

August 2015

**Questionable care: avoiding ineffective treatments**

Methodological supplement

By Stephen Duckett, Peter Breadon and Danielle Romanes

## Table of contents

1	Choosing do-not-do treatments .....	3
2	Measuring do-not-do treatments.....	5
3	Comparing hospitals .....	11
4	Limitations.....	15
	References.....	17

## 1 Choosing do-not-do treatments

*Questionable care* analyses the use of five do-not-do and three do-not-do-routinely treatments. This section explains the selection process we followed in choosing these treatments from the much larger number available.

### 1.1 Finding do-not-do recommendations

We chose do-not-do treatments from five sources, which together identified over 1200 forms of ineffective care.<sup>1</sup>

Our government advisory sources were:

- The UK's National Institute of Health and Care Excellence (NICE). NICE makes explicit do-not-do recommendations in consultation with clinical experts. We considered 961 of its recommendations.
- Australia's Medicare Services Advisory Committee (MSAC). MSAC provides funding and de-funding advice to the Minister of Health. We considered 200 of its recommendations.

Our academic sources were:

- An Australian study of potentially low-value health care practices listed on the Medicare Benefits Schedule.<sup>2</sup> We considered all 156 of these treatments.

---

<sup>1</sup> There is some overlap between the recommendations on this list.

<sup>2</sup> Elshaug, *et al.* (2012)

- A US study of contradicted medical practices, drawn from an analysis of articles published between 2001 and 2010 in the *New England Journal of Medicine*.<sup>3</sup> We considered all 146 of these contradicted treatments.

### 1.2 Inclusion criteria

We selected do-not-do treatments from the recommendations outlined above by using the following inclusion criteria:

- **Site:** the treatment had to take place in a hospital. This reduced our potential list significantly, as many of the treatments were specific to primary care.
- **Timing:** the recommendation had to have been published prior to financial year 2010-11, when our data were collected.<sup>4</sup>
- **Evidence:** the treatment had to be either disproven or unproven. In practise, this meant compelling evidence of *ineffectiveness in the case of two procedures*,<sup>5</sup> and no

---

<sup>3</sup> Prasad, *et al.* (2013)

<sup>4</sup> For example, NICE guidance from 2013 recommended vertebroplasty as an option to treat osteoporotic spinal fractures (under some circumstances), but this was not included because it was released after our data period, NICE (2013).

<sup>5</sup> Arthroscopic debridement for osteoarthritis of the knee, and vertebroplasty for osteoporotic spinal fractures. Each of these procedures had been disproven by two high quality randomised control trials.

compelling evidence of effectiveness for another.<sup>6</sup> We did not assess the quality of underlying evidence for the remaining three procedures, because they were explicitly formulated as do-not-do recommendations by NICE in consultation with clinical experts. We therefore relied on NICE findings about the suitability of these treatments. All recommendations were excluded if they were persuasively contradicted by evidence that was published after the recommendation but before data was collected.

- **Measurable:** recommendations were excluded if the do-not-do treatment and relevant patient group descriptions could not be accurately measured using the demographic characteristics, procedure and diagnosis codes recorded by the hospitals in our dataset,<sup>7</sup> or if we could not exclude patient groups for which use of the procedure may be legitimate.<sup>8</sup> Recommendations were also excluded if they required information (such as medications, test results, or the timing of an intervention) that our data did not include.<sup>9</sup>

---

<sup>6</sup> Hyperbaric oxygen therapy for various indications. The most common of these, non-diabetic wounds and ulcers, was not supported by any compelling evidence of benefit in human trials, Kranke, *et al.* (2004).

<sup>7</sup> These codes are described below.

<sup>8</sup> For example, we eliminated one potential do-not-do – radiotherapy for age-related macular degeneration (AMRD) – because the procedure codes available for radiotherapy were not site-specific. This meant that it was unclear whether patients were having radiotherapy for their eyes (which AMRD affects) or other conditions. All the patients with both AMRD and radiotherapy also had cancer diagnoses, which are legitimate reasons to have radiotherapy. We therefore could not say conclusively that patients were receiving a do-not-do treatment.

<sup>9</sup> For instance, we could not link individual patients' records and so could not analyse recommendations specific to the sequence of interventions.

- **Prevalence:** the procedure had to happen at least five times in a year. We imposed this filter because some of the instances may be due to coding errors.

### 1.3 Clinical review

After inclusion criteria were applied, we ran the remaining do-not-do treatments first past a panel of general clinical experts and then a selection of specialists relevant to each treatment. In several cases we excluded patients with specific comorbidities on clinical advice, so long as the advice did not directly contradict the clinical evidence behind the do-not-do recommendations.

#### 1.3.1 Choosing do-not-do-routinely treatments

Our process with do-not-do-routinely treatments was slightly different. We used an opportunistic approach to selection, by selecting three recommendations that met all of the inclusion criteria.

## 2 Measuring do-not-do treatments

This chapter describes the data and coding used in this project.

### 2.1 Data

This report used patient-level data from the 2010-11 National Hospital Morbidity Database. The Database contains a relatively comprehensive picture of hospital activity in Australia, with extensive clinical information recorded for 8,720,771 patient admissions.

The Database includes data for hospital sites of all sizes (even very small clinics). It has public hospital data for all states except the ACT, and private hospital data for all states except the ACT, NT and Tasmania. Hospitals are distinguished by random codes, but are not identifiable. All private hospital data are grouped under a single variable, at state rather than hospital level.

The data were provided by the Australian Institute of Health and Welfare, with use and analysis approved by the states and territories.

### 2.2 Defining do-not-do patients

Every do-not-do recommendation was specific to a certain patient group. These groups were defined using treatment and diagnosis codes, which are provided in this chapter.

The diagnosis codes used are from the 10<sup>th</sup> revision of the International Statistical Classification of Diseases and Related Health Problems (Australian modification, 7<sup>th</sup> edition). The

Procedure codes used are from the Australian Classification of Health Interventions (7<sup>th</sup> edition).

In many cases, we needed to define exclusions to the patient group, as well as inclusions. This is because a patient could have an additional diagnosis that legitimated use of the procedure, and we were unable to tell which indication the procedure was performed for. These exclusions are listed below.

Researchers using our methodology should note that these exclusions are not exhaustive. They were selected from the procedures and diagnoses of the patients in our do-not-do groups. Additional exclusions may need to be added if the analysis is repeated on a different group of patients, or on a different time period.<sup>10</sup>

#### 2.2.1 Arthroscopic debridement for osteoarthritis of the knee

Patients had to have the following procedure code, which is for arthroscopic debridement of the knee:

- 4955800.

Patients also had to have at least one of the following diagnostic codes, which are for osteoarthritis:

---

<sup>10</sup> The appropriateness of clinical choices depends on contemporary clinical evidence, which analyses should take into account.

- M170, M171, M172, M173, M174, M175, M179.

Patients could not have any of the following diagnosis codes, as they could potentially be legitimate reasons to use the procedure. The codes principally include meniscal tears, which were referenced by the source articles as a potentially legitimate reason to perform arthroscopic debridement.

- M1126, M224, M23, M230, M2303, M2304, M2306, M2309, M231, M2313, M2316, M232, M2320, M2321, M2322, M2323, M2324, M2325, M2326, M2329, M233, M2330, M2332, M2333, M2334, M2335, M2336, M2339, M2340, M234, M2341, M2342, M2343, M2344, M2345, M2346, M2347, M2349, M2350, M2351, M2352, M2353, M2359, M238, M2380, M2381, M2382, M2383, M2384, M2385, M2386, M2387, M2389, M239, M2391, M2392, M2393, M2394, M2396, M2399, M6596, M6786, M9486, S832, S833, 4950002, 4955702, 4955801, 4956000, 4956003, 4956101, 4956201, 4956300, 4956600, 5012401.

### **2.2.2 Vertebroplasty for osteoporotic spinal fractures**

Patients had to have one or both of the following procedure codes, which are for vertebroplasty:

- 3540000, 3540001.

Patients also had to have at least one of the following diagnostic codes, which are for osteoporotic fractures:

- M8008, M8028, M8058, M8088, M8098.

Our reading of the source articles and subsequent clinical consultation did not result in any exclusion recommendations for this procedure.

### **2.2.3 Removal of healthy ovaries during hysterectomy**

Patients had to have at least one of the following procedure codes, which are for oophorectomy:

- 3563801, 3563802, 3563803, 3563811, 3563812, 3571307, 3571311, 3571704.

Patients also had to have at least one of the following procedure codes, which are for hysterectomy:

- 3565300, 3565301, 3565304, 3565700, 3566100, 3566400, 3566401, 3566700, 3566701, 3567000, 3567302, 3575000, 3575302, 3575600, 3575603, 9044800, 9044801, 9044802.

Patients could not have any of the following diagnosis codes, as they could potentially legitimate use of the procedure. The codes principally include gynaecological cancers, endometriosis, and risk factors cited as reasons for prophylactic oophorectomy:

- C56, C796, D27, D391, N801, C541, C539, C55, N802, N851, C772, C570, D069, C578, D069, Z4001, Z4008, N809.

### **2.2.4 Laparoscopic uterine nerve ablation for chronic pelvic pain**

Patients had to be identified as women, and had to have the following procedure code, which is for laparoscopic uterosacral

nerve ablation:

- 3563814.

Patients also had to have one of the following diagnosis codes, which are for pelvic pain:

- R102, N731.

Our reading of the source articles and subsequent clinical consultation did not result in any exclusion recommendations for this procedure.

### 2.2.5 Hyperbaric oxygen therapy for various indications

Patients had to have one of the following procedure codes, which are for hyperbaric oxygen therapy sessions of varying lengths:

- 9619100, 1302000, 1302500.

Patients also had to have one of the following diagnosis and procedure codes, which are for acute ankle sprains, carbon monoxide poisoning, Crohn's disease, non-diabetic wounds and ulcers, osteomyelitis, sudden idiopathic hearing loss, peripheral vascular disease, skin grafts and flaps, and cancer:

- T58, T597, M8617, M8618, M8646, M8666, M8667, M8668, M8687, M8688, M8694, M8695, M8696, M8697, M8698, H903, H904, H905, H912, H919, H931, K50, K500, K501, K508, K509, S934, S9340, S9341, S9342, S9343, K626, L891, L892, L893, L899, L97, L984, S010, S0131, S0188, S211, S3180, S510, S519, S6181, S6188, S619, S810, S817,

S8181, S8188, S819, S910, S912, S913, S9181, T013, T793, T813, T8141, T8903, T930, I73, I730, I731, I738, I739.

- Any diagnosis code from the ICD-10 neoplasms chapter (C00 - D48).
- 4520000, 4520608, 4523901, 4540000, 4543900, 4544200, 4544804, 4544809, 4544810, 4545109, 4545126, 4549600, 4556200.

Patients could not have any of the following diagnosis codes, as they could potentially be legitimate reasons to use the procedure. The codes principally include potential sources of radiation injuries, diabetes, osteoradionecrosis, decompression sickness, air or gas embolisms, gas gangrene, and necrotising soft tissue infections:

- M8731, M8738, M8785, M8788, M8795, M8798, Z298, K520, K627, L598, L599, M962, N304, T66, Z923, E10-E14, G374, I775, K041, K102, M31, M318, M319, M319, M726, N498, N768, O24, O240, O241, O242, O243, O244, O249, P77, T703, T790, T800, T875, A480, A690.

### 2.2.6 Do not routinely perform fundoplication for gastro-oesophageal reflux disease

Patients had to have one of the following procedure codes, which are for fundoplication:

- 3052702, 3052700, 3052704.

Patients also had to have one of the following diagnosis codes,

which are for gastro-oesophageal reflux disease:

- K210, K219.

Patients could not have any of the following diagnosis codes, as they could potentially be legitimate reasons to use the procedure. The codes are for diaphragmatic, abdominal and hiatus hernias:

- K44, K440, K441, K449, K45, K450, K451, K458, K46, K460, K461, K469, Q790, Q401.

### **2.2.7 Do not routinely perform amniotomy in normally progressing labour**

Patients had to have the following procedure code, which is for amniotomy augmenting (rather than inducing) labour:

- 9046601.

Patients also had to have the following diagnosis code, which is for single spontaneous delivery:

- O80.

Patients could not have any of the following diagnosis or procedure codes, in part because the guidance was restricted to normal births, excluding them on that basis, and in part because they could potentially be legitimate reasons to use the procedure. The diagnoses include diabetes, pre-eclampsia, and hypertension, in addition to multiple, obstructed, complicated, prolonged and induced deliveries:

- O60, O600, O601, O602, O603, O140, O141, O142, O149, O15, O150, O151, O152, O159, E10, E100, E101, E102, E103, E104, E105, E106, E107, E108, E109, E11, E110, E111, E112, E113, E114, E115, E116, E117, E118, E119, E12, E120, E121, E122, E123, E124, E125, E126, E127, E128, E129, E13, E130, E131, E132, E133, E134, E135, E136, E137, E138, E139, E14, E140, E141, E142, E143, E144, E145, E146, E147, E148, E149, O63, O630, O631, O639, O364, P95, O312, O300, O301, O302, O308, O309, O840, O841, O842, O848, O849, P050, P051, P059, O81, O640, O688, O698, O682, O665, O691, O830, O831, O64, O641, O642, O643, O644, O645, O648, O649, O65, O650, O651, O652, O653, O654, O655, O658, O659, O66, O660, O661, O662, O663, O664, O668, O669, O68, O680, O681, O683, O689, O69, O690, O692, O693, O694, O695, O699, 9046500, 9046501, 9046502, 9046503, 9046504, 9046505.

### **2.2.8 Do not routinely perform episiotomy during spontaneous vaginal birth**

Patients had to have the following procedure code, which is for episiotomy:

- 9047200.

Patients also had to have the following diagnosis code, which is for single spontaneous delivery:

- O80.

Patients could not have any of the following diagnosis or procedure codes, in part because the guidance was restricted to

normal births and excluded them on that basis, and in part because they could potentially be legitimate reasons to use the procedure. The diagnoses include diabetes, pre-eclampsia, and hypertension, in addition to multiple, obstructed, complicated, and prolonged deliveries:

- O60, O600, O601, O602, O603, O140, O141, O142, O149, O15, O150, O151, O152, O159, E10, E100, E101, E102, E103, E104, E105, E106, E107, E108, E109, E11, E110, E111, E112, E113, E114, E115, E116, E117, E118, E119, E12, E120, E121, E122, E123, E124, E125, E126, E127, E128, E129, E13, E130, E131, E132, E133, E134, E135, E136, E137, E138, E139, E14, E140, E141, E142, E143, E144, E145, E146, E147, E148, E149, O63, O630, O631, O639, O364, P95, O312, O300, O301, O302, O308, O309, O840, O841, O842, O848, O849, P050, P051, P059, O81, O640, O688, O698, O682, O665, O691, O830, O831, O64, O641, O642, O643, O644, O645, O648, O649, O65, O650, O651, O652, O653, O654, O655, O658, O659, O66, O660, O661, O662, O663, O664, O668, O669, O68, O680, O681, O683, O689, O69, O690, O692, O693, O694, O695, O699.

### 2.3 Costing do-not-do treatments

Our report is about quality of care, rather than cost, but we did make an estimate of the cost of do-not-do procedures to illustrate that these treatments come at a cost.

We used the National Hospital Cost Data Collection (NHCCD), which was provided by the Independent Hospital Pricing Authority (IHPA). Our analysis of this dataset was authorised by States and Territories. This dataset does not include private hospitals or small rural hospitals.

We were interested in the cost of specific do-not-do procedures (e.g. vertebroplasty), independent of other procedures that a patient might have. For that reason, we looked at the cost of admissions in the NHCCD which involved the relevant procedure (e.g. vertebroplasty) and very few other procedures.

For arthroscopic lavage and vertebroplasty, only admissions with a total of one or two procedures were included. For uterine nerve ablation, the threshold was increased to three to ensure that a sufficient number of admissions were included. For hyperbaric oxygen therapy, only admissions with a single procedure were used (this treatment can be offered without any form of anaesthesia, unlike the other do-not-do treatments).

Do-not-do oophorectomies only occur at the same time as a hysterectomy. For this reason, we did not attempt to isolate the cost of the oophorectomy.

For each procedure the average cost (among low-procedure admissions) was adjusted to 2014-15 dollars using IHPA's indexation rate of 4.7%. The average cost of each of the procedures is in the table below.

**Table 1: Average admission cost, 2010-11 (\$2014-15)**

Uterine nerve ablation	\$4,412
Arthroscopy	\$3,566
Vertebroplasty	\$3,252
Hyperbaric oxygen	\$1,298

*Note: Only admissions with a few procedures (maximum of 3) included.  
Source: Grattan Institute analysis of NHCCD*

By excluding admissions with many procedures, we avoided attributing some irrelevant costs to the do-not-do treatment. However, this may also have excluded patients who had several procedures arising from their do-not-do treatment.

More importantly, because our estimate includes the fixed costs of admission, it is only approximate. In other words, we assume that all the costs of a relatively 'simple' admission (one with very few procedures) are associated with the do-not-do procedure. Some patients who got a do-not-do procedure would have been hospitalised even if they did not receive that specific treatment. For these patients, our figures are an over-estimate.

### 3 Comparing hospitals

*Questionable care* compares the rate of do-not-do procedures at different hospitals. This chapter explains how we calculated these rates. Wherever possible, our approach attempts to compare treatment choices by clinicians and to exclude variation from other causes such as hospital case-mix or specialisation.

#### 3.1 Calculating rates

##### The numerator: do-not-do treatments

The numerator for calculating the rate of do-not-do treatments at a hospital is the number of do-not-do treatments given in 2010-11 (these treatments are defined in Section 2.2).

##### The denominator: relevant patient groups

Instead of measuring the rate of do-not-do treatments among all of a hospital's admissions, we look at the rate in the relevant patient group. These are the patients who should *not* get the treatment. They are defined by the diagnostic and demographic inclusion markers explained in Section 2.2.

By measuring variation in treatment within the relevant groups, we ensure that hospital results are not skewed by irrelevant admissions – patients who would never receive the do-not-do treatment, or for whom it is appropriate.

##### The comparators: relevant hospitals

We compare hospital rates with an average. It is not the average across all hospitals, but across a comparator group of hospitals it varies for each do-not-do treatment.

To be included in a comparator group, a hospital must treat more than five patients a year in the relevant patient group (for example, patients with compression fractures). In addition, hospitals must provide the relevant procedure (for example, a vertebroplasty) at least once a year. This applies to all procedures except hyperbaric oxygen therapy, which the hospital must provide at least five times a year.

We use a higher threshold for hyperbaric oxygen therapy because false positives are more likely for this do-not-do treatment. A false positive from data errors can happen in three ways:

- an incorrect diagnosis code with a correct procedure code
- an incorrect procedure code with a correct diagnosis code
- an incorrect diagnosis code and an incorrect procedure code.

The first type of error is more likely if there are many patients who have the relevant do-not-do procedure. The second type is more likely if many patients have the relevant do-not-do diagnoses. Hyperbaric oxygen therapy is among the most prevalent do-not-do procedures (4659 instances in 2010-11) with by far the most

do-not-do diagnoses (1.3 million patients suffer from these conditions, compared to between 6,000 and 100,000 for the other do-not-do patient groups). The high frequency of the procedure and the diagnoses both increase the risk of false positives. In addition, it seems unlikely that a hospital site would use costly, highly-specialised equipment only one or two times a year, especially as it is used to treat relatively common diagnoses.

This avoids the average being skewed by hospital specialisation. Without the comparator groups, the average would be dragged down by zero rates in hospitals that have no opportunity to provide the do-not-do treatment.

The minimum requirement for the number of relevant patients avoids including extremely high rates. For example, if a hospital only treats one compression fracture patient, and gives them a vertebroplasty, they would get a do-not-do rate of 100%. Comparator groups vary in size, and are sometimes small, so extreme values of this kind could have a big impact on the average.

Overall, the comparator groups result in higher average do-not-do treatment rates (Figure 1). This makes our benchmarking rules, discussed in the next selection, more conservative, selecting hospitals which are further from normal patterns of care.

**Figure 1: Using comparator groups results in fairer, and higher, average do-not-do treatment rates**

Do-not-do	'Raw' average across all hospitals	Average across hospitals that provide relevant procedure	Average among hospitals with >5 relevant patients	Average with both filters (procedure and patient)
HBOT	0.04%	1.40%	0.04%	2.93%
Arthroscopy	0.96%	3.40%	2.17%	3.34%
Vertebroplasty	0.22%	5.22%	0.82%	5.40%
Ovary removal	0.08%	1.34%	0.32%	1.40%
Nerve ablation	0.06%	2.10%	0.21%	2.10%

### 3.2 Comparing hospitals

#### 3.2.1 Outlier hospitals

The rates we calculate can never be perfect (limitations are discussed in Chapter 4). In particular, they may be influenced by false positive do-not-do treatments caused by coding errors. In addition, in rare cases an unproven treatment could arguably be warranted. For example a patient may be ineligible for all other treatments, all alternatives may have failed, or available evidence may not apply to the patient's particular combination of health problems and/or other characteristics.

For these reasons, only hospitals that provide do-not-do treatments at above the average rate are considered ‘outliers’. As discussed in the body of the report, we use this term to denote hospitals with unusual and concerning patterns of care, despite the fact that a statistical measure would not label all above average observations as outliers.

For do-not-do routinely treatments, the top 10 per cent of hospitals (by rate of provision) in a comparator group are considered outliers. We considered other measures, such as defining outliers relative to the average rate (e.g. with a threshold of double or triple the national rate) or a certain number of standard deviations above the mean.

The first option may not work well for do-not-do routinely treatments with very high average rates of provision. The second may be hard to interpret for clinicians and system managers.

No rule for identifying outliers can be used without evaluating the distribution to see if the rule makes sense. For example, for our rule to identify do-not-do routinely outlier hospitals, if rates are too uniform the top 10 per cent of hospitals will not have treatment patterns that are meaningfully different from most other hospitals.

### 3.2.2 Low-volume do-not-do treatments

Some do-not-do treatments are rare. In these cases, a small number of treatments can cause a hospital to be an outlier and there is a risk that a hospital’s outlier status will change frequently from year to year.

This could be considered a weakness in our approach to

comparing hospitals. However, our policy recommendations have been designed to reduce any negative impacts from this risk.

Under our recommendations, outlier hospitals only face serious consequences after they have been given a chance to improve, failed to improve, had an adverse clinical review, and again failed to improve. Only then do we propose financial sanctions or changes of management. If a hospital sustains outlier status through this entire process, they are certainly providing inappropriate care.

### 3.2.3 Hospital characteristics

We compared rates of do-not-do treatments at hospitals of different sizes, and with different scopes of service.

**Size** was defined simply as the number of patient admissions a hospital had during financial a year.

**Scope** was defined using an Information Theory Index (ITI), which measures how similar a hospital’s mix of patients is compared to the system as a whole.<sup>11</sup> We calculate this in terms of how patients are distributed across different specialties (e.g. orthopaedics, rheumatology etc.). If there are N different specialties, the scope index for provider h is calculated as:

---

<sup>11</sup> The index was originally designed to measure gains in information, but the index is also an established measure of hospital specialisation (dating back to at least the late 1980s). See Kobel and Theurl (2013) for discussion.

$$scope_h = \sum_{j=1}^N p_{jh} \ln \left( \frac{p_{jh}}{\varphi_j} \right)$$

Where:

- $p_{jh}$  is the proportion of provider h's patients that are in specialty  $j$ ,
- $\varphi_j$  is the proportion of all patients that are treated by specialty  $j$ .

A provider with a scope index of zero has the exact same mix of patients as a summation of all public hospitals contributing to the NHCDC data. As the concentration of services increases, so does the scope index. The index has no upper bound but in practice even the most focussed hospitals (for example providers in which 95% of the separations are coded to dialysis) have a scope index of around 3.

## 4 Limitations

*Questionable care* provides an incomplete picture of potentially ineffective care in Australian hospitals. This chapter outlines the key limitations of the report.

### 4.1 Selection of do-not-do treatments

The do-not-do treatments that we analyse are not representative. They comprise a small fraction of do-not-do guidance and a tiny proportion of clinical activity. As such, our findings cannot be used to generalise about the characteristics of hospitals that provide do-not-do treatments, the patients that receive them, or the level of system-wide adherence to all do-not-do guidance.

### 4.2 Data limitations

Our version of the National Hospital Morbidity Database did not include consistent patient identifiers. For this reason, we could not correct for readmissions.<sup>12</sup> This may deflate our hospital do-not-do rates, making our analysis conservative but potentially reducing the accuracy of comparisons between hospitals.

We had no data for the ACT or for private hospitals in the NT and Tasmania, and all private hospital data were aggregated at the

---

<sup>12</sup> To understand this limitation, consider the following scenario. A patient arrives in hospital in January and is diagnosed with osteoarthritis of the knee. They are asked to come back in February for an arthroscopy (a do-not-do treatment). In March they return for a different treatment. In each of the visits they are recorded as having osteoarthritis of the knee. We would count this person three times in the 'relevant patient group', but they should only be counted once because the relevant clinical decision (to provide a knee arthroscopy) was only made once.

state level. These factors made it impossible to provide a complete analysis of variation between states, or among individual private hospitals.

No data are perfectly accurate and coding errors may affect our results.<sup>13</sup> However, most states use coding audits and data validation techniques to review and improve the accuracy of the dataset we used.<sup>14</sup> In addition, our recommendations include several measures to minimise the negative consequences of any false positives.

First, our recommendations only apply to outlier hospitals, reducing the risk of false positives caused by a small number of data errors. Second, as mentioned above, we recommend provision of information and clinical reviews before consideration of any harsher sanctions.

Most of our analysis is at the level of hospital sites. Our data agreement prohibited us from aggregating data into local hospital networks. Since hospital network CEOs and boards bear the ultimate organisational accountability for quality of care, the hospital network may be a preferable unit of analysis. Analysis of hospital site performance may also mask the impact of individual clinicians who work across different hospital sites.

---

<sup>13</sup> This includes recording incorrect codes, or failing to record codes for diagnoses which could be legitimate reasons to provide a 'do-not-do' treatment.

<sup>14</sup> AIHW (2014)

In the main report, we recommend that the Australian Commission for Safety and Quality in Health Care expand our approach. Many of the limitations in our research would not apply to the Commission's work. The Commission would have consistent patient-level identifiers, data for all states and data for individual private hospitals.

## References

- AIHW (2014) Australian hospital statistics 2012-13, Australian Institute of Health Welfare, accessed 25 November 2014, from <http://www.aihw.gov.au/publication-detail/?id=60129546922>
- Elshaug, A. G., Watt, A. M., Mundy, L. and Willis, C. D. (2012) 'Over 150 potentially low-value health care practices: an Australian study', *Med J Aust*, 197(10), p 556-560
- Kobel, C. and Theurl, E. (2013) Hospital specialisation within a DRG-Framework: The Austrian case, Working Papers in Economics and Statistics, University of Innsbruck from <http://eeecon.uibk.ac.at/wopec2/repec/inn/wpaper/2013-06.pdf>
- Kranke, P., Bennett, M., Roeckl-Wiedmann, I. and Debus, S. (2004) 'Hyperbaric oxygen therapy for chronic wounds', *Cochrane Database Syst Rev*, 2(2).
- NICE (2013) NICE technology appraisal guidance 279: Percutaneous vertebroplasty and percutaneous balloon kyphoplasty for treating osteoporotic vertebral compression fractures, National Institute for Health and Care Excellence from <https://www.nice.org.uk/guidance/ta279/resources/guidance-percutaneous-vertebroplasty-and-percutaneous-balloon-kyphoplasty-for-treating-osteoporotic-vertebral-compression-fractures-pdf>
- Prasad, V., Vandross, A., Toomey, C., Cheung, M., Rho, J., Quinn, S., Chacko, S. J., Borkar, D., Gall, V. and Selvaraj, S. (2013) A decade of reversal: an analysis of 146 contradicted medical practices, *Mayo Clinic Proceedings*