and trade-offs to be made about treatment, then ensuring that patients are well informed and able to make their own choices is a good response to observing variation. This is particularly the case where services are preference-sensitive, and patients need support and information from health professionals in order to weigh up the advantages and disadvantages of treatment.

Good examples of areas where improved patient information and shared decision-making can reduce variation include decisions about surgery, and about screening programmes and their consequences. Examples in New Zealand include patient information campaigns on the effective use of antibiotics, and

guidance to support general practitioners to inform patients about the value of prostate testing.

Patient expectations are a very important driver of clinical decision-making, and it is important to consider variation in the context of both patient and clinician information and preferences.

Work with practitioners at extremes of variation

It is important to be clear that a practitioner who is at the extreme end of some measure of service may be there for a good reason, or even purely by chance – particularly if they see a small number of patients, or have a patient population with special characteristics. If there are practitioners who appear to be well outside the usual range, then they should be provided with comparative information which informs them about their practice (which they may be unaware of), and a professional peer could be assigned to work with them in order to understand whether their practice genuinely is different, whether there is a need for change, and what support the practitioner needs.

Some organisations may be tempted to take a punitive attitude towards professionals who are outliers on service measures, but a measured approach based upon good information and peer expertise is often more likely to result in service improvement.

Using and disseminating clinical information

Good analysis is important, but it depends upon the availability of good information in the first place, and upon effective use of the information with clinicians. The key elements here are:

- identifying strengths and weaknesses in information at practice level, and providing support so that information can be continuously improved. For example, some practices and PHOs adopt specific coding schemes, and generate their own codes for activities which they want to collect robust information about.
- providing tools for clinicians within practices to interrogate and compare their own clinical information, allowing for exploration and dissection of data within individual practice teams. Improved information tools can particularly help with effective clinical audit, and with patient reviews.
- providing well analysed material back to clinicians, comparing their activity to their peers within a network or across an area. It is easy for practitioners to become isolated, and to be unaware that there is a range of variation in practice. Providing information on the range of practice, and where individuals lie within that range, is a powerful technique for informing clinicians and providing context for them to reflect upon their own practice.

Current practice in providing feedback varies considerably, with different approaches adopted in different parts of New Zealand. Some PHOs provide anonymous feedback to general practitioners, who know only their own level within the overall range. Others provide named, open feedback to all professionals, so everybody has the same view of the information. In some cases practitioners receive information which compares them with a small group of peers, while in others they may receive information which compares them to a whole network, or across a whole geographic area.

The important principles are that the collection and dissemination of information should be friendly to clinicians, providing clinicians with the tools and support they need to answer questions about practice, and to provide information and evidence in an accessible, useable fashion. These are core clinical governance functions of PHOs and networks, although the degree to which they are delivered can be variable across New Zealand.

Consensus management and service design

Where evidence is ambiguous or unclear, there may be benefit in attempting to establish local consensus on disease management, based upon evidence and international best practice, and acknowledging the role of varying patient preference. While it may be reasonable, in the absence of clear evidence, to have a range of practice, if the result is inequity in treatment across the population, or a mismatch of clinical resources which could be used more effectively, then it may be important to establish more standardised approaches to care. For the topics covered in the HQSC Atlas, links are provided to evidence on best practice and patient pathways.

Examples of consensus patient management built into service design include standardised referral or post-discharge protocols. Increasingly, in New Zealand, a range of disease management programmes provide resources more consistently to clinicians, and potentially reduce variation in the management of key conditions. A number of PHOs and DHBs across New Zealand have established specific disease management programmes, such as for chronic obstructive pulmonary disease or diabetes, while broader systems design, such as post-discharge planning and shared clinical information systems provide an environment in which clinical decisions are made with more information, for both patients and clinicians. Providing ready access to specialist opinion when needed, with clarity about the information needed with a referral, can support primary care professionals and reduce variation in the use of diagnostics and referral patterns.

5. Case study: polypharmacy in the elderly

5.1 Background

Polypharmacy is defined in various ways, including: the use of multiple medications or medications which are not indicated; inappropriate or unnecessary drug use; use of many or more concurrent medicines; mismatch between medicine and diagnosis; or potentially inappropriate prescribing. Where many definitions are possible, it is important to identify the definition which is most meaningful in the local context, and to the clinicians who will be involved in the discussion about addressing variation.

Since elderly people have a higher rate of comorbid conditions than younger people, applying routine treatment guidelines can easily lead to polypharmacy. This can bring risks of adverse effects and interactions, non-adherence, medication errors, falls, hospital admission and increased mortality. Polypharmacy raises difficult issues about the application of clinical evidence to the case of complex patients with multiple comorbidities.

Atlas information

The HQSC Atlas provides several views of polypharmacy data. It provides data for people aged 65 and older, and specific measures for:

- people receiving five or more long-term medications
- people receiving five to seven long-term medications
- people receiving eight to ten long-term medications
- people receiving 11 or more long-term medications.

There are particular issues of risk, especially for falls, where sedatives and other psychoactive medicines are involved. Combinations of other medicines with anticoagulants can also flag clinical risk. The HQSC Atlas therefore provides specific cuts of data for people who receive:

- an antipsychotic
- benzodiazepine/zopiclone
- both an antipsychotic and a benzodiazepine
- both an antiplatelet and an anticoagulant.

Atlas results for Canterbury

In 2009, the HQSC Atlas shows that moderate polypharmacy (5 to 7 long-term medicines) for people aged over 64 in Canterbury is not higher than across the rest of New Zealand, but that Canterbury is significantly higher than the rest of New Zealand in the number of people aged over 64 with 8 to 10, and 11 or more, medications.

Canterbury is shown as purple in Figures 13–18 and Figure 20 below.

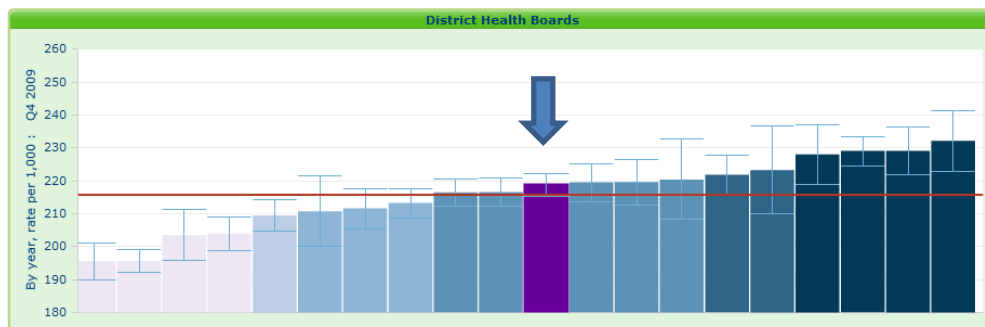


Figure 13: Canterbury – 5 to 7 long-term medications 2009

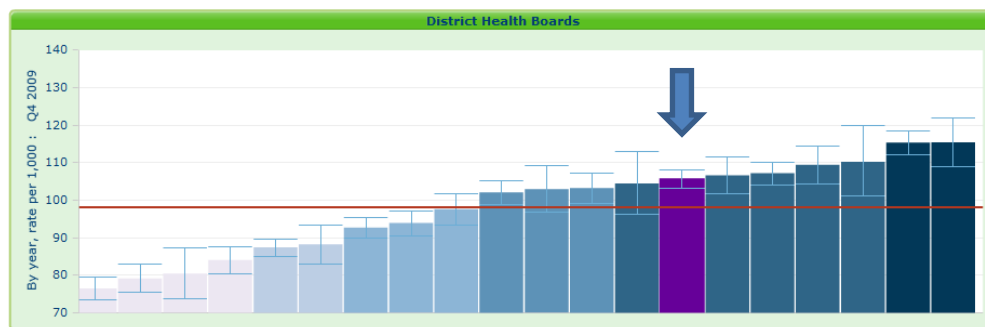


Figure 14: Canterbury – 8 to 10 long-term medications 2009

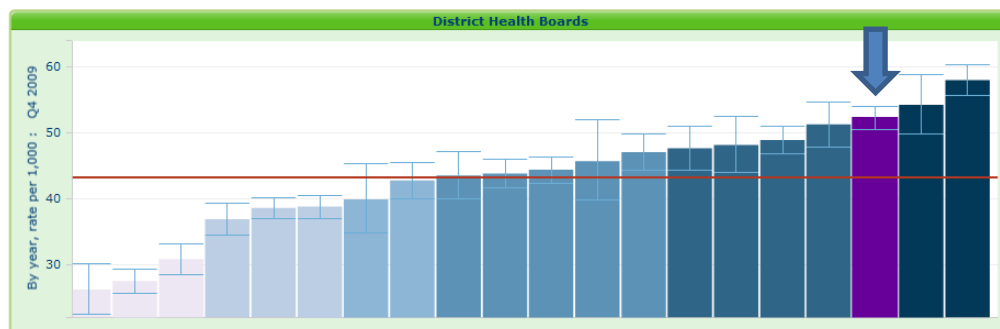


Figure 15: Canterbury – 11 or more long-term medications 2009

In 2009 Canterbury had the highest rate of antipsychotic prescribing for the elderly in New Zealand, by a substantial margin.

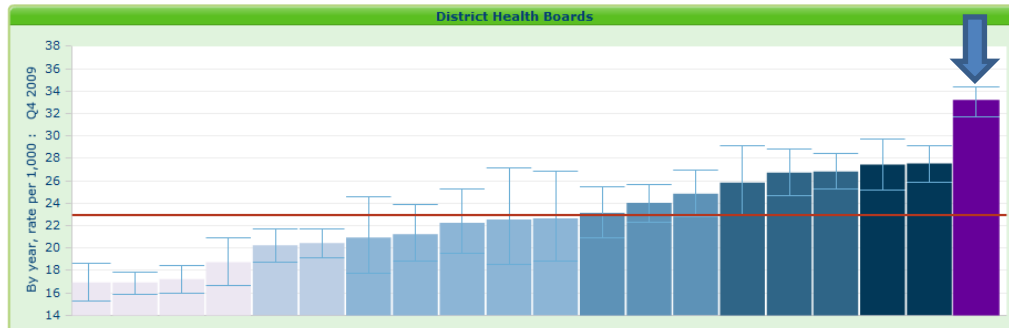
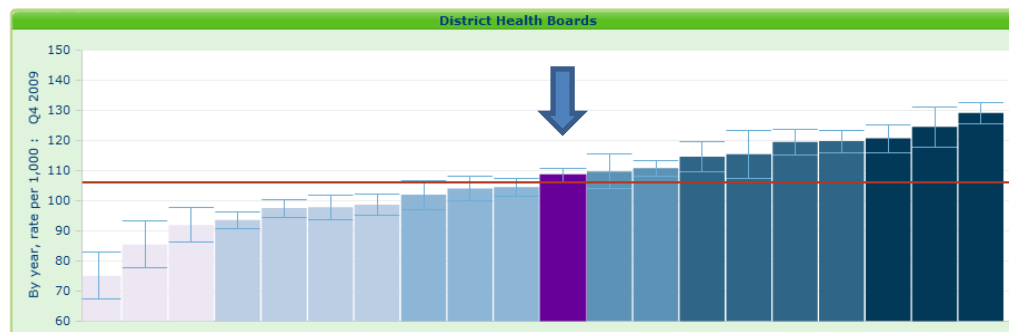


Figure 16: Antipsychotics for people aged 65 and over 2009

By contrast the rate of benzodiazepines for people aged over 64 was not significantly higher in Canterbury than the New Zealand average.



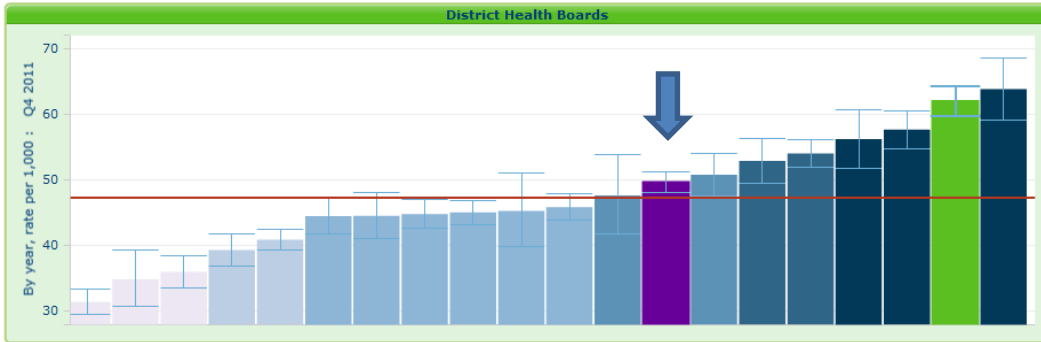


Figure 18: Canterbury – 11 or more long-term medications 2011

This result is in part due to the trend in polypharmacy declining slightly in Canterbury, while increasing across the rest of New Zealand. Canterbury clearly has a different polypharmacy trend from the rest of the country over this time period in the high polypharmacy group, where patients have 11 or more long-term medicines.



Figure 19: Patients aged 65+ with 11 or more medications 2009–2011

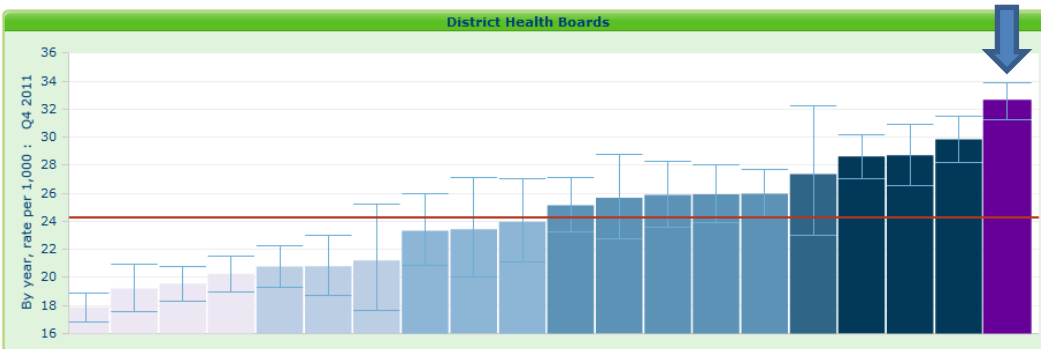


Figure 20: Antipsychotics for people aged 65 and over 2011

While Canterbury is high for antipsychotics across all age bands for the elderly, this is particularly the case for the very elderly, aged 85 and over.

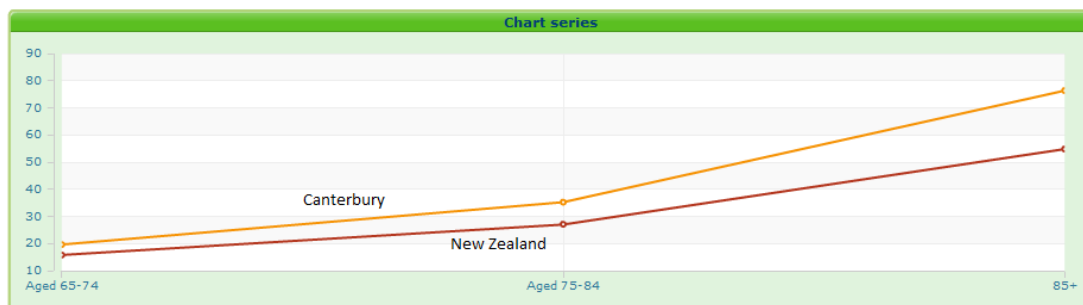


Figure 21: Antipsychotics by age band

5.2 Review

Are the data complete and accurate?

The data used to analyse variation are very complete and accurate in terms of dispensed prescription medicines, although prescribed (but not dispensed) and over-the-counter medication are not included in the data. The national dataset on dispensing has the advantage of being nationally validated and regularly updated. The challenge is that these data cannot easily be linked to patient denominator information within a PHO, making analysis at the level of individual prescriber very difficult, if not impossible, to conduct robustly. This is particularly the case in polypharmacy, because it is inherent in the nature of the problem that there will often be many prescribers involved in the care of the patient, so attributing prescribing decisions to any one practitioner is difficult to do in a robust fashion. In using the HQSC Atlas data as a starting point for looking at polypharmacy in Canterbury, the analysis therefore focuses upon describing patterns of polypharmacy across the district, and is complemented by providing query files to general practitioners to facilitate their analysis of data from their own practice systems.

The difficulty with facilitating analysis of data from practice management systems is that, since it is clear that many different prescribers are commonly involved in the care of this patient group, a practice management system may only capture part of the prescribing picture for a given patient. In this case the data, while readily accessible to the general practitioner, may be incomplete, and provide limited support for reviewing patient management.

The answers to these challenges involve working to improve information at two levels. The first is to continue the implementation of shared care records, allowing general practitioners, pharmacists and hospital clinicians to share information for the care of individual patients. In the case of Canterbury, electronic shared care records are well down the path of implementation, and are supporting clinicians to have more complete information about the care provided to patients under their management.

The second challenge is to improve the management of national pharmaceutical datasets so that they can be linked, securely and with appropriate restrictions upon access and use, to denominator data from patient registers. From the perspective of analysing variation in prescribing within a district, this allows significantly more valuable analysis to be performed in a robust fashion, and information to be provided to clinicians about their own practice.

Is there uncertainty or ambiguity in the clinical evidence?

There is both clarity and uncertainty in clinical evidence around polypharmacy, summarised in the HQSC Atlas. Evidence of the downsides of polypharmacy is clear, both in New Zealand and international data, with increased medicine interactions, and hospitalisations (Davis et al 2002). Polypharmacy is clearly associated with increased risk of side effects, interactions, adverse events and hospital admissions (Patterson et al 2012). There is a lack of evidence for the use of antipsychotics in the elderly.

At another level, evidence about polypharmacy in disease management is often unclear, or the applicability of evidence is uncertain. If practice guidelines focus on advice for the management of patients with a single disease, the prescriber may be left with considerable uncertainty about the best management of a patient with complex comorbidities. Polypharmacy for the elderly, then, is a phenomenon which emerges in the presence of considerable uncertainty at the individual level about the clinical best practice, and there is scope to provide advice and support to prescribers who work within this uncertain environment.

Are there quality issues?

Variation in polypharmacy primarily raises questions about quality of care. There is extensive material which can be provided to local prescribers about adverse events, and about addressing polypharmacy for elderly people (eg, Patterson et al 2012), while New Zealand studies have been conducted which document the level of hospital admission which can be attributed to medication error, strongly associated with polypharmacy (Davis et al 2002).

Is there inequity for patients?

Polypharmacy is not an equity issue in itself, although the consequences of polypharmacy may fall out differently for patients in different categories of age, sex and ethnicity. Polypharmacy may need to be balanced against the potential for unmet need for medicines for some population groups.

Is there inefficient use of resources?

Most variation in polypharmacy does not directly raise issues about the inefficient use of health system resources, although where the consequences of polypharmacy involve pharmacists in addressing complex medication issues, or

avoidable hospital activity as a consequence of adverse events, polypharmacy may be associated with avoidable use of health care resources.

Where polypharmacy variation may raise a more direct question about the efficient use of health care resources is in the use of antipsychotics. The antipsychotics olanzapine, quetiapine and risperidone accounted for \$2.9 million in medicine cost for the Canterbury health system in the 2009 calendar year, with prescribing analysis suggesting that low-dose off-label prescribing is likely to account for a significant element of this, particularly with quetiapine. There may well be potential to redirect some of this health care resource to other more beneficial uses.

What environmental and population factors are relevant?

There are many complex environmental factors which could increase or decrease the prevalence of polypharmacy in a given area, or for a given prescriber. In Canterbury these include:

- information systems, and access to shared care records, allowing clinicians to have a comprehensive view of care provided to individual patients. These have been widely implemented in Canterbury since 2010.
- the presence of multidisciplinary medicines review programmes, such as medicines therapy assessment or medicines management programmes. Medicines management has been implemented in Canterbury, with a large number of patients having their medication reviewed by a community pharmacist.
- hospital discharge policies, and secondary care clinical practices
- the general health of the population. In an area of high morbidity, a greater degree of polypharmacy might be expected. While Canterbury has a large population of elderly people, there is no specific reason to think that they have a higher general level of morbidity than other parts of New Zealand.
- the number of psycho-geriatric beds in the district, which is high in Canterbury
- the Canterbury earthquakes. These have likely had an adverse impact on the physical and mental health of the population, potentially creating marked increase in health need for some population groups. Evidence from other natural disasters finds a particular impact on mental health across the population, which anecdotally is supported by the experience in Canterbury.

Each of these factors can have a strong influence upon the environment in which individual primary care practitioners manage patients with complex comorbidity. System-level structures, such as shared information systems and medicines review programmes, have the potential to make a marked difference to the

information available to prescribers, and to support the decisions they make with individual patients.

The other important system-level factor arises from hospital practice. As with many aspects of care, particular practices or procedures in a hospital can affect a whole population, across many different primary care services. In the case of Canterbury, hospital services have engaged in a specific programme for reviewing polypharmacy upon discharge (known as the Pill Pruning Project), which in this case is complementary to education programmes delivered to general practitioners.

How much variation should you expect?

At the inter-district level it might be expected that there would be substantial variation in polypharmacy, since there are a number of environmental or system-level factors which are likely to vary across districts. But within a district the numbers of elderly receiving multiple medications is large, and the effect of random variation between individual prescribers is likely to be small. It could be expected that within a district there should be relatively little polypharmacy variation, although measuring this robustly is complex because of the difficulty of linking patients to prescribers in a simple and robust fashion. If there is substantial variation between practitioners this is likely to represent an issue of quality of care, and it could reasonably be expected that this should narrow with improved clinical guidance and information systems.

5.3 Analysis

The data issues in analysing polypharmacy are complex. The dataset available to Pegasus comes from the national pharmaceutical data warehouse, available to most DHBs and many PHOs and networks, and is the same basic source of information used by HQSC to develop the Atlas. This dataset is comprehensive, very complete and well validated.

The major challenge with the data is that, while prescribers are identified by medical council number, patient identifiers are encrypted. This means that patient information cannot be linked to other data, such as PHO registers. Consequently, attributing the prescribing for any one patient to a particular prescriber is difficult. For example, in the last quarter of 2009, 3862 people were prescribed 11 or more continued medicines, using a methodology similar to that in the HQSC Atlas. These patients had an average of 3.6 different prescribers involved in their care over a six-month period. Only 16 percent of these patients were prescribed medicines by a single doctor in the six-month period.

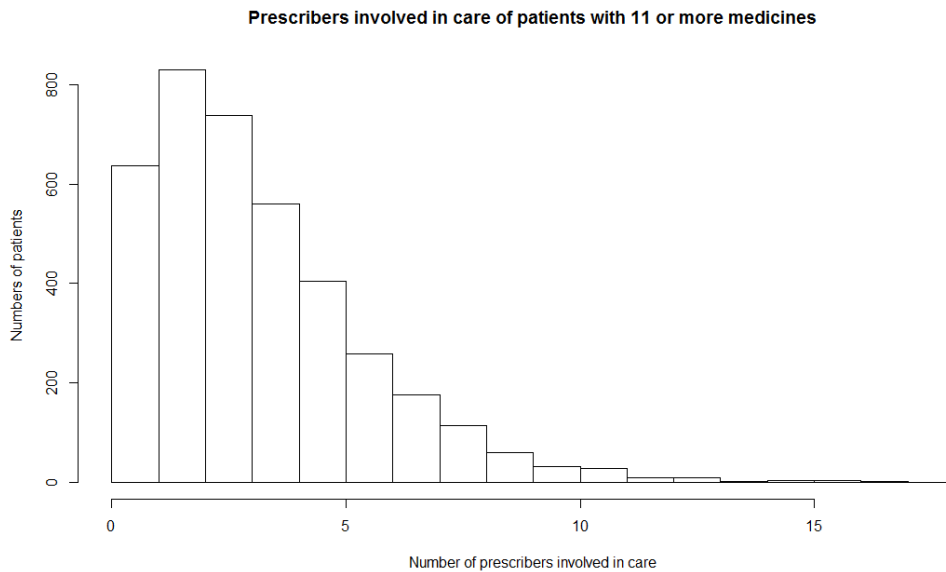


Figure 22: Number of prescribers per patient

Although in principle a patient’s general practitioner should have an overall coordinating role in prescribing medication for an individual person, and reviewing prescribing post hospital discharge is an important task for a general practitioner, the data do not support attribution of an individual’s care to registration with a specific general practitioner. This means firstly that analysis of variation within the district is very difficult, and secondly that it is not possible to use these data to feed back polypharmacy information to individual prescribers with a high degree of robustness.

The strategy to address this is twofold:

- Firstly, use the pharmaceutical warehouse data to describe the general characteristics of polypharmacy in this population.
- Secondly, provide prescribers with tools to analyse their own information from practice data.

At Therapeutic Group level 2, the most common categories of medicines for patients receiving 5–7 medicines are shown in Table 1.

Therapeutic Group 2	Patients receiving medicine
Analgesics	12,918
Antithrombotic Agents	10,085
Lipid-Modifying Agents	7738
Agents Affecting the Renin-Angiotensin	7395
Beta Adrenoceptor Blockers	6769
Antiulcerants	5163
Diuretics	5077
Vitamins	4625
Calcium Channel Blockers	4251
Minerals	3248
Antidepressants	2946
Diabetes	2455
Laxatives	1714
Drugs Affecting Bone Metabolism	1662
Thyroid and Antithyroid Agents	1561
Alpha Adrenoceptor Blockers	1532
Sedatives and Hypnotics	1434
Non-Steroidal Anti-Inflammatory Drugs	1177
Nasal Preparations	1058
Nitrates	1027

Table 1: Top 20 medicines for patients receiving 5–7 medicines 2009

The contents and rank order of medicines in this table is similar for patients receiving a larger number of medicines, and across the two periods for which data were available. It is unsurprising that the core medicines are those associated with heart disease and secondary prevention, with laxatives, diabetes, antidepressants, thyroid and sedative agents entering the list lower down.

Using an approximate population denominator, the graph in Figure 23 shows the rate of polypharmacy for the two years by age band. The overall pattern appears to be that higher levels of polypharmacy have remained fairly stable, while there has been some increase in moderate polypharmacy (5–7 medicines) for the very elderly.

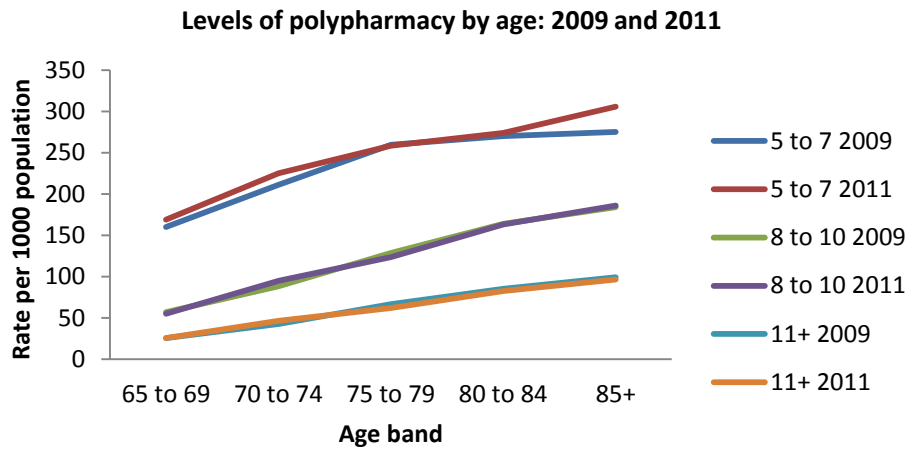


Figure 23: Polypharmacy trends in Canterbury

Canterbury shows a high proportion of antipsychotic prescribing for the elderly. In this dataset, 1435 patients received olanzapine, quetiapine or risperidone in Q4 2009, which had increased to 1512 patients in Q4 2011. Patients receiving these medications were on average receiving 7.5 different medicines in 2009, which had increased slightly to an average of 7.6 medicines in 2011.

Comparing the age profile of patients receiving these three antipsychotics between the two years, Canterbury appears to have reduced the proportion of very elderly receiving these medicines, but has seen an accompanying increase in their use in people aged 65–69. This younger group are unlikely to have their medications strongly influenced by psycho-geriatric facilities.

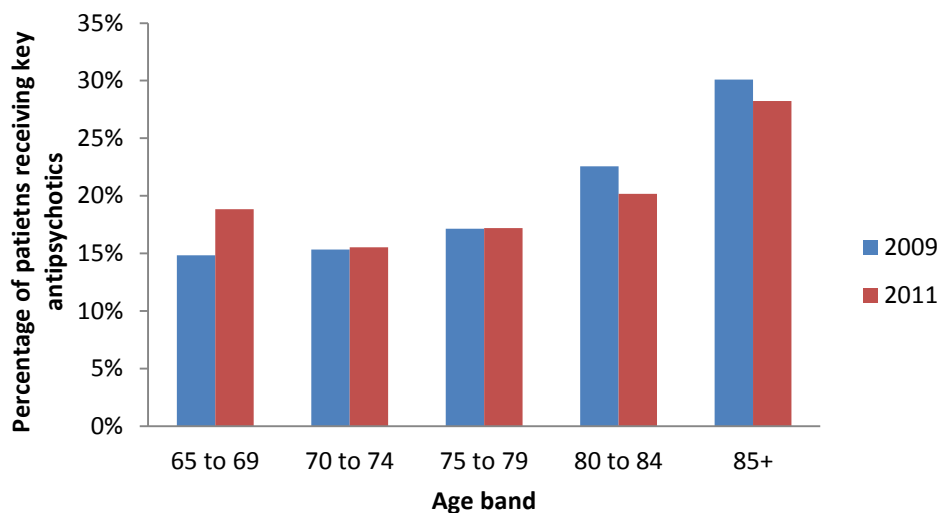


Figure 24: Age profile for key antipsychotics

5.4 Actions

Given the analysis of patterns of polypharmacy in Canterbury, and the review of causes and impacts of variation, the following quality improvement approaches are likely to be worthwhile and effective. These approaches have been undertaken in Canterbury in various ways since 2009, and in combination appear to have had some impact upon polypharmacy for the very elderly, although not in the younger old group, aged 65–80.

Approach	Rationale
Guidance for prescribers about reviewing and stopping medication	The major area of uncertainty for prescribers is the application of evidence and guidelines where patients have complex comorbidities. The Pegasus education round in 2009 provided evidence on pharmacokinetics in the elderly; a framework, supported by literature, for decision-making about treatment with the elderly; advice on specific medications which carry heightened risks; and evidence on benefits after discontinuing medications. This directly addresses the major area of uncertainty for prescribers which is likely to lead to polypharmacy. Reviewing and extending this guidance as it applies to a younger group may be appropriate.

Provide specific evidence and guidance on antipsychotic use	An education round on atypical antipsychotics was delivered in 2010, and further work may be needed to monitor and follow up in this area.
Improved sharing of patient information	Polypharmacy touches upon the activity of multiple health professionals. Improving the shared information between health professionals supports prescribers in making informed decisions with patients, and helps to make them aware of the full range of medicines being received by a patient.
Facilitate prescribers' analysis of their own data.	Pegasus provided MedTech queries to help general practitioners identify patients from their own practices who receive multiple long-term medications. Analysis of practice could be further supported by follow-up and structured comparison of data in peer groups.
Work with secondary care clinicians to align primary and secondary care prescribing patterns	Hospital activity is likely to be an important driver of prescribing patterns post discharge, supporting secondary care clinicians with evidence and guidance, as well as designing hospital processes. An important multidisciplinary approach is likely to be needed for the management of psychogeriatric beds, where there may be a particular issue in Canterbury.

As with any analysis of health data, there are many questions which arise and which are worthy of further investigation. Some of these issues are susceptible to further analysis with information which is already available in primary care in Canterbury, while other questions may require further information, or better capacity to link datasets to support robust analysis.

Key questions for further investigation are:

- What is driving the use of antipsychotics in the younger elderly?
- Can prescribing data be attributed accurately to general practice populations?
- If prescribing data could be linked to hospital discharge data, could specific analysis of patterns of medication post discharge be examined and fed back to both primary and secondary clinicians?
- How variable is the uptake of medicines management programmes and the use of multidisciplinary pharmacist input to manage polypharmacy across the district?

5.5 Commentary

This case study of variation in polypharmacy highlights a number of points.

- Because the same data were available, and the HQSC Atlas website provides explicit documentation of the methodology used, reproducing the particular definitions of polypharmacy used in the HQSC Atlas was reasonably straightforward. There were some ambiguities in the documented method, including whether the years used were calendar or financial.
- Because the HQSC Atlas definition of polypharmacy was easily reproduced, various slices of Atlas data for the particular district were a useful starting point, and provided a good overall context for the pattern and trend of variation.
- In this case, using national data to provide direct feedback to individual primary care prescribers is difficult to do robustly. Providing data to general practitioners on their own practice requires working with them to use their own practice management system information.
- It was possible to extend some elements of the analysis to tease out additional components of variation at a local level, although this was limited by the data available.
- In this case, a number of quality improvement initiatives in this area had already been developed and implemented, and may have contributed to the stable level of polypharmacy, compared to the national trend.
- Reviewing polypharmacy for Canterbury reveals a number of environmental drivers of polypharmacy, suggesting that a quality improvement focus should be both upon primary care prescribers and upon system factors, including clinical information systems, hospital prescribing and discharge, and delivering multidisciplinary medicine review programmes.
- Pharmaceutical data are very rich, but they are also complex to analyse. It is important to be able to manage a large dataset with a suitable database programme, and to be meticulous about the logic of calculating patient ages and rates. It is easy for simple mistakes to produce artefactual conclusions, and it is important that analysis be peer reviewed thoroughly before being shared with clinicians.

Overall, even in an area as complex and multifactorial as polypharmacy, where the ability to link data to individual practitioner is limited, a systematic approach suggests a range of quality improvement interventions which can be undertaken within a district. In a complex field such as this it is likely that a number of simultaneous approaches to the issue will be more successful than a single approach, so the combination of working with individual prescribers with system-level approaches such as information sharing, the pharmacist medicines management programme, and hospital “pill pruning project” will all contribute

towards a result. The example in Canterbury, where the trend in polypharmacy has remained steady by comparison with a nationally increasing trajectory of polypharmacy, suggests that these approaches can, in combination, have some impact.

6. Case study: ventilation tubes

6.1 Background

In New Zealand, ventilation tubes (commonly known as grommets) are most commonly used to treat recurrent acute otitis media and otitis media with effusion (OME, or 'glue ear'). Over 90 percent of children in New Zealand are believed to experience an episode of OME before they reach school age. Alternative treatments for these conditions include watchful waiting and antibiotics.

A review of evidence for a New Zealand PHO in 2012 found that while a number of Cochrane reviews had shown that grommets were beneficial in improving hearing over a period of six months post-surgery, longer term outcomes were no better under surgery than without. The conclusion is that in children with OME, the effect of grommets on hearing is small and diminishes after 6 to 9 months, by which time natural resolution also leads to improved hearing in non-surgically treated children. While grommets may reduce recurrent acute otitis media, they also bring complications, including an increased risk of mild tympanosclerosis, and tympanic membrane abnormalities.

Atlas information

The HQSC Atlas provides several views of grommet data for patients aged under 15. The unit of measure is the number of grommet surgeries performed per 1000 population. This is summarised by the rate per year by DHB (for three years, ending June 2012). The data can be broken down by ethnicity and five-year age band.

Atlas results for Waitematā

Between July 2009 and July 2012, Waitematā had an average annual rate of 7.5 grommets per 1000 population (for patients under the age of 15). This is 14 percent higher than the national average of 6.6. The rate in Waitematā increased over the three years, while the national rate remained constant over the time period.

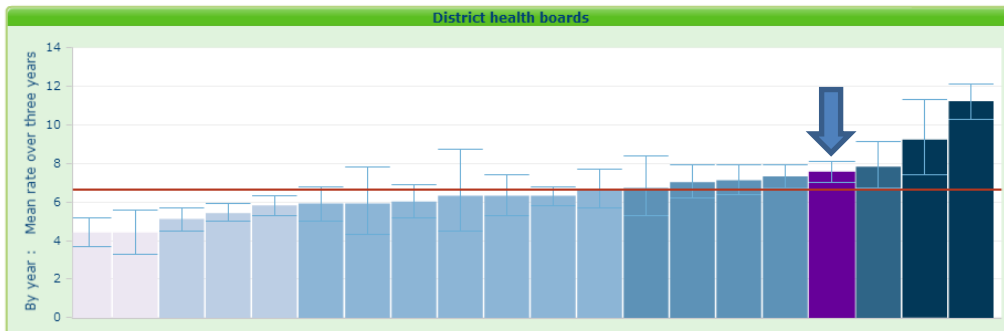


Figure 25: Grommet surgeries per 1000 population by DHB, three-year average June 2009 to July 2012

Across the country the highest rate of grommets is in those aged 0–4 and who are Māori. In Waitematā this difference is greater than the national average. This can be seen in Figures 26 and 27. Waitematā appears to differ from the national average, particularly in the rate of grommet use for younger children, under 5, rather than for children aged 5 to 14 (orange = Waitematā, red = national average).



Figure 26: Grommet surgery by age group 2012

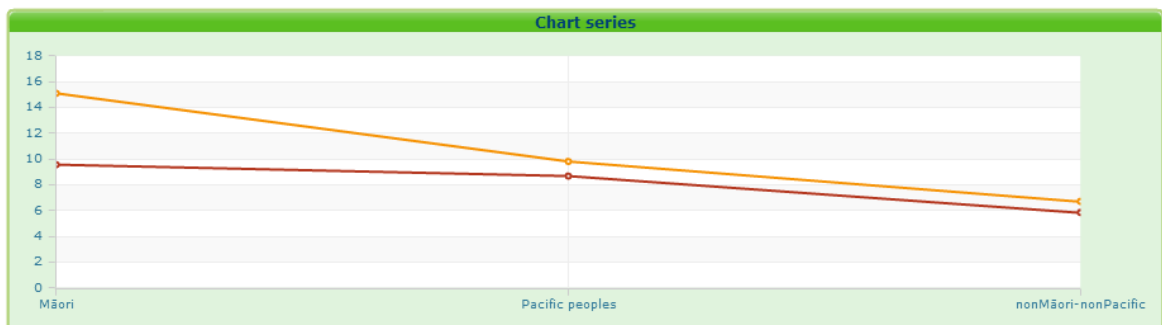


Figure 27: Grommet surgery by ethnic group 2012

The significant difference in grommets for Waitematā therefore appears to be in the high rate which they provide for children under 5, and for Māori children, in particular.

Approximately half of New Zealand’s DHBs have guidelines for referral and for surgery. Waitematā have guidelines for both.

	Referral	Surgery
Waitematā	Yes	Yes
Other DHBs	10 Yes 9 No	9 Yes 10 No

Table 2: OME guidelines in DHBs

6.2 Review

Are the data complete and accurate?

The data in the HQSC Atlas derive from the National Minimum Dataset. While this is therefore complete and accurate on its own terms, there are a number of issues which might limit the conclusions which could be drawn.

- Where there is a vigorous private health care market, as is the case in Waitematā, privately funded surgery will not appear in the National Minimum Dataset. This could have the impact of underestimating the true level of referral for grommets in some areas, and of creating an appearance of variation where none exists. However, in the case of Waitematā, the overall rate is higher than in the rest of New Zealand, especially for younger children, even without considering the additional volume of surgery provided privately. The impact of missing private sector data would therefore be to underestimate Waitematā’s rate of grommet use, making the difference from the national rate even greater than estimated in this dataset.
- The HQSC methodology does not distinguish between grommet procedures for a single ventilation tube, or where two tubes are inserted at once. Different practice across different hospitals could potentially distort comparisons, if some hospitals have a policy of doing only one ear at a time while others will do both.

There is considerable potential for analysing local data, but this would require collaboration between PHOs and the DHB. The best approach would probably be to try to link data, with suitable privacy protocols in place, between the DHB for inpatient and outpatient care, and practice registers, for the denominator of enrolled children. Analyses which could be performed would include:

- general practice rates of grommet insertion
- more detailed analysis of deprivation and ethnicity effects in the rate of grommet insertion, and how they interact
- potentially linking to prescribing data to look at how many episodes of antibiotics have been used in the population of children who have received grommets
- if diagnosis data are available, measuring the proportion of children diagnosed with acute otitis media and OME who receive grommets.

As an initial starting point for locality analysis, it is possible to use national data to examine whether there is variation within Waitematā, albeit on a geographical basis rather than on the basis of enrolment with different practices.

Is there uncertainty or ambiguity in the clinical evidence?

The evidence on benefits and risks from use of grommets is relatively clear. There are several Cochrane reviews showing relatively limited and short-term benefit, and guidelines have been made available in Waitematā. The National Health Committee issued a technology note in January 2013 summarising the evidence for use of grommets, which noted the substantial variation in use of the procedure across DHBs, and that no standard pathway of care exists across DHBs.

Overall, there is relatively little uncertainty in the evidence for the clinical impact and appropriate use of grommets.

Are there quality issues?

Inappropriate use of grommets can risk exposing children to complications such as tympanosclerosis, which may in turn have a moderate impact upon hearing. Use of grommets also exposes children to anaesthetic risk. The quality issue is generally about whether children are exposed to unnecessary care which brings little benefit.

Is there inequity for patients?

At one level there is clearly a difference in the use of grommets between Māori children and those of other ethnicities. The high use among Māori children could reflect a greater prevalence of OME and acute otitis media in this population, or it could reflect later presentation with greater recurrence or more severity, thereby raising the probability of referral. The major question of inequity which variation in grommet use raises is about the effective provision of preventive care for children of Māori ethnicity. The Communicable Disease Centre recommends avoiding second-hand smoke and air pollution, ensuring full immunisation and breastfeeding as preventive measures for paediatric ear infections.

Is there inefficient use of resources?

Elective surgical procedures are in high demand in New Zealand, and the shortage of supply and length of waiting time to receive surgery are prominent and politically controversial issues. If elective surgery is being used inappropriately, or with a low level of benefit, this represents an inefficient use of resources which could potentially be better used for the health of the population. In this case the resources include the specialist surgical workforce, both in assessment consultations and in theatre time, as well as the nursing and allied workforce involved in delivering the services, the physical resources of theatre, and the time and resources of the patient and their family, who may have to set aside time and possibly travel to hospital services. There is considerable scope for variation in use of grommets to suggest inefficient use of health system resources.

What environmental and population factors are relevant?

As well as the individual referral decisions of clinicians in primary care, a number of environmental factors could influence the rate of grommet use.

- *Patient expectations and community experience.* Grommets are a good example of preference-sensitive care, in which the benefits and downsides are not necessarily clear. In this case community expectations of the benefit of grommet use could be an important factor in the overall utilisation of ventilation tubes, and there may be a place for improving consumer information for parents.
- *Surgical practice.* Since hospital services tend to be relatively centralised within a district, the practice of a small number of hospital specialists can influence patterns of care across a whole population.
- *High prevalence rates.* If there is a higher prevalence rate of OME and acute otitis media, then the issue is less one of patterns of health care, and more one of epidemiology and prevention
- *High levels of severity.* If patients present later, having experienced more occurrences of OME and acute otitis media, this may generate a higher level of referral. Higher levels of severity can be related to health-seeking behaviour, which can vary markedly across different cultures, and with barriers to accessing primary care.

Environmental and population factors could potentially be very important in determining local rates of use of grommets.

How much variation should you expect?

While there is variation across New Zealand, at a rate of approximately 10 per 1000 under-5-year-olds, grommets are still a relatively rare procedure in population terms. The statistics of this are robust at district level (the HQSC Atlas

shows clear differences with confidence intervals between DHBs), but analysis may be more difficult within districts, as a relatively unusual procedure with smaller numbers of children could be difficult to compare across practices and small communities. Random variability is likely to be a challenge for analysis. This issue is examined further in the analysis below.

6.3 Analysis

The data used for this analysis were extracted from the National Minimum Dataset, and therefore have the limitation that they do not include information on private sector procedures, nor can they be linked to primary care denominator data. But these limitations do not mean that the information is not useful for drawing conclusions about variation, equity and effective use of public health resources. The data do include more detailed area information, based upon patient residence, which allow some degree of analysis of whether there is variation within the district.

Compared to the rest of New Zealand, Waitematā has:

- approximately half the proportion of Māori people, and nearly twice the proportion of Asian. Since Māori have a high rate of grommets insertion, this would predict a lower overall rate for Waitematā, all things being equal.
- less deprivation (mean 2.7 vs 3.1, on the NZDep06 quintile scale), with only 10 percent of people in the most deprived quintile
- similar sex and age distribution to the national average.

Within the Waitematā district, there are 144 area units, with an average of 250 children under 5 living in each one. In the year ending June 2012 there were 552 ventilation tube surgeries for children in Waitematā. The rate of surgery for children across the area units in Waitematā is shown in Figure 28.

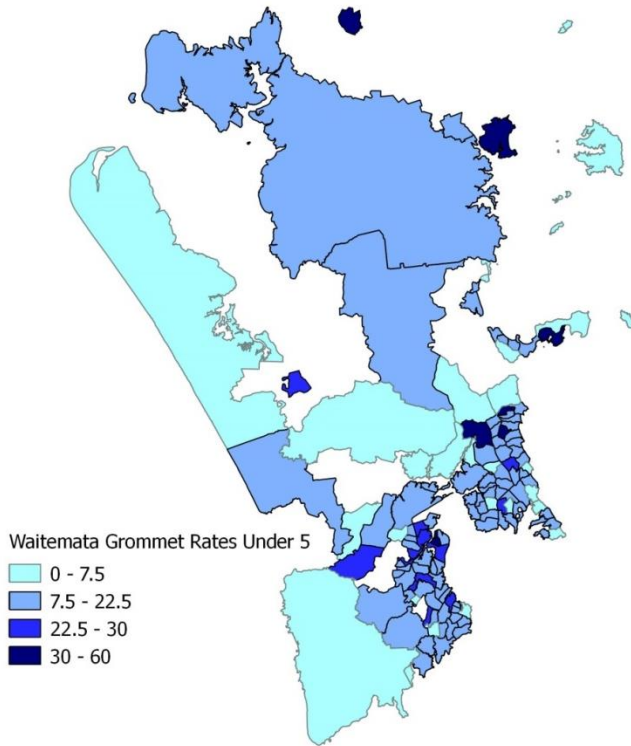


Figure 28: Grommet surgery/1000 children aged 0–4 by census area unit

Rates of grommet insertion per 1000 children under five appear to vary markedly across the district. However, while the mean number of children living in a census area unit is approximately 250, there is a considerable range, and some areas have smaller numbers of young children. A simulation of how much variation is expected across Waitematā area units shows that the expected variation is almost identical to the variation actually observed. While there is a wide distribution with a long tail of higher rates, this is exactly what would be expected on the basis of a relatively rare intervention at population level (see Figure 29 – the coloured bars show the actual distribution, the black line shows the simulated distribution, with 95 percent confidence intervals in broken lines above and below).

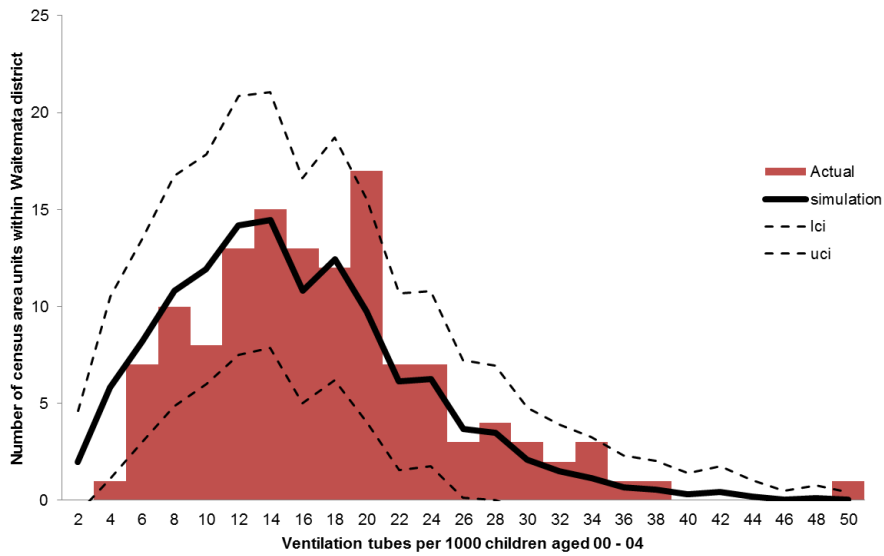


Figure 29: Expected and observed variation in grommet surgery rates by census area unit

Within Waitematā, there was no strong relationship observed between the deprivation of patients and the rate of grommets, although there is a slightly higher rate in children from deprivation quintile four. The interaction between deprivation and ethnicity could be a subject of further analysis at local level.

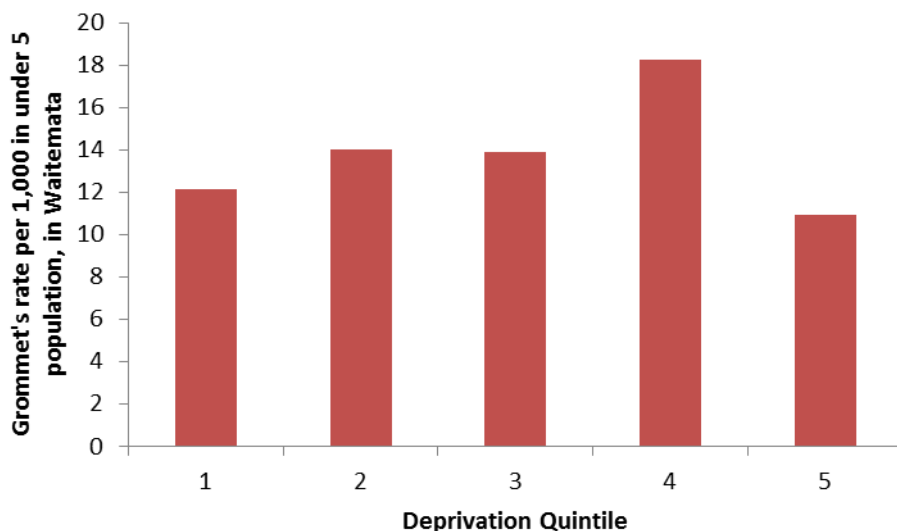


Figure 30: Grommet operations by deprivation category/1000 children 0–4

6.4 Actions

Given this analysis, and what is already known about the issue of ventilation tubes, the following approaches are likely to be worth investigating. These are premised upon the basis that Waitematā has a high rate of grommet use as a district, although there is little evidence of strong variation within the district. An analysis on the basis of enrolled patient denominators might change this view, but on the present information the focus of quality improvement should be much more upon shifting the whole curve of grommet use within Waitematā, rather than focusing on variation across individual referrers and specific practice populations. Essentially, reviewing this variation suggests that the issue is one of significance, with implications for both the quality of care and the efficient use of health resources, and that environmental factors operating across the district are likely to be more important than specific aspects of practice style and difference in clinical decision-making among referrers.

Approach	Rationale
Review patient information resources	If patient demand is a major factor, then resources to support general practice in informing patients about the evidence on benefit and risk of ventilation tubes may help referrers to manage demand, and support referrers to engage in watchful waiting, rather than early referral. There may be potential to work with preschool education facilities and other community organisations to disseminate information to the wider population, and to modify demand for the procedure.

Undertake clinical audit	Support referrers to undertake clinical audit of children referred with grommets, collecting and critiquing information on presentation, recurrence, and management prior to referral.
Undertake work on prevalence of OME and acute otitis media	A possible approach could be to establish a number of pilot practices to code consistently for OME and acute otitis media, allowing accurate assessment of the local burden of disease. While Gribben et al (2012) have analysed the prevalence of acute otitis media nationally, those results do not allow differential analysis by district, and understanding whether Waitematā was different in this respect would help to inform future action. An informed planning and funding role across the health system is likely to seek the best available information on cost-effectiveness and need for grommets, and to work with funding mechanisms and pathways to match the estimate of most appropriate and cost-effective need.

6.5 Commentary

This brief example of variation in ventilation tubes highlights a number of points:

- the interconnectedness of the issue across the health system, from patient information and knowledge, to referrer evidence, to surgical practice
- the importance of environmental factors in determining local patterns of health care
- apparent large local-level variation can be an artefact of small numbers and low rates, meaning that within a district, variation is a less important factor than the overall level of utilisation in the district
- the need for further information, but much of this is centred around the actual incidence of the disease, and understanding the nature of need and demand in the population, rather than necessarily being focused upon clinicians and their practice.

7. References

- Adams G, Gulliford MC, Ukoumunne OC et al. 2004. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *Journal of Clinical Epidemiology* 57(8): 785–94.
- Davis PB, Lay-Yee R, Briant R et al. 2002. Adverse events in New Zealand public hospitals I: Occurrence and impact. *New Zealand Medical Journal* 115(1167).
- Diehr P. 1990. What is too much variation? The null hypothesis in small-area analysis. *Health Services Research* 24(6): 741–71.
- Expert Advisory Group. 2014. *Integrated performance and incentive framework: Expert Advisory Group final report*. Wellington: Ministry of Health. URL: <http://www.hiirc.org.nz/page/45527/>
- Gribben B, Salkeld LJ, Hoare S et al. 2012. The incidence of acute otitis media in New Zealand children under five years of age in the primary care setting. *Journal of Primary Health Care* 4(3): 205–12.
- Patterson SM, Hughes C, Kerse N et al. 2012. Interventions to improve the appropriate use of polypharmacy for older people. *Cochrane Database of Systematic Reviews* 5(CD008165).
- Turner RM, Omar RZ, Thompson SG. 2001. Bayesian methods of analysis for cluster randomized trials with binary outcome data. *Statistics in Medicine* 20(3): 453–72.
- Wennberg J. 2010. *Tracking medicine: A researcher's quest to understand health care*. Oxford: University Press.
- Wennberg J. 2011. Time to tackle unwarranted variations in practice. *British Medical Journal* 342(7799): 687–90.



HEALTH QUALITY & SAFETY
COMMISSION NEW ZEALAND
Kupu Taurangi Hauora o Aotearoa

New Zealand Government