

Hospital standardised mortality ratio: A reliable indicator of quality of care?

J.A. van Erven¹, L.S. van Galen¹, A.A. Hettinga-Roest¹, E.P.J. Claessens¹,
J.C. Roos², M.H.H. Kramer¹, P.W.B. Nanayakkara^{1*}

Departments of ¹Internal Medicine, Section Acute Medicine, ²Bureau of Medical Affairs, VU University Medical Centre, Amsterdam, the Netherlands, *corresponding author: email: p.nanayakkara@vumc.nl

ABSTRACT

Background: This study investigates (1) whether the hospital standardised mortality ratio (HSMR) model underestimates or overestimates disease severity and (2) the completeness of the data collected by administrators to calculate HSMR in a cohort of deceased patients with the diagnosis of pneumonia.

Methods: In this retrospective cohort study Pneumonia Severity Index (PSI) and Abbreviated Mortality in Emergency Department Sepsis (abbMEDS) scores and associated mortality probabilities were obtained from 32 deceased pneumonia patients over the year 2014 in the VU University Medical Centre. These were compared with mortality probabilities of the Central Bureau for Statistics (CBS) calculated for every patient using the HSMR model. Clinical charts were examined to extract relevant comorbidities to determine the reliability of data sent to the national registration of hospital care.

Results: Risk categories determined by using the PSI and the abbMEDS were significantly higher compared with the risk categories according to HSMR ($p = 0.001$, respectively $p = 0.000$). The mean difference between the number of comorbidities in our registration and the coders' registration was 1.97 ($p = 0.00$). The mean difference was 0.97 ($p = 0.000$) for the number of comorbidities of influence on the Charlson Comorbidity Index (CCI) and 1.25 ($p = 0.001$) for the calculated CCI.

Conclusion: The results of this study suggest that the mortality probabilities as calculated by the CBS are an underestimation of the risk of dying for each patient. Our study also showed that the registration of data sent to the CBS underestimated the actual comorbidities of the patient, and could possibly influence the HSMR.

KEYWORDS

Data registration, disease severity, HSMR, patient outcomes, quality indicator

INTRODUCTION

Since March 2014, Dutch hospitals are obliged to be transparent about their mortality rates.¹ To be able to compare the quality of hospital care using their mortality rate, these rates have to be standardised in order to correct for the differences in the case-mix.² This standardised ratio is represented in the hospital standardised mortality ratio (HSMR) and is the ratio of the observed to the expected deaths, derived from data from the national registration of hospital care, the LBZ.³ The expected deaths are calculated with the use of a statistical model that corrects for certain factors such as age, socioeconomic status and comorbidity.³ In 2014, this model contained standardised mortality ratios (SMRs) for 50 diagnosis groups, which account for 80% of in-hospital deaths. This was extended to SMRs for 157 diagnosis groups in 2015.

Over the year 2014, the VU Medical Centre, Amsterdam (VUmc) had a relatively high HSMR, in part caused by a high SMR for the diagnosis group 'pneumonia'. The SMR of a diagnosis group can be used to investigate the cause of unexpected high mortality in a hospital more specifically than by solely using the HSMR.⁴ For this reason a commission of independent external investigators in the VUmc were asked to investigate this high SMR. The aim was to investigate whether preventable/avoidable factors contributed to these deaths. Their report showed no avoidable causes of death in this cohort. These findings

suggest that the cause of the high SMR for pneumonia is probably due to other unknown factors. It could, for example, be caused by insufficient registration of comorbidities or wide variations in disease severity. In the clinical setting, physicians and nurses use several different scoring systems to determine the severity and to predict the mortality of pneumonia using patient characteristics such as age, blood urea and respiratory rate. Two of the best-validated and most used scoring systems are the Pneumonia Severity Index (PSI)⁵ and the Abbreviated Mortality in Emergency Department Sepsis score (abbMEDS).⁶

The HSMR is calculated by the Central Bureau of Statistics (CBS) and the data used for this calculation are registered by Dutch Hospital Data (DHD) within the context of the LBZ. The Medical Administration Office of each hospital provides the information that is used. The HSMR is, among other covariates, derived from the primary diagnosis and the Charlson Comorbidity Index (CCI),⁷ which are obtained from patients' charts and documented by coders. This underlines the importance of a complete administration, as deficient or faulty data might directly influence the HSMR. Van der Laan et al. (2013)⁸ showed that the effect of registering 10% more comorbidities could result in a decrease of 5 points of the HSMR.⁸ Although the administration of data has improved significantly since the implementation of the HSMR as an indicator of quality of care, there still might be inconsistencies in the comorbidity data extracted by coders and registered by DHD, when compared with the actual data extracted by doctors from the patients' charts.^{8,9}

Therefore, the main aim of this study was to examine whether the HSMR model underestimates or overestimates the disease severity of pneumonia patients when compared with routinely used clinical severity scores. Our secondary aim was to investigate the completeness of the data sent to DHD to calculate the HSMR.

MATERIAL AND METHODS

In 2014, 32 deceased patients were registered in the 'pneumonia' group at the VUmc. In order to obtain PSI and abbMEDS scores for these patients, patients' charts were examined for information needed to calculate these scores from which corresponding mortality probabilities could be calculated. Missing information was considered as not contributing to the score.

The HSMR is calculated by logistic regression using the below-mentioned covariates with data provided by hospital coders. With this information, regression coefficients for these covariates are estimated and are used to calculate mortality probabilities for each individual admission.³ The results of the calculations are sent to each hospital in the annual HSMR report.

The HSMR is calculated using the following covariates³:

- Age at admission
- Sex
- Socioeconomic status (SES) of the postal area of the patient's address. The SES classification per postal code is compiled by the Netherlands Institute for Social Research (SCP)
- Severity of main diagnosis. Instead of CCS diagnosis subgroups (Clinical Classifications Software: a tool to cluster patient diagnoses into a manageable number of clinically meaningful categories, based on the International Classification of Diseases. The CCS makes little distinction in regard to disease severity when categorising diagnosis codes), a classification of severity of the main diagnosis in terms of mortality rates is used, as suggested by Van den Bosch et al. (2011)¹⁰
- Urgency of admission (elective, acute)
- Comorbidity (17 comorbidity groups of the Charlson Comorbidity Index⁷)
- Source of admission (home, nursing home or other institution, hospital)
- Year of discharge
- Month of admission

In order to compare the mortality probabilities derived from the PSI and abbMEDS scores (which correspond with ordinal risk categories) and the mortality probabilities calculated by the CBS (which can be considered a continuous variable), new categories needed to be formed for the latter. It was decided to form three sets of categories from the CBS data, one for each of the scores. *Table 1* shows the risk categories and corresponding mortality probabilities of the three scoring systems. The consensus was that the best way to establish limits for new categories was by using the median between each of the mortality probabilities, as those are the means of that risk category. As can be seen in *table 1*, the lowest risk categories of the PSI predict a risk of 0.1% and of 0.6%. The median between these risks is 0.35, therefore, the limits of the

Table 1. Risk categories and corresponding mortality probabilities of the scoring systems

PSI ⁵		abbMEDS ^{6,13}	
Low risk I	0.1%	Low risk	3.6%
Low risk II	0.6%	Intermediate risk	19.5%
Low risk III	0.9%	High risk	46.2%
Medium risk	9.5%		
High risk	26.7%		

The mortality probabilities of the risk categories for the abbMEDS score are derived from a study by Roest et al. (2015)¹³

PSI categories used are 0-0.35; 0.35-0.75; 0.75-5.2; 5.2-18.1; 18.1-100 and the limits for the abbMEDS are 0-11.55; 11.55-32.85; 32.85-100.

The newly formed categories of the CBS calculated mortality probabilities were compared with the categories of the PSI and abbMEDS scores. A Wilcoxon sign-rank test was used for statistical analysis to test for conformity.

To investigate whether data sent to DHD significantly differed from what is found in patients' charts, data were gathered on the total amount of comorbidities that were present in charts, which of these were directly of influence to the CCI (excluding the comorbidities that are not in the Charlson Comorbidity Index) and finally the estimated CCI by the hospital itself. The coders in VUmc primarily look at the discharge letter and only broaden their scope when they presume this to be insufficient. In this study one researcher (JVE) thoroughly checked every patient's chart which included the discharge letter. If there was any uncertainty concerning a possible comorbidity or diagnosis, a second researcher (PN) was consulted and consensus was reached. The data that the CBS used were obtained from the Medical Administration Office. A paired t-test was used to analyse the difference between our registration and the coders' registration. For all analyses, a two-tailed p-value of less than 0.05 was considered statistically significant.

RESULTS

Table 2 gives an overview of the patient characteristics of our population. Ten patients had a cause of death other than respiratory failure or sepsis.

Mortality probabilities

Table 3 illustrates the dispersion of mortality probabilities calculated by the CBS using the HSMR model and those of the two clinical scoring systems. It can be seen that for the majority of patients the estimated risk of dying within 28-30 days is much higher according to the clinical scoring systems than the estimated risk of dying as calculated using the SMR model. Especially the abbMEDS assesses the risk to be significantly higher than the CBS does. In our cohort of patients, the abbMEDS seemed to estimate the severity of pneumonia the best. This is why we categorised the table according to the risk categories of the abbMEDS.

Descriptive statistics of conformity were performed and this showed that for the PSI 18 patients were in a higher risk category than according to the CBS (SMR), 3 were in a lower category and 11 were in the same category. When looking at the abbMEDS, all patients were either in the same risk category (10) or in a higher risk category (22) compared with SMR.

Further analysis showed a significant increase in assigned risk categories for the PSI ($p < 0.001$) and for the abbMEDS ($p < 0.001$) compared with the SMR. This indicates that risks of dying of these patients, according to clinical scoring systems, were significantly higher than the risks of dying according to SMR calculated by the CBS.

Registration of data

Figure 1 shows the number of comorbidities, the number of comorbidities influencing the CCI and the calculated CCI itself from our own registration and those same outcome measures which medical coders registered. For each of the outcome measures the mean of our registered number is higher than the mean of what the coders registered.

As table 4 shows, the mean difference between the number of comorbidities in our registration and the coders' registration is 1.97. The mean difference between our

Table 2. Demographic and clinical characteristics of the deceased patients in the 'pneumonia' group over the year 2014

Characteristics	Deceased patients (n = 32)	
	Number	Percentage
Demographic factor		
Age > 65 years	25	78.13
Female sex	12	37.5
Nursing home resident	6	18.75
Admissions		
≥ 2 in last 12 months	16	50
ICU admissions in last 12 months	14	43.75
Unexpectedly long admission*	6	18.75
Cause of death		
Respiratory failure	14	43.75
Sepsis	12	37.5
Myocardial infarction	5	15.63
Heart failure	4	12.5
Other	1	3.13
Other clinical characteristics		
Immunocompromised [†]	10	31.25
Do not resuscitate	26	81.25
Polypharmacy [*]	28	87.5
Limited mobility ^{&}	24	75
Delirium	8	25
Malnutrition [§]	12	37.5

Cause of death: In some patients the respiratory failure or heart failure was a direct result of sepsis.

* An admission minimally 50% longer than expected for a specific primary diagnosis. The calculation of the expected length of admission takes into account the age of the patient, primary diagnosis and any possible interventions.

[†] Immunodeficiency by the use of immunosuppressive drugs, by neutropenia or leukopenia or other causes.

^{*} The chronic use of ≥ 5 medications.

[&] Patient uses devices for mobility or was bedridden.

[§] Patient has a Short Nutritional Assessment Questionnaire (SNAQ) score of ≥ 2 or when the patient was described as cachexic.

Table 3. Mortality probabilities calculated by the CBS and those derived from the scoring systems, categorised according to the risk categories of the abbMEDS

Risk category abbMEDS	Patient no.	abbMEDS	PSI	Mortality probabilities CBS (SMR)
Low risk	1	3.60%	0.60%	0.87%
	11	3.60%	9.50%	8.76%
	28	3.60%	26.70%	2.89%
Intermediate risk	2	19.50%	26.70%	6.29%
	5	19.50%	9.50%	10.01%
	6	19.50%	26.70%	3.64%
	7	19.50%	26.70%	18.02%
	9	19.50%	9.50%	10.62%
	10	19.50%	26.70%	12.19%
	13	19.50%	9.50%	18.60%
	14	19.50%	0.90%	3.45%
	16	19.50%	26.70%	5.76%
	17	19.50%	26.70%	11.02%
	19	19.50%	26.70%	6.68%
	20	19.50%	9.50%	7.74%
	21	19.50%	9.50%	5.92%
	22	19.50%	9.50%	6.18%
	23	19.50%	26.70%	24.59%
	24	19.50%	26.70%	10.54%
	25	19.50%	9.50%	13.70%
	26	19.50%	9.50%	5.01%
	27	19.50%	9.50%	1.54%
30	19.50%	9.50%	12.87%	
32	19.50%	9.50%	21.24%	
High risk	3	46.20%	26.70%	14.70%
	4	46.20%	9.50%	6.43%
	8	46.20%	26.70%	15.33%
	12	46.20%	26.70%	18.05%
	15	46.20%	26.70%	5.39%
	18	46.20%	26.70%	9.34%
	29	46.20%	26.70%	6.32%
	31	46.20%	26.70%	13.37%

Green represents higher probability than calculated by the CBS (using the SMR-model), red represents a lower probability.

registration and the coders' registration for the CCI is 1.25. All of these results are statistically significant.

An unanticipated finding was that the source of admission was 'home' in every case. It seemed as though no distinction was made between 'home' and 'nursing home'. Nevertheless, table 1 shows that 6 out of 32 patients were admitted from a nursing home.

DISCUSSION

The findings in this paper indicate that (1) the SMR model appears to underscore the severity of pneumonia compared with the validated clinical scoring systems PSI and abbMEDS in a cohort of patients who died of pneumonia; and (2) the total number of comorbidities and the number of comorbidities influencing the CCI is higher according to our registration than according to the coders' registration.

The results in this study further support the suggestion that was made by Pleizier et al. that the SMR for more diagnosis groups besides cerebrovascular diseases will also decrease when adjusted for the severity of disease.¹¹ They concluded that within the SMR group 'cerebrovascular diseases' there is no distinction between 'stroke', 'cerebral haemorrhage' and 'subarachnoid haemorrhage' while their mortality rates differ greatly.¹¹ The mortality rates were 18, 43 and 35%, respectively, and when these differences were not taken into account, the influence on the SMR

Figure 1. Compared means and SD of the outcome measures extracted during our study versus what coders registered and the CBS used to calculate the HSMR

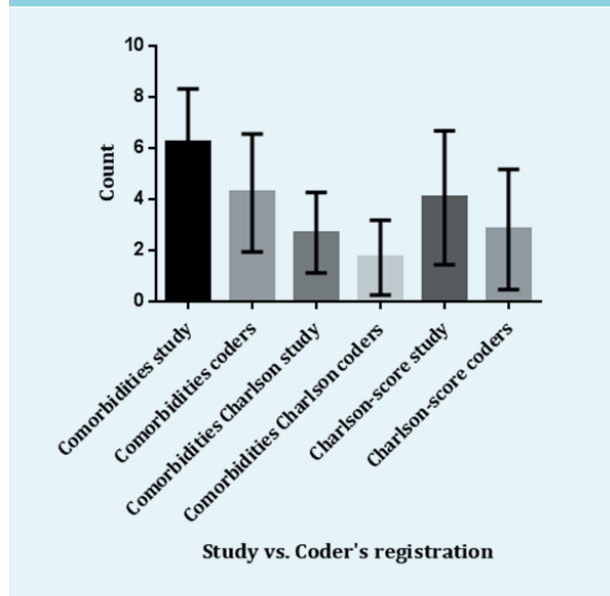


Table 4. Statistical analysis of the outcome measures

	Mean	SD	SE	95% CI	p-value
Comorbidities study – comorbidities CBS	1.96875	2.53345	.44785	1.05535-2.88215	.000
Charlson Index Study – Charlson index CBS	1.25000	1.95101	.34489	.54658-1.95342	.001

Outcome measures in the table are the number of comorbidities and the Charlson Comorbidity index. SD = standard deviation; SE = standard error of the mean.

could be considerable.¹² They recalculated the SMR for ‘cerebrovascular diseases’ after correcting for the above-mentioned sub-diagnoses and found that this gave a reduction from 119 (95% CI 105-133) to 102 (95% CI 91-115).¹¹ Beside this diagnosis group, this is possibly also true for other SMR groups such as the ‘pneumonia group’. A subdivision for ‘cerebrovascular diseases’ was easily made by just looking at the mortality rates for several subdiagnoses within that group. This, however, is a lot harder for a diagnosis group such as ‘pneumonia’, where there are no known distinct subdiagnoses. To make a subdivision for ‘pneumonia’, two different scoring systems were used that indicate severity of disease. The best way to prove that a subdivision by each of these scores has a direct effect on the SMR is by adjusting the SMR model in the same way Pleizier et al. did.¹¹ They incorporated a division in risk categories into the logistic regression model, just as the other covariates. In our study, it was decided to compare the mortality probabilities of the validated scores with the mortality probabilities calculated by the CBS with the use of the SMR model.

The results show that for the large majority of patients the expected mortality within 28-30 days is much higher according to the two scoring systems than to the score calculated by the CBS. This is probably partly caused by underscoring the number of comorbidities, but also a lack of proper adjustment for the severity of the disease pneumonia in the individual patient. These two scoring systems are widely used in clinical settings when dealing with pneumonia patients and have been validated.^{5,13} They are specifically designed to assess the severity of pneumonia/sepsis and should therefore be taken seriously as predictors of death. This suggests that the mortality probabilities according to the HSMR model of CBS are an underestimation of the real risk of dying for each patient. Naturally, estimating disease severity with the use of nine variables results in a simplification of reality. In addition, it is known that university medical centres predominantly provide tertiary care for a case-mix of patients with a higher severity of disease than peripheral hospitals. Therefore, they might falsely have a ‘higher’ HSMR.

Our results indicate that the mortality probabilities calculated by the PSI and the abbMEDS are higher than what the CBS calculated. It could be argued that the steps

between the risk categories of these scoring systems are fairly big. Therefore, when a patient is placed in the highest risk category of, for example, the abbMEDS their risk of dying could be even higher than 46.2%. However, *table 3* does compare categorical variables (the mortality probabilities calculated by the scoring systems) with a continuous variable (the mortality probabilities calculated by the CBS), which implies that these risks will almost always differ from the risks as calculated by the CBS.

The secondary aim of this study was to assess the registration of comorbidities from the patients’ charts by medical coders. As stated earlier, we discovered that the source of admission was ‘home’ in every case, although six of the patients in our cohort were admitted from a nursing home. This could potentially have an impact on the HSMR as a whole; however, this influence is probably rather small. It must be acknowledged that the source of admission is not primarily registered by coders, but they are responsible for checking this registration.

Van der Laan et al. already underlined the influence of the registered number of comorbidities on the HSMR.⁸ With this in mind, an average difference of two registered comorbidities seems significant enough to be of influence to the HSMR. For a comorbidity to be of influence to the HSMR it needs to add to the CCI; so, to make the previous assumption more likely, the CCI of every patient was also taken in consideration. It was found that there was a statistically significant difference of 1.25 points between the calculated CCI based on our registration and the calculated CCI based on the coders’ registration. This strongly suggests that the apparently insufficiently registered number of comorbidities does directly influence the HSMR. As stated earlier, coders are dependent on proper documentation by others, including doctors. They primarily look at the discharge letter and operation reports, and are not expected to go through the entire patient chart, mainly since this would be too time consuming. This lack of time might be one of the causes of the apparent under registration of comorbidities. One other cause explaining the under registration is that according to coding protocol, sometimes an International Classification of Diseases (ICD-10) code which has less impact on the HSMR than the actual diagnosis has to be selected. Although the precise impact cannot be judged by the results of this study, these findings do raise the question whether the

HSMR is reliable enough to estimate what it is supposed to do or to be published for everyone to see.

The limitations of this work must be acknowledged. The self-formed categories composed to compare categorical and continuous variables are merely based on what was thought to be the most logical way to do this. Hence, a note of caution is due here when interpreting these results. Also, in this study no control group was investigated. This withholds the opportunity to compare the mortality probabilities of the living patients with the deceased patients and therefore we were not able to investigate if the severity of disease was greater in the deceased group. Finally, it would have been interesting to calculate the HSMR/SMR using our calculated CCI and compare this with the HSMR/SMR calculated by the CBS. Unfortunately, computing our own logistic regression model to perform these calculations proved to be too time consuming.

CONCLUSION

Hospitals are obliged to publish their HSMRs, which gives patients and healthcare institutions the opportunity to judge and compare hospitals on the basis of this number. However, we demonstrated that differences in case-mix and the incompleteness of the data used to calculate the HSMR could negatively influence the HSMR. Although it seems quite logical to look at the number of deaths in each hospital as an indicator of quality of care, there are numerous pitfalls hidden in using the HSMR as a quality indicator. Therefore, HSMR should always be interpreted with caution and openly publishing HSMRs may have unfair negative consequences for some hospitals.

DISCLOSURES

All authors declare no conflict of interest. No funding or financial support was received.

REFERENCES

1. Hospital Standardised Mortality Ratio, de Praktijk Index, www.hsmr.nl. Accessed January 2017.
2. Understanding HSMRs. A Toolkit on Hospital Standardised Mortality Ratios. http://www.drfooster.com/wp-content/uploads/2014/09/HSMR_Toolkit_Version_9_July_2014.pdf. Published 2014. Accessed December 2016.
3. Van der Laan J, de Bruin A, van den Akker-Ploemacher J, Penning C, Pijpers F. HSMR 2014 methodological report, november 2015. http://www.hsmr.nl/wp-content/uploads/2016/01/2015hsmrmethodologic_alreport2014.pdf. Accessed November, 2016.
4. Jarman B, Pieter D, van der Veen AA, et al. The hospital standardised mortality ratio: a powerful tool for Dutch hospitals to assess their quality of care? *Qual Saf Health Care*. 2010;19:9-13.
5. Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. *N Engl J Med*. 1997;336:243-50.
6. Vorwerk C, Loryman B, Coats TJ, et al. Prediction of mortality in adult emergency department patients with sepsis. *Emerg Med J*. 2009;26:254-8.
7. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40:373-83.
8. Central bureau of statistics (Van der Laan J). Quality of the Dutch Medical Registration (LMR) for the calculation of the Hospital Standardised Mortality Ratio. 2013; ISSN: 1572-0314.
9. Tol J, Broekman M, Brauers M, van Gulik T, Busch OR, Gouma DJ. [Reliability of the registration of data on complex patients: effects on the hospital standardised mortality ratio (HSMR) in the Netherlands]. *Ned Tijdschr Geneeskd*. 2012;156:A4918.
10. Van den Bosch WF, Spreeuwenberg P, Wagner C. [Hospital standardised mortality ratio (HSMR): adjustment for severity of primary diagnosis can be improved]. *Ned Tijdschr Geneeskd*. 2011;155:A3299.
11. Plezier CM, Geerlings W, Pieter D, Boiten J. Patientmix influences HSMR. *Medisch Contact* 2010. 36:1777-9.
12. Bots ML, Jager-Geurts H, Berger-van Sijl M. Risk of dying after first hospital admission for a cerebrovascular accident in the Netherlands. *Cardiovascular disease in the Netherlands*. The Hague, 2006.
13. Roest AA, Tegtmeier J, Heyligen JJ, et al. Risk stratification by abbMEDS and CURB-65 in relation to treatment and clinical disposition of the septic patient at the emergency department: a cohort study. *BMC Emerg Med*. 2015;15:29.