6 Discussion

6.1 Can we produce in-hospital mortality indicators using Australian administrative data?

This study demonstrates an approach to specifying a set of indicators of in-hospital mortality and calculating values for them using currently available Australian administrative data, from the NHMD. The analytic approach was based on findings of a review of relevant literature.

The work demonstrates the technical feasibility of producing indicators of in-hospital mortality now using national data. In particular, the section based on longitudinal analysis of 3 years of data provides support for the position that current Australian morbidity data are largely of adequate quality to support this type of use.

The particular indicators specified here are not the only ones possible. However, they exemplify major types of indicators: namely those focusing on in-hospital mortality among relatively high-risk cases, those focusing on in-hospital mortality among low-risk cases, and an overall group including all cases and in-hospital deaths. They are general-purpose indicators, rather than indicators specific to particular types of diagnosis, treatment or service. They should be applicable to a wide range of hospitals, though probably less so for some (e.g. specialised hospitals with an atypical casemix, such as WCHs).

The present study was based on current holdings of Australian hospital separation data, and this was considered to provide a sufficiently robust basis for the current program of work. Our assessment of the literature, as reviewed in Section 2.4.2, led us to conclude that it may be preferable to include deaths occurring soon after discharge from hospital, and that death within 30 days of discharge is a suitable criterion. At the national level, this data set has not been routinely linked to other major national databases such as the National Death Index, although several jurisdictions have undertaken such linkages at state level demonstrating its feasibility. As we have stated, we were not able to apply this aspect of indicator definition because we did not have access to linked hospital separations and deaths data, but it is likely that this will become feasible in the near future.

The following sections present further discussion of these and other points raised by this project.

Model assessment

The model parameters generated by applying the RACM (the de facto international standard logistic regression model for in-hospital mortality) to Australian data are very similar to those reported in the international literature. The model shows good discrimination (in terms of the c-statistic, 0.87 for the high-risk 80% mortality group). As reported elsewhere, the explanatory power of the model, as indicated by pseudo R^2 values, although seemingly modest (0.24 for the 80% set), were consistent with the international literature and typical for logistic regression models that compare the fitted model with the null model, in which none of the variation is explained. (This differs from the situation with linear regression, where comparison is with the saturated model, in which

100% of variation is explained, and where higher R^2 values are often obtained.) The c-statistic and pseudo R^2 values were higher for the 20% model (0.96 and 0.34, respectively) and the 100% model (0.95 and 0.35, respectively). In these sets, subjects had been grouped according to deciles of mortality risk—based on primary diagnosis. Although this technique guarantees an increase in discrimination and explanatory power, the change in pseudo-R^2 values and c-statistics with the inclusion of the deciles were similar to the changes in these statistics with the inclusion of the primary diagnosis groups for the 80% models.

The Hosmer–Lemeshow test did not demonstrate good fit for any of the RACM models. However, as has been discussed previously, the Hosmer–Lemeshow goodness of fit method is sensitive to the very large sample sizes used here, and the RACM model is not sophisticated. The graphic plots of deciles of observed and expected risks show that the RACM model fit is closer for the deciles of higher risk than for the lower deciles, where the model seems to over-predict expected mortality. This exemplifies the problems of fitting graduated risks in relation to outcomes that can have only one of two values: alive or dead (Chassin et al. 1996).

The unevenness across the deciles of risk is likely to be much less important for large hospitals, with large volumes of patients and larger number of both observed and expected deaths, than for smaller institutions. Indeed, the caterpillar plots for the A1 peer group hospitals show HSMRs that vary steadily across a substantial range, and demonstrated that there are large hospitals with HSMRs with narrow confidence intervals that have mortality rates that are significantly below, or significantly above, the national average for that peer group. But the unevenness, which is more marked for peer groups including hospitals with lower case volumes, further confirms the inappropriateness of simply rank ordering the hospitals from end to end, rather than looking for outlier groups and institutions.

Although the HSMRs for the B1 peer group hospitals are within realistic boundaries, those for the small hospitals in C2 and D1 are hard to interpret. The small number of both observed and expected deaths generates HSMRs in those groups that range from 0 to over 300—some with very wide confidence intervals.

Analysis by peer group

Unadjusted HSMRs should be expected to differ between hospitals because of their different casemix. Adjustment models, such as those presented in this report, do much to overcome differences in casemix, but institutional level differences remain. Hospitals have been divided into peer groups to enable comparisons of like with like. Peer groups may also provide a useful basis for assessment of in-hospital mortality. The analyses presented in this report are based on overall models—based on all hospitals. It would also be possible to make peer-group-specific models, at least for the groups that treat sufficiently large numbers of patients. In any case, interpretation of HSMRs across peer-group boundaries should be undertaken cautiously.

Small hospitals

The dilemmas posed by small hospitals are substantial. In-hospital death is a relatively rare event in many of those settings, and mortality rates are likely to be subject to many extraneous influences related to the casemix of those hospitals, and to opportunities for end-of-life care in rural and remote regions.

The most straightforward way to deal with small hospitals is to exclude them from mortality monitoring: setting some mortality rate criteria (e.g. at least 50 deaths in any one of the three previous financial years, or some other mortality threshold yet to be determined). Other simple approaches to dealing with sparse data include enlarging the reporting period for small hospitals (e.g. calculate HSMRs for a rolling 2– or 3–year period) or reporting HSMRs only for clusters of small hospitals. If none of these is deemed sufficient, then analysis using a Bayesian method that creates shrunken estimates—i.e. estimates of the HSMR which are shifted towards a value obtained from known information about other hospitals (known as a 'prior' probability)—could be developed for consideration by the Commission.

Refined risk-adjustment model (ERM)

This discussion has so far revolved around findings based on the de facto international standard risk-adjustment model (RACM). The modest fit of this model prompted us to consider whether it could be improved. We developed a more refined risk-adjustment model (labelled the ERM model in this report) that allowed for the possibility that some variables, such as age, were not simply linear in relation to mortality risk. The model also allowed for interactions between the modelled variables (we found significant interactions for all the major variables modelled). We acknowledge helpful advice from Professor DW Hosmer in the course of this work.

The ERM model displayed a number of technically more acceptable characteristics. The model fit was a substantial improvement over the RACM model, and the residual differences between observed and expected mortality were spread much more evenly across the risk deciles.

At this point, we have not gone on to analyse all peer groups for every combination of mortality using the ERM modelling, and the ERM analysis is provided for the sake of comparison. Technically, it is a superior model and the improvements in model fit justify its further development. If, however, there is a concern that any Australian study should follow work done internationally, then the Commission may want to continue with the RACM model, despite its poorer performance. On a practical note, the large number of interactions that are computed within the ERM model make major demands on computing power. Interested groups lacking access to powerful desktop computers can expect long processing times to compute the ERM models.

SEIFA and other factors

The finding that a measure of social deprivation (SEIFA) did not add substantially to the discriminatory power of the risk-adjustment modelling is ambiguous. It might reflect somewhat flatter social gradients within the Australian population than in settings in which socioeconomic variables have been found to be influential—at least in relation to access to health care. However, it could also reflect insensitivity of SEIFA to relevant aspects of deprivation or other social determinants of health. Conversely, it could be the case that variables in our model took some account of any such differences. Aboriginal and Torres Strait Islander peoples, as a group, have well-known excess early mortality and other characteristics of poor health status. We did not examine the practicability of examining this subgroup separately in the present study. Although it would be possible to make such an examination, we anticipate that relatively small case numbers and uncertain identification of Indigenous status in the NHMD would be important constraints.

Longitudinal analysis

We demonstrated the feasibility of longitudinal analysis of in-hospital mortality in Australia using NHMD data for 3 years. Analysis showed that most of the variation in HSMRs was between hospitals, not within hospitals, suggesting sufficient data quality and stability in hospital specific HSMRs to provide a basis for indicators.

The results presented for the 3–year analysis show a modest decline in risk-adjusted mortality over a 3–year period. There is some very tentative indication that this kind of pattern may be emerging elsewhere (e.g. Heijink et al. 2008, Kelman and Friedman 2007). Much more detailed work needs to be done to ensure that the trend is not an artefact of a number of different factors; for example, of coding changes (between and within jurisdictions or individual institutions), a reflection of the changing demography of hospital populations (hospital populations are not simply representative of populations as a whole), or an outcome of changing locations of places of death. Analysis using additional years of data will be a stronger basis for assessing trends. But the possibility remains that the trend is real. If so, it might be the case that an increased emphasis on hospital safety is beginning to have a demonstrable effect on hospital mortality, and is possibly of sufficient interest to warrant further study.

Methods of presentation

We have demonstrated three forms of presentation of HSMRs: tables, caterpillar plots and funnel plots. Each has distinct strengths and limitations.

Tables provide ready access to specific values for an institution or a group of institutions. However, the overall pattern of HSMRs is difficult to assimilate from a large table. Also, tabulated data, ranked by HSMR values, encourages unhelpful and statistically meaningless over-interpretation of the rank position of hospitals whose HSMR values do not differ significantly. For this reason, they are not preferred as a method for public dissemination of results.

Caterpillar plots provide a good overview of the range of HSMRs and of the associated confidence intervals. HSMRs for a population of hospitals tend to include many values in a 'middle range': not different from one another to a statistically significant extent (e.g. Figure 12). Caterpillar plots show this property rather clearly, especially if they are drawn in a way that gives at least as much visual emphasis to the confidence intervals as to the point estimates. They also show outliers, if present.

Funnel plots allow the identification of those small numbers of hospitals that are true outliers, with mortality results that are either much worse, or much better, than most hospitals. One limitation is that they do not facilitate comparison of non-outlier hospitals—a matter likely to be of interest to people responsible for each charted institution. Funnel plots are, perhaps, more difficult to interpret than caterpillar plots.

We conclude that although good use can be made of all three methods of presentation, the choice for public reports (if made) should be between the two forms of chart.

We suspect that many members of the public may find caterpillar plots easier to interpret than funnel plots. We are not aware of empirical data on this matter (though a study could certainly be done).

6.1.1 Indicators specified in the project

This report presents the results of a proof-of-concept project on the development of in-hospital mortality indicators based on existing Australian administrative data.

The three indicators specified in this study are intended to represent types of indicator described in the international literature, while also reflecting a pragmatic response to characteristics of the National Hospital Minimum Dataset and to the short time available for this project. The three indicators specified in this project are:

Indicator 1: High-risk group. This was specified as the Principal Diagnoses that accounted for 80% of in-hospital deaths in Australian hospitals, and had the highest number of cases of in-hospital mortality, in 2005–06.

Indicator 2: Lower risk group. This includes all Principal Diagnoses that are not included in the first indicator, and accounted for 20% of deaths.

Indicator 3: Indicator 3 includes all principal diagnoses. Thus, it includes all cases and all deaths.

The first of the three is an example of an indicator focusing on relatively high-risk conditions. Overall, the group of 68 Principal Diagnosis codes included in it account for less than one-fifth of all cases, but the cases selected by this criterion include four-fifths of all deaths in hospital. This type of indicator (i.e. including 80% of in-hospital deaths) is quite common in the literature.

Conversely, the second is an example of an indicator focusing on a lower risk set of conditions—i.e. diagnoses which, as a group, accounted for over 80% of cases, but 20% of deaths.

The third indicator includes all cases and all deaths.

Apart from Principal Diagnosis, we applied a single set of case inclusion criteria throughout the project. These are specified in Section 4.5.1. They are similar to those reported by other recent work of similar type (e.g. CIHI 2007, Heijink et al. 2008).

Table 21 provides a demonstration of how the three generic indicators specified in this report could be applied. In this instance, specific indicators are framed in terms of a generic indicator and a hospital peer group. A similar approach could, in principle, be applied to subsets of separations grouped in other ways. Examples are the types of diagnosis, types of service and types of procedures. However, formal statistical assessment of any such groups is necessary before practical feasibility can be assured. The large number of possible variations goes beyond the scope of this report. A cautionary observation is that this approach is limited by the (fortunately) relatively low probability of most types of admitted cases ending as an in-hospital death.

Although not done in this project, it is technically possible to produce summary HSMR values for regions, jurisdictions, or other groups of cases, in much the same way as peer group summaries were produced in this project (see Table 19).

Table 21: Application of derived indicators to hospital peer group

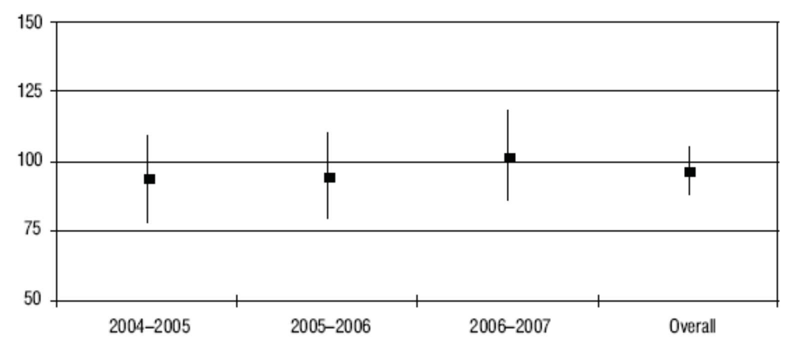
Indicator	Definition	Peer group description
Indicator 1a	Diagnoses that account for 80% of in-hospital deaths in peer group A1 Australian hospitals (high risk)	Major city hospitals with >20,000 acute casemix-adjusted separations and Regional hospitals with >16,000 acute casemix-adjusted separations per annum
Indicator 1b	Diagnoses that account for 80% of in-hospital deaths in peer group B1 Australian hospitals (high risk)	Major city acute hospitals treating more than 10,000 acute casemix-adjusted separations per annum
Indicator 1c	Diagnoses that account for 80% of in-hospital deaths in peer group C1 Australian hospitals (high risk)	Medium acute hospitals in Regional and Major city areas treating between 2,000 and 5,000 acute casemix-adjusted separations per annum, and acute hospitals treating <2,000 casemix-adjusted separations per annum but with >2,000 separations per annum
Indicator 1d	Diagnoses that account for 80% of in-hospital deaths in peer group D1 Australian hospitals (high risk)	Small Regional acute hospitals (mainly small country town hospitals), acute hospitals treating <2,000 separations per annum, and with less than 40% non-acute and outlier patient days of total patient days
Indicator 2a	Diagnoses that account for 20% of in-hospital deaths in peer group A1 Australian hospitals (high risk)	Major city hospitals with >20,000 acute casemix-adjusted separations and Regional hospitals with >16,000 acute casemix-adjusted separations per annum
Indicator 2b	Diagnoses that account for 20% of in-hospital deaths in peer group B1 Australian hospitals (high risk)	Major city acute hospitals treating more than 10,000 acute casemix-adjusted separations per annum
Indicator 2c	Diagnoses that account for 20% of in-hospital deaths in peer group C1 Australian hospitals (high risk)	Medium acute hospitals in Regional and Major city areas treating between 2,000 and 5,000 acute casemix-adjusted separations per annum, and acute hospitals treating <2,000 casemix-adjusted separations per annum but with >2,000 separations per annum
Indicator 2d	Diagnoses that account for 20% of in-hospital deaths in peer group D1 Australian hospitals (high risk)	Small Regional acute hospitals (mainly small country town hospitals), acute hospitals treating <2,000 separations per annum, and with less than 40% non-acute and outlier patient days of total patient days
Indicator 3a	Diagnoses that account for 100% of in-hospital deaths in peer group A1 Australian hospitals (high risk)	Major city hospitals with >20,000 acute casemix-adjusted separations and Regional hospitals with >16,000 acute casemix-adjusted separations per annum
Indicator 3b	Diagnoses that account for 100% of in-hospital deaths in peer group B1 Australian hospitals (high risk)	Major city acute hospitals treating more than 10,000 acute casemix-adjusted separations per annum
Indicator 3c	Diagnoses that account for 100% of in-hospital deaths in peer group C1 Australian hospitals (high risk)	Medium acute hospitals in Regional and Major city areas treating between 2,000 and 5,000 acute casemix-adjusted separations per annum, and acute hospitals treating <2,000 casemix-adjusted separations per annum but with >2,000 separations per annum
Indicator 3d	Diagnoses that account for 100% of in-hospital deaths in peer group D1 Australian hospitals (high risk)	Small Regional acute hospitals (mainly small country town hospitals), acute hospitals treating <2,000 separations per annum, and with less than 40% non-acute and outlier patient days of total patient days

6.2 How might in-hospital mortality indicators be used at different levels in Australia?

Countries that use in-hospital mortality indicators do so at different levels and for different purposes. For example the CIHI publishes HSMR trends by health region and hospital. The results are designed to be used by hospitals and health regions to monitor and understand their trends over time. An example of reporting of HSMRs by region is presented in Figure 21.

Newfoundland and Labrador
Western Regional Integrated Health Authority

Hospital Standardized Mortality Ratio, Excluding Palliative Care, 2004–2005, 2005–2006 and 2006–2007



	HSMR Excluding PC	HSMR Excluding PC 95% CI	HSMR All Cases	HSMR All Cases 95% CI
2004–2005	93	78–109	108	94–123
2005–2006	94	80–110	111	97–126
2006–2007	101	86–118	125	110–141
Overall	**96**	**88–105**	**115**	**106–123**

Hospitals included in the regional result:

- Bonne Bay Health Centre
- Calder Health Care Centre
- Dr. Charles L. Legrow Health Centre
- Rufus Guinchard Health Care Centre
- Sir Thomas Roddick Hospital
- Western Memorial Regional Hospital

Figure 21: Regional HSMR reporting by the Canadian Institute for Health Information (Taken from the report 'HSMR: a new approach for measuring hospital mortality trends in Canada.'CIHI 2007)

In Australia, in-hospital mortality indicators could be presented in a similar manner to that already used by the Canadians and others. The method of presentation will depend on the use to which the HSMRs are to be put. But, at every point, the HSMRs are always best considered as the starting points for further investigation rather than as definitive measures of a hospital's standing.

6.3 Are the in-hospital mortality indicators valid and reliable?

Validity refers to the extent to which a measurement truly measures what it is intended to measure.

If in-hospital mortality, *per se*, is the subject of interest, then the validation of indicators of the type specified in this report is relatively straightforward. Death is usually a well-defined event, though ventilators and other devices can complicate assessment. 'In-hospital death' is amenable to definition, though there is some room for ambiguity (e.g. how to treat cases of people who died while at a hospital, but had not been formally admitted, or cases where a person died before reaching a hospital, but was certified as dead after arrival?). However, the main issues are whether the available data sources are complete and reliable. These are amenable to study.

If hospital quality and safety is the subject of interest then the validation of the indicators is much more complicated. As discussed in Section 2.8.2, safety and (especially) quality are complex abstractions, which are difficult to define and measure.

The specific issue of the adequacy of administrative hospital separations data for risk adjustment could be subjected to formal study, along the lines of Aylin et al. (2007).

Reliability refers to the extent that a measurement method, if applied more than once under the same conditions, will give the same result. Repeated measurement of the same hospitals is emerging as a basis for assessing the reliability of measurements of in-hospital mortality (e.g. Heijink et al. 2008). This is based on the assumption that the true risk of in-hospital mortality in most hospitals is not likely to vary much from year to year, after adjustment for a small set of the characteristics of cases and provided that case numbers are sufficient to prevent small chance fluctuations in number of deaths from dominating results.

In this project, the 3–year analysis of in-hospital mortality, using indicator 1 (a relatively high-risk group of cases) produced the reassuring finding that within-hospital variation of HSMRs over the 3 years accounted for a generally low proportion of the total HSMR variation. In line with expectation, this was most true for groups of relatively large hospitals. Further work should be done to extend such analysis to other years, and other ways of selecting and grouping hospitals and cases.

6.3.1 Limitations

Mapping public hospitals over time

As discussed in Section 5.8.1 for the longitudinal analysis, a map was not available to track individual hospitals over time. We were able to develop a map for the purposes of this project; however, we were not able to include some hospitals (notably private hospitals) or to map some public hospitals. A map will be necessary for further longitudinal studies. Mapping is implicit in ongoing data linkage systems that include hospital separations data.

Problem of private hospitals

Private hospitals were largely excluded from this project, because they are not well-identified in the data source available to us (the NHMD). Although all records are marked as to whether the patient was in a private hospital, in many cases the information does not enable the private hospital cases to be grouped according to hospital, which was necessary for this project.

Limitations of time for project

The present project was undertaken in a short period of time. Although this did not present too great a challenge for the literature review, it did present significant challenges in the modelling aspects of the project. As indicated in Section 6.1, the long processing times to compute the models chosen for application had an impact on our ability to carry out much of the internal validation work necessary for these types of activities. With more time, we could have tried variations of indicators, applied them to hospitals grouped in additional ways, and done further development and evaluation of risk-adjustment models. It also ruled out time-consuming aspects of a more ideal study, such as attempting to obtain person-linked linked hospital and mortality data. Time constraints also had an impact on our ability to check the reliability and validity of the construction of our institutional map. We were also unable to fully explore the potential of longitudinal analysis.

6.4 Presentation and use of indicators of in-hospital mortality

6.4.1 How should in-hospital mortality indicators be presented?

In-hospital mortality indicators were presented in three different ways in the present report: as ranked tables, as caterpillar plots and as funnel plots. It is not possible to state explicitly what the best method of presentation is because any method will be governed by a number of factors including the purpose of the reporting, whether the material will be in the public or private domain, and the intended audience (experts or novices).

The main Australian example of publicly available hospital-specific reports including information on in-hospital mortality is from Queensland, where 'Measured Quality' reports are available via the Internet (e.g.
<http://www.health.qld.gov.au/quality/measured_quality/2004/bay_redl.pdf>. These are extensive reports containing a great deal of information on many aspects of hospital

performance. The mortality data in the reports is presented as numerical values in tables, with peer-group values for comparison and use of symbols to indicate differences of statistical significance, and colour to mark values assessed to be outliers.

One of the specified outcomes for the National Indicators project is to 'enable the Commission to report publicly on the state of safety and quality'. From this we are able to assume that at least some in-hospital mortality indicators, if produced, should be presented in the public domain. With this assumption in mind we recommend that primary (see below) national in-hospital mortality indicators be publicly presented, in the main, as caterpillar plots. Caterpillar plots have the advantage of simplicity compared with funnel plots and are also less likely to encourage over-interpretation of small and non-significant differences than presentation in simple 'league tables'.

Caterpillar plots can be constructed and drawn in a range of ways, some of which will be more successful than others in communicating information on in-hospital mortality. We have provided some examples of ways to construct caterpillar plots to enable consideration of this issue (Appendix 4).

Further consideration of the method of presentation will be required when considering levels of disaggregation of HSMR analysis and presentation. Inclusion of numerous hospitals (at least 10 or so; preferably 20 or more) is needed to produce a plot recognisable as a caterpillar plot. Presentation of HSMRs concerning smaller groups of hospitals could follow the methods adopted by the CIHI (2007) (see, for example, Figure 21).

6.4.2 How should in-hospital mortality indicators be used in Australia?

As discussed in the conclusion to the literature review (Section 2.10), we recommend that in-hospital mortality indicators be used as screening tools, rather than being assumed to be definitively diagnostic of poor quality and/or safety. A screening tool is a signalling device. It is intended to signal that a problem may exist and that further detailed investigation is required.

6.5 What are the methodological obstacles to producing mortality indicators in Australia now?

6.5.1 Model checking and refinement

The models used in the project (RACM and ERM) will benefit from further scrutiny and refinement. We think that the general analytic approach is satisfactory, but there is room for improvement in its details. The ERM model demonstrates the possibility of improving on the RACM model. There may be potential to improve on the current ERM model, though we did not have sufficient time to exploit this possibility. Likewise, we have demonstrated the approach when applied to general-purpose indicators, including one (the high-risk set, including 80% of in-hospital deaths) that is now common in the international literature. There has not been an opportunity in the present project to explore the performance of the approach on indicators specified in other ways (there is an almost limitless number of possible ways).

We think that data in the NHMD offer potential to develop models that improve further on the already substantial improvement of the ERM model over the RACM model. For example, probability of in-hospital death is predicted better by some four- and five-character principal diagnosis codes than by their parent three-character codes, as used in this project. The extent of the potential improvement is not yet known, nor whether the gains in model performance would outweigh the added computational burden of analysis. Other potential enhancements include inclusion of additional socio-demographic characteristics (such as Indigenous status) and peer-group specific analysis.

Better understanding of some aspects of the data might also make a useful contribution. For example, 'admission' to hospital is a complex concept, particularly for emergency cases, and there are differences between hospitals in the point at which a patient is recognised as having been admitted. Such differences could influence whether certain cases involving death soon after arrival at a hospital are recorded in the NHMD.

Of great interest is the increased precision that may result from the inclusion in the models of national level coding of variables to show whether secondary diagnoses recorded for a case were present on admission (known as C-codes in Victoria). Risk adjustment is intended to adjust for patient-level variation in risk present at the point of admission, not for adverse events and other problems that occur during hospital stays. The latter should be sought out and analysed—not included in risk adjustment. Comparisons of the precision of risk-adjustment models with and without present-on-admission codes, and the impact of that coding on HSMRs will generate considerable interest locally and nationally, and will be a major contribution to the further development of measures of hospital safety.

6.5.2 Consultation

Consultation concerning indicators of in-hospital mortality is required with technical experts and stakeholders. Engaging key stakeholders in the finalisation of a 'standard Australian method' for producing in-hospital mortality ratios has the potential to improve on the methodological work reported here. An important step is to consult with state agencies and hospital groups: can they provide evidence of jurisdiction-level (or hospital-level) data issues that might influence findings and can be taken into account in models or risk adjustment? For instance, we have excluded records that were designated neither as elective nor as emergency. This third category may have different meanings in different jurisdictions. Because omission of palliative-care cases forms part of the approach that we have taken, possible differences in identification of such cases between jurisdictions or between hospitals would also benefit from scrutiny.

6.5.3 Suggested improvements to data collections

One of the National Indicators project objectives is to 'Enable the Commission to advise Ministers on whether existing reporting processes and collections should be continued, enhanced or replaced.' The NHMD has been demonstrated—at least in the context of this report—to be adequate for producing in-hospital mortality indicators. However, a small number of enhancements of the NHMD would contribute greatly to the usefulness of the NHMD for this, and other, purposes.

Data linkage

There are two aspects of data linkage that are particularly relevant to the production of in-hospital mortality indicators. The first is internal linkage within the NHMD and the second is external linkage of the NHMD with the National Death Index. We consider both forms of linkage to be vital enhancements to the NHMD to enable more valid and reliable in-hospital mortality indicators to be produced in Australia.

Variation in definitions and practices concerning hospital admission also has potential to influence measured in-hospital mortality. An argument akin to that concerning inclusion of deaths soon after discharge could be made for the inclusion of cases in which death occurs at a hospital, but before formal admission. Whether this would have an important effect on results is not known, but warrants investigation.

Internal linkage

At the present time, separations within the NHMD are not internally linked by person. Individuals—some with serious and persisting conditions—are likely to experience more than one episode of in-hospital care within a period covered by a study of in-hospital mortality. Without the ability to link related separations, it is not possible to be sure whether a person whose episode of hospital care ended with transfer to another hospital, or with a 'statistical type change', died during the next episode of inpatient care. Even a person who separates with discharge home might have been re-admitted soon after, with the possibility of fatal outcome of that episode. We are unable to take these factors into account when modelling because of the lack of internal linkage in the National Hospital Minimum Dataset to group the separations belonging to an individual patient.

External linkage

The second role of data linkage relevant to this type of work is linkage between hospital records and death registers (i.e. the National Death Index). This is necessary to enable studies that include deaths soon after discharge (i.e. to assess 30–day mortality).

Timeliness of availability of data

Reasonable expectations for timeliness of national indicators based on hospital inpatient data are not met at present. Although case records are generally processed, coded and accessible at state or territory level within a few months of separation, the NHMD file is released only annually, and records in it are from 1–2 years old by the time they become available for use. This prompts the question: can hospital morbidity data be made available more rapidly and frequently for purposes such as reporting indicators of hospitalised mortality?

Investigations into the feasibility of a more timely release of NHMD data—perhaps quarterly—should be considered.

Validation of coding

Mortality indicators depend on the reliability and quality of coding of hospital records. The most important variable for this purpose is Principal Diagnosis.

The quality of Principal Diagnosis coding is the subject of various coding audits in which a selection of records undergo independent recoding and the results are compared with the codes originally assigned. The Australian Coding Benchmark Audit—a method for auditing the diagnosis codes assigned to separation records in Australia—has been published by the

National Centre for Classification in Health (NCCH 2000). Neither this nor any other auditing method is mandated. The extent of auditing undertaken is difficult to assess; results are usually treated as confidential and are usually not published.

A second type of tool for quality checking is also exemplified by a product of the NCCH. Performance Indicators for Coding Quality (PICQ) (NCCH 2006). PICQ is software that screens coded records for compliance with the Australian Coding Standards for the ICD-10-AM and the Australian Classification of Health Interventions. It flags errors, and probable errors: allowing checking and recoding. This sort of tool can also be used to detect patterns of errors (e.g. a high prevalence of doubtful codes for records from a particular specialty in a hospital), which can be used to prompt investigation and corrective action. As with audits, application of such tools is not mandatory, and the extent of their use is unknown.

Introduction of indicators of in-hospital mortality is likely to heighten interest in the quality of the data on which they are based. Confidence in the indicators is likely to be enhanced by undertaking and publishing data-quality audits. An example of a project and study design that could be adapted for this purpose is the study of the quality of external causes coding in a sample of records from a sample of hospitals in four states, which has recently been undertaken by a team led by Dr Kirsten McKenzie of the Queensland University of Technology (a paper relevant to this point is in preparation but has not yet been published)

6.6 International benchmarking

In order to provide an accurate point of comparison with OECD countries, the model used to calculate in-hospital mortality should be consistent with the models and methods produced elsewhere. As yet, there is no internationally governed or stipulated standard practice for calculating HSMRs; however, the RACM model is consistent with how HSMRs are calculated in a number of different countries. The ERM model makes significant improvements to the RACM and we would suggest that, with proper peer-reviewed scrutiny and replication, it may be suggested as a potential candidate for an International standard.

Alternatively, the best performing model developed on the basis of Australian data could be used for national purposes. Additional analysis using a poorer-performing, but more widely-used, model (i.e. RACM) could be undertaken for the specific purpose of international comparisons.

6.7 Conclusion

The literature review in this report shows an emerging international consensus on best practice for national studies of hospital mortality, concerning a measure (the risk-adjusted HSMR factors to be included in risk-adjustment models, modelling methods, and types of cases to exclude (e.g. palliative-care cases). While discussion continues on the adequacy of administrative data for measuring in-hospital mortality, administrative data from good-quality systems appear to be adequate. In-hospital mortality rates are now reported regularly and publicly in several countries or jurisdictions within countries (United Kingdom, the Netherlands, Canada, and Queensland, Australia).

We applied two models: the most widely used approach, the RACM model, and the better-performing ERM model. This demonstrates that national indictors of in-hospital mortality can be produced using the Australian NHMD, and that findings have statistical properties similar to those reported elsewhere.

A longitudinal study of 3 years of data—following the approach of a recently-reported national study of in-hospital mortality in the Netherlands—provides evidence suggesting that although some unexplained variation in risk-adjusted HSMRs remains after modelling, Australian administrative data provide a strong 'signal' related to hospital-specific values. The findings were similar to those reported for the Netherlands.

Although further work is required to confirm the findings of this project, to elaborate them (e.g. to review and refine specifications for indicator case inclusion) and to extend them to issues that we could not deal with in this study (e.g. data linkage to include deaths within 30 days), it appears that Australian hospital data—like data from Canada, England and the Netherlands—can be used to measure risk-adjusted in-hospital mortality.

Variations in hospital mortality appear to fulfil the necessary criteria to qualify as a performance measure. The questions that remain are exactly which indicators, used in exactly which ways.

The literature review pointed to the continuing uncertainty concerning the relationship between variations in hospital mortality and other measures of hospital structure and process. This does not argue against the use of mortality-based indicators. In our view, it does mean that variations in hospital mortality measures should be viewed as screening tests. High or rising HSMRs should not be assumed to be definitively diagnostic of poor quality or safety. Nor should low or declining HSMRs be assumed to mean that all is well. Such results produced by a screening tool signal that further investigation is warranted to understand "What goes on here?'