

The hospital standardised mortality ratio: a powerful tool for Dutch hospitals to assess their quality of care?

B Jarman,¹ D Pieter,² A A van der Veen,³ R B Kool,² P Aylin,¹ A Bottle,¹ G P Westert,⁴ S Jones^{5,6}

¹Dr Foster Unit, Faculty of Medicine, Imperial College, London, UK ²Prismant, Utrecht, The Netherlands ³de Praktijk Index, Ondiep Zuidzijde, Utrecht, The Netherlands ⁴RIVM National Institute for Public Health and the Environment and Tranzo, Tilburg University, Bilthoven, The Netherlands ⁵Division of Health and Social Care Research, James Clerk Maxwell Building, King's College London, London, UK ⁶Dr Foster Intelligence, London, UK

Correspondence to

Brian Jarman, Dr Foster Unit, Faculty of Medicine, Imperial College London EC1A 9LA, UK; b.jarman@imperial.ac.uk

Accepted 21 July 2009

ABSTRACT

Aim of the study To use the hospital standardised mortality ratio (HSMR), as a tool for Dutch hospitals to analyse their death rates by comparing their risk-adjusted mortality with the national average.

Method The method uses routine administrative databases that are available nationally in The Netherlands—the National Medical Registration dataset for the years 2005–2007. Diagnostic groups that led to 80% of hospital deaths were included in the analysis. The method adjusts for a number of case-mix factors per diagnostic group determined through a logistic regression modelling process.

Results In The Netherlands, the case-mix factors are primary diagnosis, age, sex, urgency of admission, length of stay, comorbidity (Charlson Index), social deprivation, source of referral and month of admission. The Dutch HSMR model performs well at predicting a patient's risk of death as measured by a c statistic of the receiver operating characteristic curve of 0.91. The ratio of the HSMR of the Dutch hospital with the highest value in 2005–2007 is 2.3 times the HSMR of the hospital with the lowest value.

Discussion Overall hospital HSMRs and mortality at individual diagnostic group level can be monitored using statistical process control charts to give an early warning of possible problems with quality of care. The use of routine data in a standardised and robust model can be of value as a starting point for improvement of Dutch hospital outcomes. HSMRs have been calculated for several other countries.

In recent years, there has been an increasing interest in monitoring standards of clinical care in many countries. In the UK, the Bristol Royal Infirmary Inquiry into paediatric cardiac surgery deaths from 1999 to 2001¹ raised national awareness of the subject. In The Netherlands, an analysis of the death rates in cardiac surgery at the Radboud University by the Health Inspectorate led, in 2006, to a temporary six-month closure of the cardiac surgical department.

Mortality is a “hard” outcome with special relevance to the patient. Measuring death rates has the advantage that death is a definite unique event unlike morbidity, which often represents a spectrum of severity and can be difficult to record accurately. Death rates can, when adjusted for the factors that affect death rates, act as markers of a hard outcome of healthcare.

In England, the hospital standardised mortality ratio (HSMR), an overall measure of in-hospital

mortality, has been used since 1999.² About 67% of English acute hospital trusts nowadays use Real Time Monitoring (RTM)³ for monitoring and analysing HSMRs and their component diagnosis-level SMRs in order to deploy possible patient safety improvements. RTM makes the data available, updated monthly, to hospitals via the internet. HSMRs have also been calculated for the USA, Canada, Sweden, Wales, Australia (New South Wales), France, Japan, Hong Kong and Singapore and could be used to assess mortality, identify areas for possible improvement and monitor performance over time.

Until now, Dutch mortality figures, as measures of outcome of hospital care, were based on clinical databases and related to certain patient groups or procedures—for example, intensive care admissions,⁴ high-risk surgery⁵ and elderly patients.⁶ In the Dutch healthcare system, assessment of quality by calculating HSMRs has attracted considerable attention from government, patient organisations and the media. A study estimated that every year, more than 1700 avoidable deaths occur in Dutch hospitals.⁷ Following this study, a national patient safety programme was launched in 2007 by the associations of hospitals, medical specialists and nurses aimed at reducing the number of avoidable deaths. Monitoring the quality of hospitals within this programme by measuring HSMRs is one of the tools being used.

Two research organisations—Prismant and De Praktijk Index—developed, with Jarman and colleagues from Imperial College London, and Dr Foster Intelligence, a model to calculate HSMRs using data from the National Medical Registration (LMR) files, which contain all inpatient and day case admissions to hospitals. During the last years of calculating HSMRs for Dutch hospitals⁸ and using them for improving quality of care, several questions have been raised. The Dutch Minister of Health has announced that all Dutch hospitals should publish their HSMR in 2010.⁹ This article is intended to explain the current Dutch model and its statistical performance. It is hoped that it may be of assistance to hospitals by helping them to understand the method and that it may be of use for monitoring improvements in the quality of care for their patients.

METHODS

The HSMR compares the actual number of hospital deaths with the expected number for those patients with a primary diagnosis within the set of diagnostic groups that account for 80% of all deaths in hospital nationally.



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://qshc.bmj.com/site/about/unlocked.xhtml>.

The national LMR dataset for 2005–2007 was used as the data source for the logistic regression calculations. The HSMRs were calculated for 2005–2007. In the LMR dataset, diagnoses are coded using the International Classification of Diseases, Ninth Revision (ICD-9), and these are converted to 259 Clinical Classification System (CCS) groups developed by the US Agency for Healthcare Research and Quality.¹⁰ From these CCS groups, those responsible for 80% of hospital deaths nationally were determined. Day cases (which have very few deaths) and inpatient admissions were included in the analysis. Logistic regression models were fitted for each of the CCS groups separately in order to generate an expected risk of death for each patient. The HSMR is derived from the sum of the observed deaths and expected risks across the CCS groups.

Details of Dutch HSMR calculations

1. The 2005–2007 LMR data were used to form the model. These were data made available by Prismant, with permission of the Dutch Hospital Association (NVZ) and the Dutch Association of Medical Specialists. Seventeen thousand fifty-six ICD-9 codes in the Dutch hospital data were assigned to the 259 US Agency for Healthcare Research and Quality CCS groups.
2. After removing vague or undetermined diagnoses, the 50 CCS groups that give rise to 80% of all deaths in 2005–2007 were determined (table 1). Patients with lengths of stay under one year were used. The 50 CCS diagnoses covering 159 987 deaths were used for the model. The reported HSMR is then calculated using data from 2005 to 2007.
3. The calculation for non-average hospitals, hospitals with a case mix very different from the national average, which were excluded, was done by:
 - a. calculating the percentage of expected deaths nationally for each of the diagnostic groups making up the HSMR (leading to 80% of all deaths nationally);
 - b. calculating as in (a) for each hospital;
 - c. scale up or down the number of expected deaths by a scaling factor (Sf) for each diagnostic group to make the percentage of expected deaths at each hospital the same as the national %;
 - d. scale the observed deaths at each hospital by the same scaling factor for each diagnostic group;
 - e. use the scaled values of the numbers of observed and expected deaths at each hospital to calculate a “scaled HSMR”;
 - f. calculate the difference, D, between the normal (unscaled) HSMR and the scaled HSMR and
 - g. for the “average”, or non-specialist, hospitals’ D tends to be less than 7.5 for the hospitals that are not specialist hospitals.
4. Patients’ age was determined from the date of admission—date of birth age groups used were those for the English Hospital Episode Statistics (HES) (ie, <1 year=1, 1–4 years=2, 5–9 years=3, 10–14 years=4, 15–19 years=4, 20–24=6, 25–29 years=7, 30–34 years=8, 35–39 years=9, 40–44 years=10, 45–49 years=11, 50–54 years=12, 55–59 years=13, 60–64 years=14, 65–69 years=15, 70–74 years=16, 75–79 years=17, 80–84 years=18, 85–89 years=19, 90+ years=20)
5. The number of days of care was coded into length of stay (LOS) categories: 1 day=1; 2–7 days=2; 8–16 days=3; 17–23 days=4; 24–1000 days=5 (but only LOS to 365 was used in the data analysis).
6. The age group, sex, urgency, LOS group, CCS diagnosis, month of admission, social deprivation and year categories were determined for each patient.

7. The source of referral for each patient was coded as 0=own habitat; 1=nursing/elderly home; 2=born in hospital; 21=hospital—academic/top clinical; 22=hospital—general; 23=hospital specialised; 24=other care organisations; 29=hospital—unknown.

The statistical performance of the model was measured by the c statistic (area under the receiver operating characteristic curve) for each SMR and for the hospital level HSMR.¹¹ The c statistic is the probability of assigning a greater risk of death to a randomly selected patient who died compared with a randomly selected patient who survived. A value of 0.5 suggests that the model is no better than random chance in predicting death. A value of 1 suggests perfect discrimination. In general, values above 0.75 suggest good discrimination.

RESULTS

In the 2005–2007 data and for the HSMR CCS groups only, there were 2 363 332 admissions and 90 873 deaths (crude death rate 3.85%). The quality of the data of 15 hospitals did not fulfil the national registration standards in 2007, so we did not include them in a national comparison. Seven out of these 15 hospitals did not fulfil the standard for two or more criteria. Six of the 15 hospitals had more than 5% vague diagnosis, eight hospitals had less than 33% urgent admissions and ten hospitals had a ratio of comorbidity diagnosis to main diagnoses of <0.2. Another six hospitals were excluded because they had a patient population that differed too much from the national average. Four of these hospitals had less than 100 expected deaths in 2007, and finally two more hospitals were excluded because they are non-average hospitals in terms of their case mix. A funnel plot of the HSMRs of the remaining 65 hospitals is shown in figure 1.

Hospitals in figure 1 that lie within the control limits are said to exhibit common cause variation and those outside special cause variation unlikely to be due to natural random variation (in Shewhart’s original terminology). Funnel plots provide a simple and easily understandable way to plot institutional comparisons.¹² They have been used to plot anonymised mortality rates by surgeon for paediatric cardiac surgery¹³ and have been promoted as providing a strong visual indication of divergent performance, with the advantage of displaying actual event rates and allowing an informal check of a relationship between outcome and volume of cases.¹⁴

Dutch HSMRs differ widely among hospitals. According to this analysis, the chance of death in the hospital with the highest HSMR is 2.3 times the chance of dying in the hospital with the lowest HSMR, after adjusting for available case-mix factors.

The c statistic of the Dutch HSMR model is 0.91, similar to the values found for the other countries. Table 1 shows also the c statistics of all CCS groups: they vary from 0.68 to 0.96.

Significant factors determining the total hospital mortality were: primary diagnosis, age, sex, admission urgency (urgent/not-urgent, equivalent to emergency/elective (planned)), LOS, comorbidity (measured by the Charlson Index),¹⁵ area-level social deprivation (from the Dutch Central Office of Statistics), month of admission, type of organisation that made the referral and the CCS subgroup. These factors and their coefficients vary among each CCS group. Table 1 gives the significant factors ($p<0.05$) for every CCS group.

DISCUSSION

HSMRs have been calculated for The Netherlands in a manner similar to that used in several other countries. Currently, almost every Dutch hospital has asked for their HSMR without any pressure from the government or Healthcare Inspectorate. In addition, more than 50 hospitals have ordered

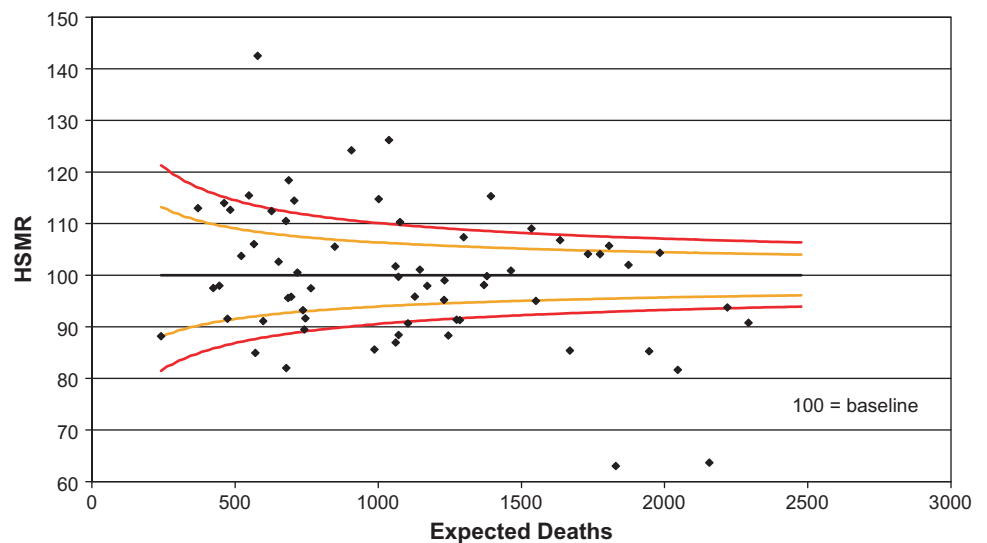
Table 1 CCS groups included in the model with their c statistics and relevant variables

Group	C statistic	Age	Charlson	Deprivation	LOS	Month	Sex	Source organisation type	CCS subgroup	Urgency	Year
Septicemia (except in labour)	0.827	Yes	Yes	Yes	Yes		Yes	Yes	Yes		Yes
Cancer of oesophagus	0.840	Yes	Yes		Yes	Yes		Yes		Yes	Yes
Cancer of stomach	0.811	Yes	Yes		Yes		Yes			Yes	Yes
Cancer of colon	0.857	Yes	Yes	Yes	Yes		Yes	Yes		Yes	Yes
Cancer of rectum and anus	0.858	Yes	Yes	Yes	Yes		Yes	Yes		Yes	
Cancer of pancreas	0.776	Yes	Yes	Yes	Yes		Yes			Yes	Yes
Cancer of bronchus, lung	0.873	Yes	Yes	Yes	Yes		Yes	Yes		Yes	Yes
Cancer of breast	0.957	Yes	Yes		Yes					Yes	
Cancer of prostate	0.925	Yes	Yes	Yes	Yes					Yes	
Cancer of bladder	0.939	Yes	Yes	Yes	Yes			Yes		Yes	Yes
Non-Hodgkin's lymphoma	0.923	Yes	Yes	Yes	Yes		Yes	Yes		Yes	Yes
Leukaemias	0.930	Yes	Yes	Yes	Yes	Yes		Yes		Yes	Yes
Secondary malignancies	0.908	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes
Neoplasms of unspecified nature or uncertain behaviour	0.916	Yes	Yes		Yes		Yes	Yes		Yes	Yes
Diabetes mellitus with complications	0.848	Yes	Yes		Yes		Yes		Yes	Yes	Yes
Fluid and electrolyte disorders	0.807	Yes	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Deficiency and other anaemia	0.911	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	
Coma, stupor and brain damage	0.728	Yes	Yes		Yes		Yes	Yes		Yes	Yes
Heart valve disorders	0.809	Yes	Yes	Yes	Yes			Yes	Yes	Yes	
Acute myocardial infarction	0.782	Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes	Yes
Coronary atherosclerosis and other heart disease	0.832	Yes	Yes	Yes	Yes			Yes	Yes	Yes	Yes
Pulmonary heart disease	0.798	Yes	Yes		Yes			Yes		Yes	Yes
Cardiac dysrhythmias	0.874	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cardiac arrest and ventricular fibrillation	0.809	Yes	Yes		Yes			Yes		Yes	Yes
Congestive heart failure, non-hypertensive	0.677	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Acute cerebrovascular disease	0.775	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Peripheral and visceral atherosclerosis	0.906	Yes	Yes		Yes		Yes	Yes	Yes	Yes	Yes
Aortic, peripheral and visceral artery aneurysms	0.866	Yes	Yes		Yes				Yes	Yes	Yes
Aortic and peripheral arterial embolism or thrombosis	0.880	Yes	Yes		Yes		Yes	Yes	Yes	Yes	
Other circulatory disease	0.862	Yes	Yes		Yes		Yes	Yes		Yes	Yes
Pneumonia	0.810	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Chronic obstructive pulmonary disease and bronchiectasis	0.778	Yes	Yes		Yes	Yes		Yes	Yes	Yes	Yes
Aspiration pneumonitis, food/vomit	0.718	Yes	Yes		Yes	Yes		Yes			Yes
Pleurisy, pneumothorax, pulmonary collapse	0.834	Yes	Yes	Yes	Yes			Yes	Yes	Yes	
Other lower respiratory disease	0.877	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes
Intestinal obstruction without hernia	0.831	Yes	Yes		Yes			Yes	Yes	Yes	Yes
Diverticulosis and diverticulitis	0.903	Yes	Yes		Yes		Yes	Yes	Yes	Yes	Yes
Biliary tract disease	0.920	Yes	Yes		Yes		Yes	Yes	Yes	Yes	
Liver disease, alcohol-related	0.728	Yes	Yes		Yes			Yes		Yes	Yes
Other liver diseases	0.843	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes
Gastrointestinal haemorrhage	0.812	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes
Other gastrointestinal disorders	0.943	Yes	Yes		Yes		Yes	Yes	Yes	Yes	Yes
Acute and unspecified renal failure	0.777	Yes	Yes		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Chronic renal failure	0.881	Yes	Yes		Yes			Yes		Yes	Yes
Urinary tract infections	0.880	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes
Fracture of neck of femur (hip)	0.782	Yes	Yes		Yes		Yes	Yes		Yes	Yes
Intracranial injury	0.884	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Complication of device, implant or graft	0.858	Yes	Yes		Yes		Yes	Yes	Yes	Yes	Yes
Complications of surgical procedures or medical care	0.873	Yes	Yes		Yes		Yes	Yes	Yes	Yes	Yes
Shock	0.802	Yes	Yes	Yes	Yes						Yes
Average ROC curve	0.845										
Overall ROC curve	0.910										

All models included an intercept term.

CCS, Clinical Classification System; LOS, length of stay; ROC, receiver operating characteristic.

Figure 1 Funnel plot showing HSMR variation, 2005–2007 in Dutch hospitals (excluding 24 hospitals) with 95% and 99.8% control limits.



a “Hospital Mortality Profile” over the last two years—a brief report giving the HSMR of a hospital broken down into its constituent diagnostic group SMRs and by age group, urgency and length of stay. The following applications of HSMRs are used in Dutch hospitals:

- ▶ With the Hospital Mortality Profile to identify high and low risk “areas” within the hospital. Such a retrospective profile enables more directed intervention for patient safety.
- ▶ Dr Foster’s RTM—a tool used in 16 hospitals for early warning, continuous monitoring and analysis of their mortality by diagnosis and procedures using the same risk models underpinning the HSMR. Hospitals use this tool to follow their own progress in decreasing patient safety risks.
- ▶ Some hospitals use HSMRs in combination with clinical audits. They drill down to the level of the mortality risk of individual patients admitted. By doing so, hospitals can select “unexpected cases”. These are patients who die in the hospital but have a relatively low risk of dying in hospital. These cases are perhaps the most useful for case note review and complication analysis and can aid improvement initiatives.

Our analysis of data completeness found no missing values of the date of admission, date of discharge, age, sex, urgency of admission or postal code (for social deprivation). However, for the recording of secondary diagnosis in particular, we cannot tell whether there is no comorbidity present or if comorbidity has simply not been recorded. Miscoding may also affect the HSMR.

The LMR data use a limited number of clinical variables but for the HSMRs examined in this study, the discrimination of the risk prediction model was very good. A recent UK study concluded that, at least for three common procedures, risk prediction with discrimination comparable with that obtained from clinical databases is possible using routinely collected administrative data.¹⁶

Although simplified models of risk prediction might be as effective in predicting outcome as some complex models currently in use,^{17 18} further improvements to the case-mix model are being evaluated. The numbers of previous admissions within a given time period, which requires the linking of admissions of the same patient, could be of potential use. Other features of the healthcare system that could potentially affect hospital mortality ratios include admission thresholds, the proportion of people in the area dying in hospital, discharge policies or underlying disease rates in the catchment population.

It is unclear, however, whether and how one should measure and adjust for these factors.

A relevant discussion is also whether the length of stay and the procedure group are factors that are part of the case mix or determine quality. Both are related to the patient’s illness but also to treatment.

Based on experience in other countries, the introduction of HSMRs raises various questions.^{19–22} Most recently, attention has been focused on the so-called “constant risk fallacy”²³ in which some SMRs—for example, for some Charlson scores, differ from the overall HSMR. One paper suggests at least two mechanisms that might contribute: the first involves differential measurement error, and the second involves inconsistent proxy measures of risk.²⁴ Measurement error, including poor coding, will have an impact on HSMRs, and this is the first thing that a hospital should check. The variation in SMRs can be interpreted in two ways, either as bias or as real differences in risk. Either way, further investigation using local data sources and case note reviews rather than more statistical analysis is suggested.

Another often heard query is that the methodology should correct for regional variation in health conditions or in the organisation and performance of healthcare facilities adjacent to the hospital. A multiple regression analysis has been developed for the Dutch HSMRs to find the factors that best explain the variation of HSMRs throughout The Netherlands.²⁵ Depending on the extension of the dataset, further yearly refinements can be made to the models for the yearly releases of the HSMRs and SMRs.

The HSMR for The Netherlands appears to be a statistically robust model that can be used as an indicator for hospital deaths to help Dutch hospitals improve their quality of care. The statistical model is robust enough to include all hospitals with more than about 100 deaths per year, an average case mix and good quality data, varying in size and function, into one analysis. However, random variation and data quality issues need to be considered when interpreting the results. HSMRs can be used to highlight hospitals that have significantly high mortality, which may merit further investigation by the hospitals concerned. Furthermore, the impact of interventions designed to reduce mortality can be tracked using this measure.

The Dutch Ministry of Health^{26 27} has put HSMR high on its quality agenda and commissioned RIVM (the National Institute for Public Health and the Environment) to use HSMRs as one of the performance indicators in the Dutch Health Care

Performance report. In the future, international comparisons might also be possible.

Acknowledgements We would like to thank Laurens Touwen who has been immensely helpful in developing the Dutch HSMRs since the start of the project in 2004. We also would like to thank the Dutch Hospital Association (NVZ) and the Federation of Medical Specialists (de Orde) for granting permission to use the Dutch hospital data.

Contributors All the authors have been involved in writing, reading and commenting on the manuscript. BJ, DP, PA, AB and SJ were involved in developing the model. BJ is the guarantor of the content.

Funding There was no separate funding for this study. BJ, AB and PA are employed within the Dr Foster Unit at Imperial College London (BJ, part-time), which is partly funded by a grant from Dr Foster Intelligence (an independent health service research organisation). The unit is also partly funded for its HSMR work by the Rx Foundation of Cambridge, Massachusetts, USA, and is affiliated with the Centre for Patient Safety and Service Quality at Imperial College Healthcare NHS Trust, which is funded by the National Institute of Health Research. The Department of Primary Care and Social Medicine is grateful for support from the National Institute for Health Research Biomedical Research Centre Funding Scheme.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- Learning from Bristol: the report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984–1995. <http://www.bristol-inquiry.org.uk/>
- Jarman B, Gault S, Alves B, *et al.* Explaining differences in English hospital death rates using routinely collected data. *BMJ* 1999;**318**:1515–20.
- Bottle A, Aylin P. Intelligent Information: a national system for monitoring clinical performance. *Health Serv Res* 2008;**43**:10–31.
- De Jonge E, Bosman RJ, van der Voort PH, *et al.* Intensive care geneeskunde in Nederland 1997–2001. I. Patientenpopulatie en behandelresultaten [Patient populations and results of treatment]. *Ned Tijdschr Geneesk* 2003;**147**:1013–7.
- Obertop H. Defening baart kunst. Gunstige effecten van ervaring op behandelresultaten [Practice makes perfect. Favourable effects of routine on results of treatment]. *Ned Tijdschr Geneesk* 2004;**148**:1327–9.
- Beenen E, Simons MP, Vahl AC. Determinanten van ziekenhuissterfte bij chirurgische patiënten van 80 jaar en ouder [Decisive elements of hospital mortality of surgical patients of 80 years and older]. *Ned Tijdschr Geneesk* 2003;**147**:1915–8.
- de Bruijne MC, Zegers M, Hoonhout LHF, *et al.* *Onbedoelde schade in Nederlandse ziekenhuizen* [Adverse Events in Dutch Hospitals]. Amsterdam: EMGO Instituut/VUmc en NIVEL, Nederlands Instituut voor onderzoek van de gezondheidszorg, 2007.
- Hospital Standardized Mortality Ratio. <http://nl.wikipedia.org/wiki/HSMR> (accessed 9 January 2010).
- Parlando (Dutch Parliament). Zoekresultaat. <http://parlando.sdu.nl/cgi/login/anonymous> (accessed 9 January 2010).
- Elixhauser A, Andrews RM, Fox S. *Clinical classifications for health policy research: discharge statistics by principal diagnosis and procedure. Provider Studies Research Note 17*. Rockville, Maryland: Agency for Health Care Policy and Research, 1993. AHCPR Pub. No. 93-0043. <http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>.
- Pepe MS. *The statistical evaluation of medical tests for classification and prediction*. New York: Oxford, 2003.
- Spiegelhalter D. Funnel plots for comparing institutional performance. *Stat Med* 2005;**24**:1185–202.
- Stark J, Gallivan S, Lovegrove J, *et al.* Mortality rates after surgery for congenital heart defects in children and surgeon's performance. *Lancet* 2000;**355**:1004–7.
- Spiegelhalter D. Funnel plots for institutional comparison. *Qual Saf Health Care* 2002;**11**:390–1.
- Charlson ME, Pompei P, Ales KL, *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseases* 1987;**40**:373–83.
- Aylin P, Bottle A, Majeed A. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *BMJ* 2007;**334**:1044.
- Sutton R, Bann S, Brooks M, *et al.* The Surgical Risk Scale as an improved tool for risk-adjusted analysis in comparative surgical audit. *British Journal of Surgery* 2002;**89**:763–8.
- Jenkins J, Gauvrau K, Newburger J, *et al.* Consensus based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg* 2002;**123**:110–18.
- Iezzoni LI. The risks of risk adjustment. *JAMA* 1997;**278**:1600–7.
- Penfold RB, Dean S, Flemons W, *et al.* Do hospital standardised mortality ratios measure patient safety? HSMRs in the Winnipeg Regional Health Authority. *Healthc Pap* 2008;**8**:8–24.
- Wright J, Shojania KG. Measuring the quality of hospital care. *BMJ* 2009;**338**:b569.
- Godlee F. Measuring quality. *BMJ* 2009;**338**:b1356.
- Nicholl J. Case-mix adjustment in non-randomised observational evaluations: the constant risk fallacy. *J Epidemiol Community Health* 2007;**61**:1010–13.
- Mohammed MA, Deeks JJ, Girling A, *et al.* Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. *BMJ* 2009;**338**:b780.
- Heijink R, Koolman X, Pieter D, *et al.* Measuring and explaining mortality in Dutch hospitals; the hospital standardised mortality rate between 2003 and 2005. *BMC Health Serv Res* 2008;**8**:73.
- Westert GP, van den Berg MJ, Koolman X, *et al.*, eds. *Dutch health care performance report 2008*. Houten: Bohn Stafleu en Van Loghum, 2008.
- Westert GP, *et al.* *Dutch health care performance report*. <http://www.healthcareperformancereport.nl/> (accessed 22 June 2009).