Quality of the Dutch Medical Registration (LMR) for the calculation of the Hospital Standardised Mortality Ratio

Jan van der Laan

Discussion paper (201308)

Statistics Netherlands

The Hague/Heerlen, 2013

Explanation of symbols

.	data not available
*	provisional figure
**	revised provisional figure (but not definite)
x	publication prohibited (confidential figure)
–	nil
–	(between two figures) inclusive
0 (0.0)	less than half of unit concerned
empty cell	not applicable
2012–2013	2012 to 2013 inclusive
2012/2013	average for 2012 up to and including 2013
2012/'13	crop year, financial year, school year etc. beginning in 2012 and ending in 2013
2010/'11– 2012/'13	crop year, financial year, etc. 2010/'11 to 2012/'13 inclusive

Due to rounding, some totals may not correspond with the sum of the separate figures.

Publisher
Statistics Netherlands
Henri Faasdreef 312
2492 JP The Hague

Prepress
Statistics Netherlands
Grafimedia

Cover
Teldesign, Rotterdam

Information
Telephone +31 88 570 70 70
Telefax +31 70 337 59 94
Via contact form:
www.cbs.nl/information

Where to order
E-mail: verkoop@cbs.nl
Telefax +31 45 570 62 68

Internet
www.cbs.nl

ISSN: 1572-0314

60083201308 X-10

Quality of the Dutch Medical Registration (LMR) for the calculation of the Hospital Standardised Mortality Ratio

Jan van der Laan

Abstract: Since 2011 Statistics Netherlands is responsible for the calculation of the Hospital Standardised Mortality Ratio (HSMR) for the Dutch hospitals. Since there were some indications that there are quality issues with some of the variables used in the calculation of the HSMR, it was decided to investigate the quality of these variables. The variables investigated in this paper are the comorbidities and the urgency of the admission, i.e. whether the admission was planned or not. It is attempted to estimate the increase in variance in the HSMR caused by coding differences between hospitals in these variables. Although there are some indications that there are coding differences between hospitals for urgency that affect the HSMR, these differences appear to be small. For the comorbidities the coding differences appear to be much larger.

Keywords: Hospital Standardised Mortality Ratio (HSMR), Dutch Medical Registration (LMR), data quality

1 Introduction

Since 2011 Statistics Netherlands is responsible for the calculation of the Hospital Standardised Mortality Ratio (HSMR) for the Dutch hospitals. The HSMR is the ratio of the observed and expected mortality in a hospital and aims to make mortality better comparable between hospitals. In order to monitor the quality, the quality of some of the variables used in the calculation of the HSMR was investigated.

The HSMR is calculated using the Dutch Medical Registration (LMR) that contains information on hospital stays of patients. The LMR contains detailed information on approximately 88% of the inpatient hospital stays in Dutch hospitals. Included in this research are general and university hospitals and one specialty hospital that have registered inpatient data in the LMR. The exact method of calculating the HSMR is described elsewhere (Israëls et al., 2011, 2012). Here we will give only a global overview of the method. Using the main diagnosis of the hospital stay, the inpatient hospital stays are divided into main diagnosis groups. The fifty most important groups are selected accounting for approximately 80% of hospital mortality. For each of these fifty groups a logistic regression model is estimated that predicts the probability of mortality for each hospital stay. Summing these probabilities for all hospital stays in a hospital gives the expected mortality E_h of that hospital. The HSMR_h for hospital h is then given by $\text{HSMR}_h = 100 \cdot O_h/E_h$ where O_h is the observed mortality in that hospital for the fifty diagnosis groups. The goal of the HSMR is to enable fair comparison

of mortality between hospitals. The unstandardised gross mortality rate (mortality divided by the number of hospital stays) cannot be used, because the hospitals can differ significantly in their patient populations: some hospitals such as university hospitals can have more complex patients than other hospitals and can therefore also show relatively higher mortality. The HSMR tries to correct for these differences in patient population.

The model uses variables from the LMR. There are two main reasons for including variables in the model. One, there has to be a strong relation between the variable and mortality. Second, the hospital populations differ for these variables. Since the HSMR is often used as an indicator for quality of care, variables should not depend on the care given in the hospital. A variable such as type of treatment received, although having probably a strong relation with mortality, should therefore not be included in the model. The following variables are used in the model: age, gender, socio-economic status, severity of main diagnosis, urgency of admission (planned, not planned), comorbidities (secondary diagnoses), source of admission (home, nursing home, general hospital, or university or top-clinical hospital), year of discharge, month of admission (Israëls et al., 2011, 2012).

In a hypothetical experiment we could send the same patient to different hospitals. For the variables used in the standardisation we would like the hospitals to code the same values for this patient. However, for some of the variables used in the HSMR there are indications that this is not the case. This can have different reasons. First, the hospitals can have different coding practices: given the examinations and treatment etc. different values are coded in the registration. Second, there can be differences in the level of examination of patients. For example, a hospital that more thoroughly examines a patient might discover more comorbidities. For the calculation of the HSMR both reasons are unwanted. For the sake of brevity we take into account both types of differences when we speak of differences in coding practices or coding differences.

Assuming that the different coding practices do not depend on the quality of care, the different coding practices add noise to the results. Therefore, the observed spread in the HSMR is larger than would be expected if the hospitals would all have the same coding practice. One of the goals of present research is to estimate this additional spread. We will do this for the two variables that are suspected of having the largest differences in coding practices namely the comorbidities and urgency of the admission (in short urgency). These variables are also reported elsewhere as being possibly problematic, namely in Jarman (2008), Pieter et al. (2010) and van den Bosch et al. (2010). For administrative variables such as age and gender we do not expect large coding differences. Variables such as source of admission or the indicator for socio-economic status used in the model, have a relatively small effect on the HSMR, making possible coding differences less important.

The next section discusses more in detail the variables investigated. In section 3 we look at different aspects of these variables and try to estimate the increase in variance of the HSMR caused by the coding differences.

2 Methods

2.1 Decomposition of the HSMR

The observed HSMR of hospital i HSMR_i can be written as

$$\text{HSMR}_i = \text{HSMR}_i^0 + \text{HSMR}_i^e, \tag{1}$$

where HSMR_i^0 is the real HSMR and HSMR_i^e is the error introduced by coding differences. If we assume that the coding differences are independent of the true HSMR then the variance can be written as

$$\text{var}\left(\text{HSMR}_i\right) = \text{var}\left(\text{HSMR}_i^0\right) + \text{var}\left(\text{HSMR}_i^e\right). \tag{2}$$

One of the goals of the present research is to estimate the size of $\text{var}\left(\text{HSMR}_i^e\right)$ compared to $\text{var}\left(\text{HSMR}_i\right)$.

2.2 The variables investigated

In the following sections two variables are investigated, namely comorbidities and urgency. In this section a more precise description is given of both variables. All details concerning the variables and model can be found in the two methodological papers on the HSMR (Israëls et al., 2011, 2012).

The variable *comorbidities* is actually not one variable but consists of in total seventeen variables, each of which indicates whether or not a certain comorbidity from at least one of the seventeen comorbidity groups is present. Table 1 shows the comorbidity groups. These groups are the same groups as used in the Charlson Index (Charlson et al., 1987). All secondary diagnoses registered in the LMR and belonging to the seventeen comorbidity groups are used, but if a secondary diagnosis is identical to the main diagnosis, it is not considered a comorbidity. Complications are also excluded.

Two additional variables were derived for the analysis. First, for each hospital stay an indicator variable was derived that indicates whether or not one of the comorbidity groups is present for that patient. Second, a variable was derived containing the total number of comorbidity groups present for that patient. By averaging this last variable over all hospital stays in a hospital, it is also possible to derive the average number of comorbidities per hospital.

The other variable investigated is *urgency*, which indicates whether a hospital stay was planned (elective admissions) or not (urgent/emergency admissions).

3 Results

3.1 Registration differences between hospitals

Figure 1 shows the development of the average number of comorbidities per hospital stay and the average fraction of urgent hospital stays for each of the hospitals. One

TABLE 1 *The comorbidity groups used in the HSMR with corresponding ICD9-CM codes.*

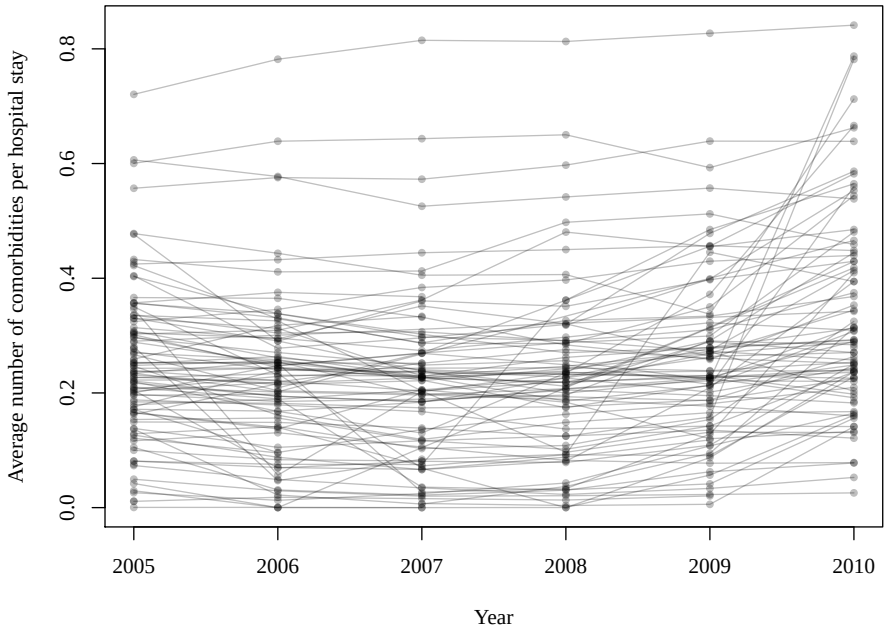
No.	Comorbidity groups	ICD9-CM codes
1	Acute myocardial infarction	410, 412
2	Congestive heart failure	428
3	Peripheral vascular disease	441, 4439, 7854, V434
4	Cerebral vascular accident	430–438
5	Dementia	290
6	Pulmonary disease	490, 491, 492, 493, 494, 495, 496, 500, 501, 502, 503, 504, 505
7	Connective tissue disorder	7100, 7101, 7104, 7140, 7141, 7142, 71481, 5171, 725
8	Peptic ulcer	531, 532, 533, 534
9	Liver disease	5712, 5714, 5715, 5716
10	Diabetes	2500, 2501, 2502, 2503, 2507
11	Diabetes complications	2504, 2505, 2506
12	Paraplegia	342, 3441
13	Renal disease	582, 5830, 5831, 5832, 5836, 5837, 5834, 585, 586, 588
14	Cancer	14, 15, 16, 18, 170, 171, 172, 174, 175, 176, 179, 190, 191, 192, 193, 194, 1950, 1951, 1952, 1953, 1954, 1955, 1958, 200, 201, 202, 203, 204, 205, 206, 207, 208
15	HIV	042, 043, 044
16	Metastatic cancer	196, 197, 198, 1990, 1991
17	Severe liver disease	5722, 5723, 5724, 5728

thing that can be noted is that the average number of comorbidities varies strongly between the hospitals. Some hospitals code hardly any comorbidities while others code on average more than 0.5 comorbidities per hospital stay. Although differences between the hospitals are likely, it is unlikely that these are this large. What also can be noted is that there is a group of hospitals that has started to code much more comorbidities in 2010. In some cases the number of comorbidities coded per hospital stay has doubled. It is very unlikely that the patient population in these hospitals will have changed that much in one year. Both the large spread between hospitals and the large spread in the year-to-year developments indicate that there are differences in the number of recorded comorbidities that are not caused by differences in patient population. For urgency there seem to be less differences between the hospitals. The bulk of the hospitals have between 50% and 70% urgent hospital stays. There are a few outliers that code hardly any urgent hospital stays in some or all years. However, these outliers are covered by the quality criteria of the HSMR: the HSMR is not calculated for hospitals coding urgency for less than 30% of the hospital stays (Israëls et al., 2011). The strong increase in coding in 2010 is not present for urgency. Therefore, except a few outliers (which in 2010 seems to be only one) there are no clear indications that there are coding differences between the hospitals for urgency.

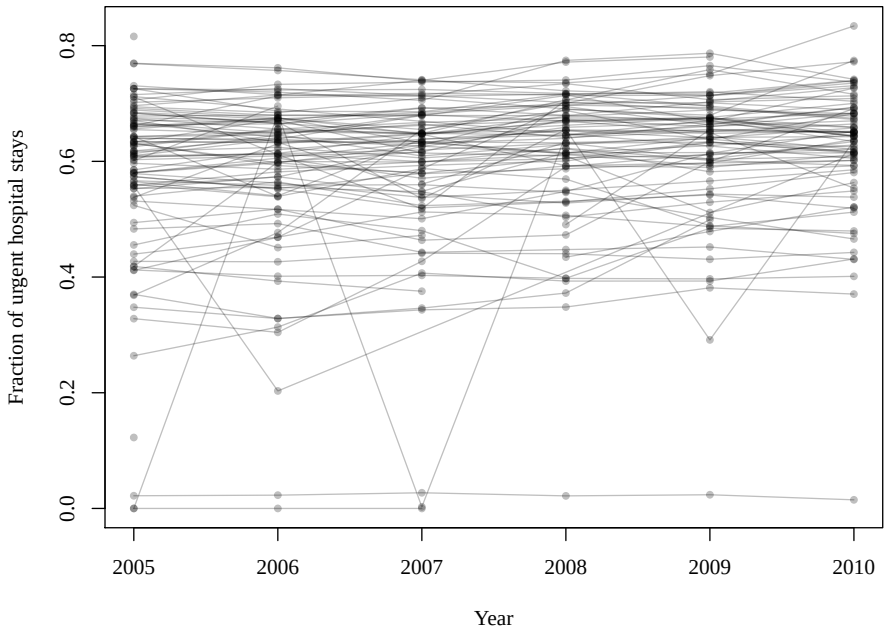
Coding more comorbidities or more urgency, given the type of patient, will lead to a decrease of the HSMR, as the larger number of comorbidities or urgencies will lead to a larger probability of death which in turn leads to a larger expected mortality. Therefore, if some hospitals code more comorbidities or more urgency for a given hospital stay, it is expected that these hospitals have a lower HSMR than those who code less. Figure 2 shows the HSMR for each of the hospitals plotted against the average number of comorbidities and against the fraction of urgent hospital stays. Hospitals coding more comorbidities have on average a lower HSMR than those coding less. If the one outlier in the graph is removed this effect is not seen for urgency.

Pieter et al. (2010) showed the same type of graphs for the HSMR of 2005. There it was concluded that there was a small effect of coding differences. As can be seen in figure 1 the coding differences in 2005 were less than in 2010. In principle it is possible that hospitals receiving more complex patients (having more comorbidities) provide relatively better care or vice versa. Therefore, the significant slope for the comorbidities does not necessarily mean that there are coding differences. However, it does indicate that coding differences can lead to significant effects on the HSMR. Coding ten percentage points more comorbidities for the same patients leads to a decrease of the HSMR of five points.

In figure 3 the change in the HSMR between 2009 and 2010 is plotted against the absolute change in the average number of comorbidities and the fraction of urgent hospital stays. As in the previous figure, the slope for urgency is not significant, while that for comorbidities is. As previously, an increase in the number of comorbidities leads to a decrease of the HSMR. However, in this case it seems unlikely that during one year hospitals start receiving both much more difficult patients and during the same period significantly improve their care. Therefore, the decrease seen in this figure is probably not realistic and is caused by changes in coding practice and not by changes in care or patient population.

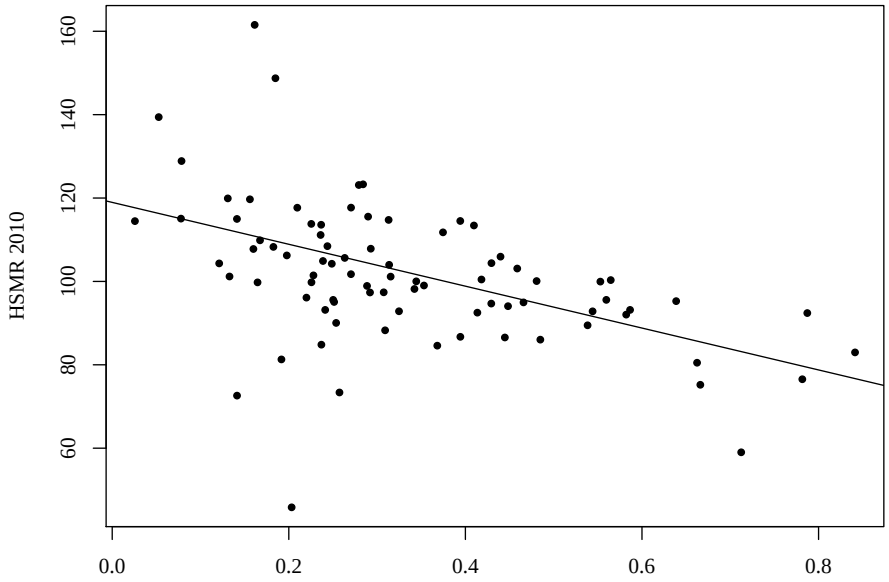


(A) *Average number of comorbidities*

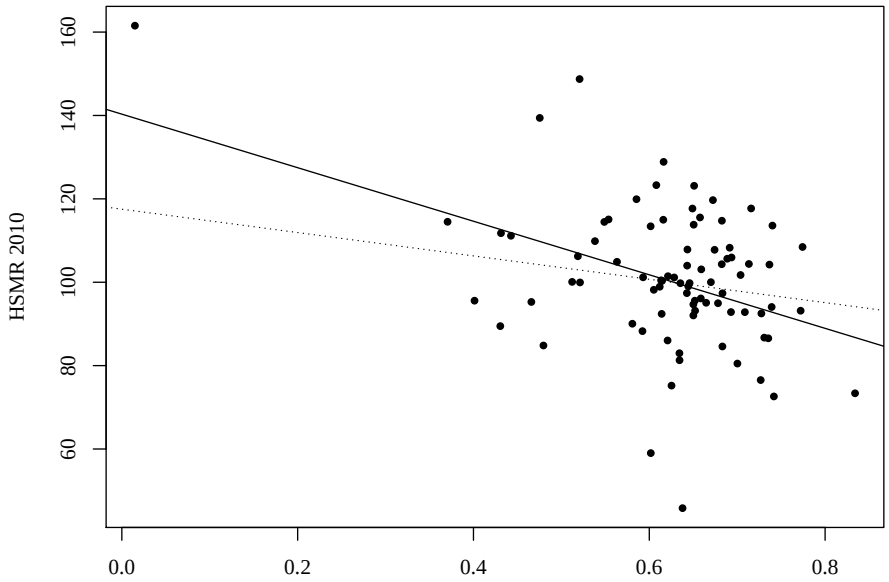


(B) *Fraction of urgent hospital stays*

FIGURE 1 *Development in the average number of comorbidities per hospital stay and the fraction of urgent hospital stays for each of the hospitals.*

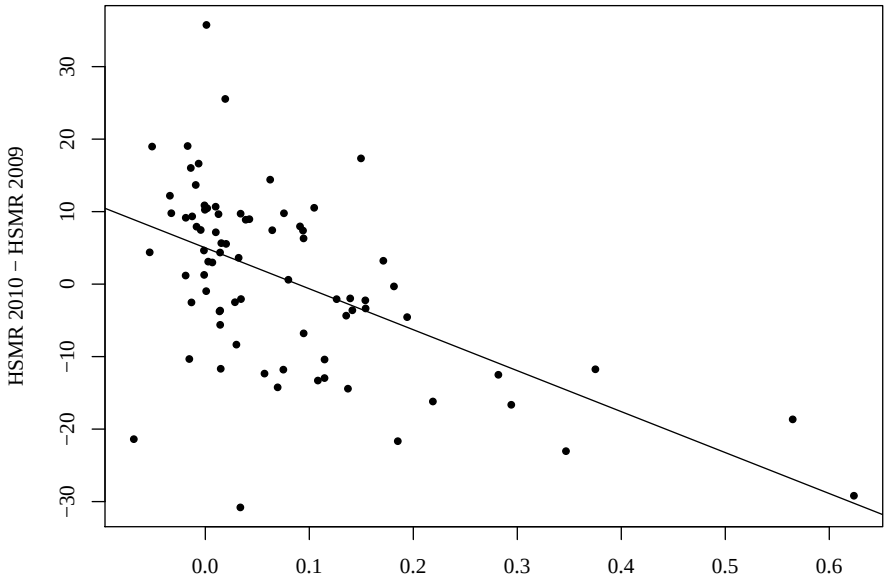


(A) *Average number of comorbidities*

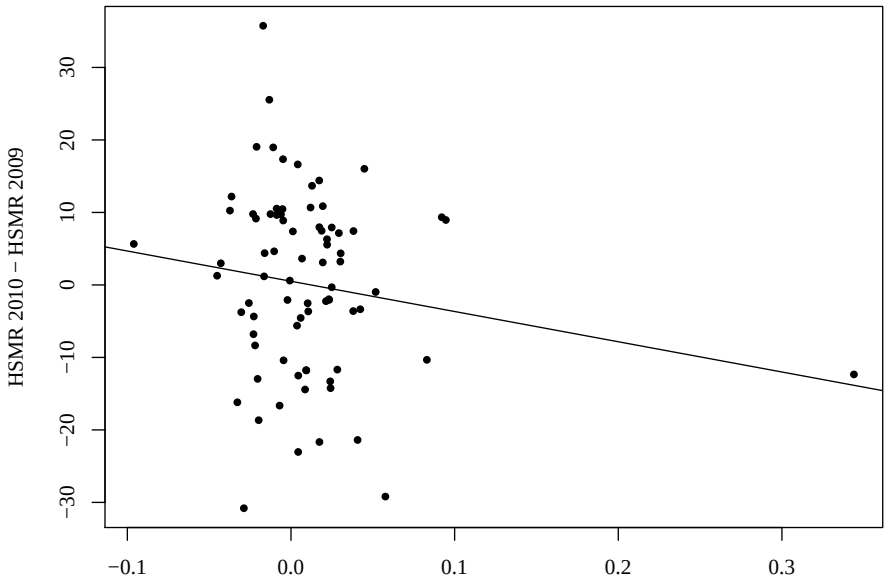


(B) *Fraction of urgent hospital stays*

FIGURE 2 *The HSMR of 2010 against the average number of comorbidities per hospital stay and the fraction of urgent hospital stays for each of the hospitals. The solid line is obtained using a weighted (using the number of hospital stays) linear regression and has a significant slope of -50 for the comorbidities and a significant slope of -64 for urgency. However, removing the one outlier in the latter graph results in the dotted line which has a non significant slope.*



(A) *Average number of comorbidities*



(B) *Fraction of urgent hospital stays*

FIGURE 3 *Change in the HSMR from 2009 to 2010 plotted against the absolute change in the average number of comorbidities per hospital stay and the fraction of urgent hospital stays. The solid line is obtained using a weighted (using the number of hospital stays) linear regression and has a significant slope of -56 for the comorbidities and a non-significant slope for urgency.*

From figure 3 it can be concluded that coding differences introduce a slope of minus five HSMR points per ten percentage points increase of comorbidity coding. This is approximately the same slope as found in figure 2. Previously it was concluded from figure 3 that this slope is probably not realistic and caused by coding differences for comorbidities. Therefore, the slope in figure 2 is probably also caused by coding differences and not by real differences between hospitals. By removing the slope from figure 2 a rough estimate can be obtained for the variance introduced by the coding differences for comorbidities. The variance of HSMR_i is 292; the variance of the residuals, which is an estimate of the variance of HSMR_i^0, is equal to 225. Therefore, using equation 2 the variance introduced by the coding differences for comorbidities can be estimated as 67, and the coding differences increase the variance by 30%.

As the slope in figure 3 for urgency was not significant, a similar analysis was not performed for urgency. For urgency it is concluded that there does not seem to be a significant increase in variance because of coding differences

Even if there are coding differences in, for example, the comorbidities this does not necessarily lead to a slope in the graph of the HSMR against the average number of comorbidities. Therefore, the estimated increase in variance of 30% for the comorbidities can be considered to be a lower bound for the coding differences.

3.2 Comparison between expected and observed number of comorbidities

In the previous sections it was shown that there are strong indications that for the comorbidities there are differences in coding practice between hospitals. For urgency, the differences seem to be much smaller. In this section it is attempted to estimate the size of the coding differences. We see differences between hospitals in the number of comorbidities or the number of urgent hospital stays registered. However, these differences can be caused by differences in patient population between hospitals and by coding variation. One possibility is to try to correct the comorbidities or urgency for differences in the patient population, e.g. to standardise. This can be done in the same way as mortality is standardised in the HSMR, by dividing the observed number of hospital stays with a given comorbidity or urgency by the expected number of hospital stays. When the correction for the differences in the patient population is good enough, there should only be statistical fluctuations left. Any variation larger than the expected statistical variation can be attributed to coding differences.

For each of the seventeen comorbidities and for urgency a model is estimated that predicts the probability of this comorbidity or urgency occurring for that hospital stay. This is also done for a derived variable that indicates the presence of any comorbidity group. For each of the fifty main diagnosis groups used in the calculation of the HSMR a separate model is estimated using logistic regression. Besides the covariates used in the HSMR (urgency, admission source, gender, severity of main diagnosis, age, month of admission, social economic status, see Israëls et al., 2011) the following covariates were added (also derived from the LMR):

re-admission	categorical variable with the following categories: 'no re-admission', 're-admission - planned', 're-admission - not planned'.
type of hospital	categorical variable with the following categories: 'general hospital', 'university hospital', 'general hospitals that also deliver topclinical care'.
specialty	categorical variable with 44 categories containing the medical specialty that determined the main diagnosis.
reason of admission	categorical variable with the following categories: 'observation', 'diagnostic examination', 'therapeutic treatment', 'guest stay of people accompanying the patient'.

Since it was not clear how for each hospital the categories 're-admission - planned' and 're-admission - not planned' of the variable 're-admission' are related to the categories 'planned' and 'not-planned' of the variable 'urgency', for the model for urgency the variable 're-admission' was recoded to a dichotomous variable with categories 'no re-admission' and 're-admission'.

Since above mentioned additional variables are either hospital dependent or care dependent these cannot be used for the HSMR. However, in this case it is attempted to predict the differences observed for comorbidities and urgency as good as possible using properties of both patient and hospital. Any remaining differences indicate coding differences. These extra variables were chosen because it was expected that they predict some of the variation in the target variables. It is for example not unreasonable to expect more comorbidities for patients that have been admitted previously, for patients admitted to university hospitals or for certain types of specialties. It is assumed that when all these variables are included in the model, the prediction of comorbidities and urgency is 'perfect', i.e. that the remaining hospital level variation can be attributed to coding differences. Of course, perfect prediction will in practice not be the case, there will always be predictors lacking, so the remaining variation can be considered as a maximum level for the coding differences.

The estimated models predict the target variables rather well. For the probability of occurrence of any comorbidity the C-statistic is 0.78[1]. For urgency it is 0.88. For the five most important comorbidities 'Metastatic cancer', 'Diabetes', 'Pulmonary disease', 'Acute myocordial infarction' and 'Cancer', the C-statistics are 0.94, 0.73, 0.80, 0.81 and 0.85 respectively. All of which indicate adequate to excellent discrimination.

As with the HSMR, the models can be used to construct a so called funnel plot. On the y-axis, the observed number of comorbidities or urgent hospital stays is divided by the expected number (one hundred corresponds to equal observed and expected). On the x-axis the expected number of hospital stays is shown. Figure 4 shows the funnel plot for

[1]The C-statistic is the area under the ROC (Receiver Operator Curve) and measures how well the model is able to discriminate between subject with the target variable and those without. It can be between 0.5 and 1; values above 0.7 indicate adequate discrimination; values above 0.8 excellent discrimination (Hosmer and Lemeshow, 2000).

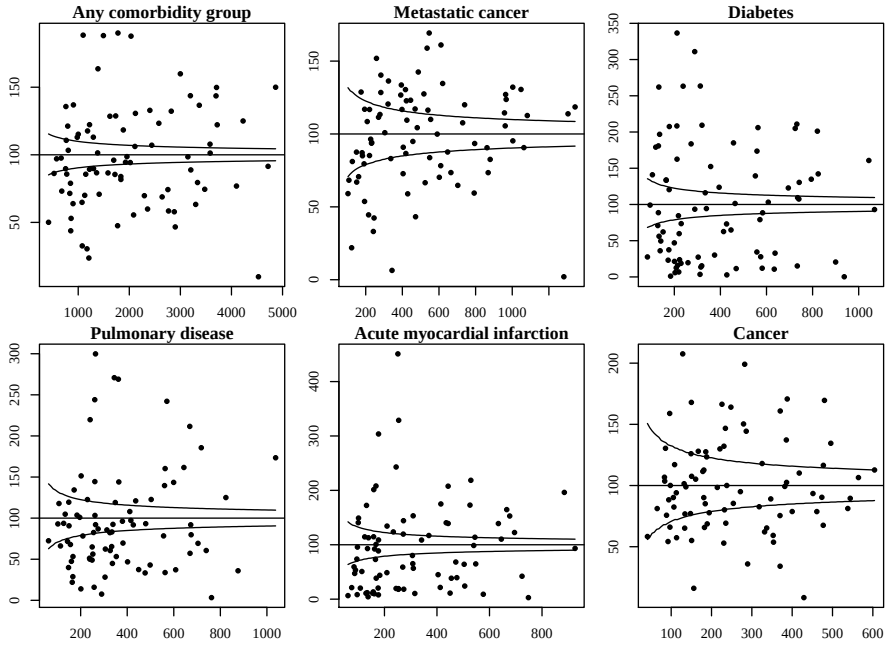


FIGURE 4 *Standardised number of hospital stays with a given comorbidity plotted against the expected number in the hospital. On the x-axis the expected number is given, on the y-axis the observed number divided by the expected number. The solid lines indicate the region which should contain 99.8% of the observations if the model is correct and there are no differences between the hospitals.*

having any comorbidity and the five most important comorbidities. The 99.8% control limits (solid line) indicate the region in which 99.8% of the observations are expected to fall. Values outside the 99.8% control limits are unlikely given the model. These control limits are calculated by assuming that the observed number follows a Poisson distribution with expectation value equal to the expected number of comorbidities or urgent hospital stays. Since the standard deviation of the Poisson distribution is equal to the square root of its expectation, the width of the control limits is approximately proportional to one over the square root of the expected number thereby creating the funnel-shaped control limits.

The figures clearly show that the variation in the number of comorbidities is much larger than would be expected based on the models. This can have two reasons. One, there is coding variation between the hospitals or, two, some explanatory variables are missing from the models. However, this then have to be variables for which significant variation exists between the hospitals that cannot be explained by any of the variables present in the model. However, the predictive power of the models on hospital stay level is good and we also included 'type of hospital' in the model which should explain possible differences between general, top-clinical and university hospitals. Therefore, it seems unlikely that the observed difference between the expected and observed number of comorbidities is caused by missing covariates alone.

Figure 5 shows the funnel plot for urgency. The spread is much less than for the comorbidities. However, there is still a substantial number of hospitals that is outside of the control limits. Therefore, although the coding differences seem to be less than for the comorbidities, there are indications for coding differences for urgency.

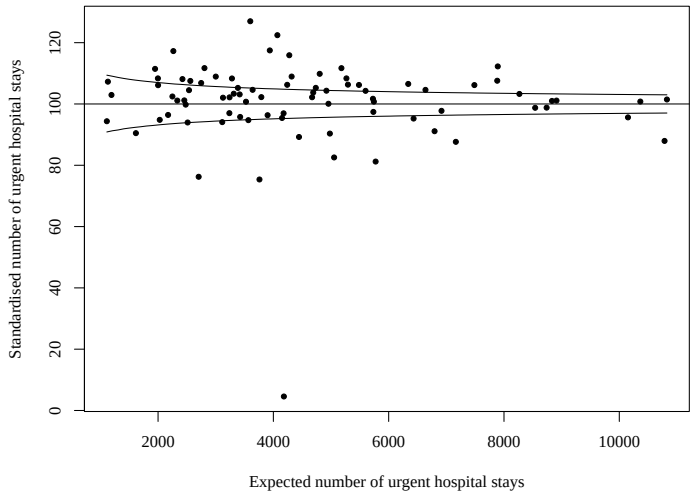


FIGURE 5 *Standardised number of urgent hospital stays plotted against the expected number in the hospital. The solid lines indicate the region which should contain 99.8% of the observations if the model is correct and there are no differences between the hospitals.*

3.3 Estimation of the effect of coding differences on the HSMR

The results presented in the previous sections indicate coding differences between hospitals for the comorbidities and to lesser extent for urgency. In section 3.1 an attempt was already made to estimate the effect of these coding differences on the HSMR. However, there only use was made of the assumption that the HSMR should not depend on the number of comorbidities or the number of urgent hospital stays coded by the hospital. Even if there are coding differences in, for example, urgency this does not necessarily lead to a slope in the graph of the HSMR against the fraction of urgent hospital stays per hospital. Therefore, the estimates from section 3.1, probably underestimate the true effect of the coding differences.

Using the models of the previous section, it is possible to estimate for each hospital stay the expected probability of occurrence for each of the comorbidities and urgency. These can be used in the estimation of the HSMR instead of the original values. So, for each hospital stay the probability of occurrence is calculated for urgency and each comorbidity. These probabilities are used in the model for the HSMR instead of the original dichotomous variables for urgency and the comorbidities. For comparison, the HSMR was also calculated without the use of the comorbidities and urgency in the model. Since the models of the previous section were only estimated using the 2010 data, the HSMR is also calculated using only the 2010 data and is, therefore, not exactly comparable to the HSMR presented in section 3.1.

Figure 6 shows funnel plots for the HSMR calculated using the three different methods: the original method of calculating the HSMR, the method using the predicted probabilities, and the method without the comorbidities or urgency. Switching from the observed to the predicted comorbidities decreases the variation in the HSMR and there are less hospitals outside of the control limits. This can also be seen in the standard deviation of the HSMR that decreases from 15.9 to 12.1. Without comorbidities in the model the standard deviation decreases to 13.2. Not shown, but what also appears

TABLE 2 *Number of hospitals inside and outside control limits in original model and in model with predicted comorbidities or in model with predicted urgency.*

(A) *Comorbidities, 99.8% control limit*

Original	Predicted comorbidities	
	Not sign.	Sign.
Not sign.	63	4
Sign.	10	5

(B) *Comorbidities, 95% control limit*

Original	Predicted comorbidities	
	Not sign.	Sign.
Not sign.	46	3
Sign.	19	14

(C) *Urgency, 99.8% control limit*

Original	Predicted urgency	
	Not sign.	Sign.
Not sign.	64	3
Sign.	2	13

(D) *Urgency, 95% control limit*

Original	Predicted urgency	
	Not sign.	Sign.
Not sign.	46	3
Sign.	7	26

is that the slope present in figure 2 disappears when using the predicted comorbidities or when the comorbidities are not included in the model.

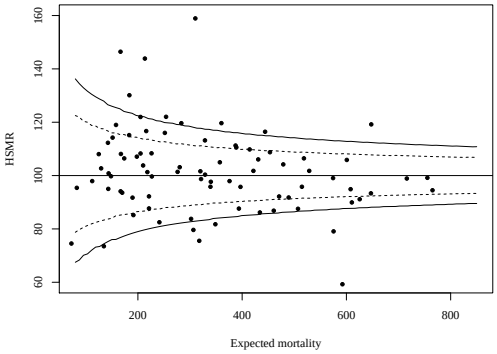
Using equation 2 and assuming that the HSMR estimated using the predicted comorbidities is a good estimate of HSMR_i^0, the variance introduced by coding differences (HSMR_i^e) is estimated as 104.3. Therefore, using these estimates the variance in the HSMR increases by 71% because of coding differences between the hospitals.

As expected, for urgency the differences are less. Using the predicted urgency the standard deviation decreases from 15.9 to 14.0. This gives an estimate for the variance introduced by the coding differences in urgency of 30%.

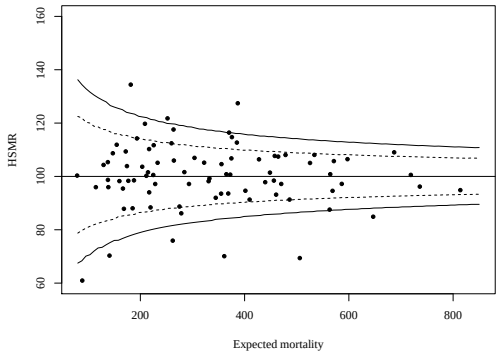
Besides looking at the spread in the HSMR, we can also look at the number of hospitals outside the control limits. Table 2 shows a cross-tabulation of the hospitals between being outside of the control limits (Sign.) in the original HSMR and in the HSMR calculated using the predicted values for the comorbidities or urgency. It shows, for example, that of the fifteen hospitals that were outside the 99.8% control limits in the original HSMR, ten are no longer outside when the predicted comorbidities are used. We can see that for the comorbidities the number of hospitals outside the control limits decreases strongly. For urgency the changes are again much smaller.

4 Conclusion

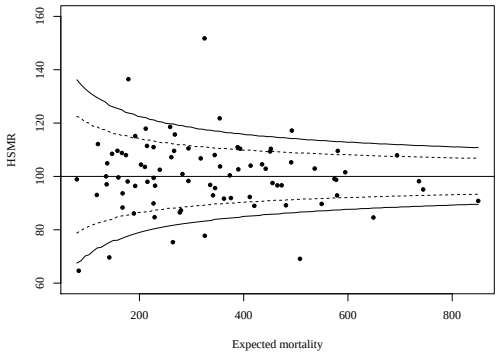
The differences between hospitals in the number of comorbidities coded are much larger than expected on the basis of the differences in patient populations alone. This is visible in the unrealistic development in the amount of comorbidities coded in 2010 compared to 2009 for a number of hospitals, and was confirmed in the results of section 3.2 where the observed number of hospital stays with a given comorbidity was compared to the expected number given the patient population. The spread observed there was much larger than could be explained by the model. Even if the model does not capture all hospital dependent patient variation, which is plausible, then it seems still unlikely that this explains all of the remaining hospital variation.



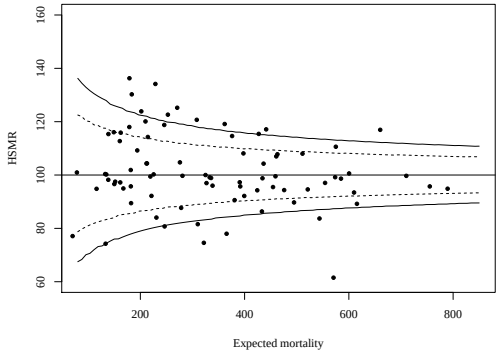
(A) *All covariates including observed comorbidities and urgency*



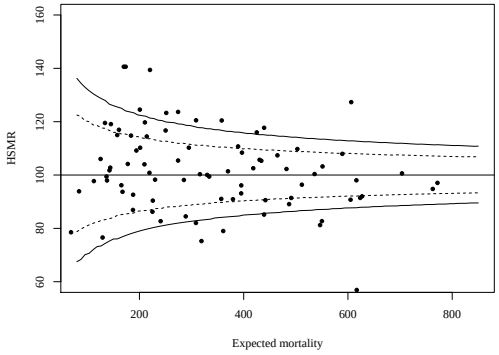
(B) *Predicted comorbidities*



(C) *No comorbidities*



(D) *Predicted urgency*



(E) *No urgency*

FIGURE 6 *Funnel plot of the HSMR 2010 for three different methods of taking comorbidities or urgency into account in the model. The model was estimated using only 2010 data.*

For the variable urgency the picture seems to be better. The variation between hospitals in the fraction of urgent hospital stays is much smaller. There are a few hospitals that code much urgent stays than the other hospitals. However, HSMR results of hospitals that code less than 30% not planned hospital stays are considered unreliable and are not distributed. Furthermore, these few hospitals will not strongly affect the results of the remaining hospitals. Although the variation for urgency is smaller than for the comorbidities, section 3.2 showed that the spread in urgency is still larger than can be explained by patient variation alone. So, even for urgency there seem to be some coding differences between hospitals.

Both in section 3.1 and 3.3 it was attempted to estimate the increase in the spread in the HSMR caused by coding variation. Of course, without additional data, some assumptions had to be made. In section 3.1 it was assumed that the HSMR should not decrease as the average number of comorbidities per hospital stay or the fraction of urgent hospital stays increases for a hospital. In section 3.3 it was assumed that the model used to predict for each hospital stay the probability that the patient has a given comorbidity or urgency, captures all between hospital patient variation. For both assumptions it is likely that they at least partially hold.

The assumption in section 3.1 that the HSMR should not depend on the number of comorbidities or urgent stays registered in the hospital seems not unrealistic, and since only this effect was removed from the HSMR, it is likely that the calculated effect of the coding differences is an underestimation. The value of 30% increase in the variance of the HSMR for the comorbidities found in section 3.1 can therefore be considered as a lower limit for the coding differences effect. The value of 70% found in section 3.3 for comorbidities can be considered as an upper limit, as it is likely that not all hospital dependent patient variation was included in the prediction models, which would imply that too much of the remaining variation is attributed to coding differences, leading to an overestimation. Thus the increase of variance caused by coding differences between hospitals is estimated to be between 30% and 70% for the comorbidities. Similarly, the estimate for urgency is between 0% and 30%, as no significant effect was measured in section 3.1 for urgency, and the extra variance calculated in section 3.3 was 30%.

The main reason for standardisation is that hospitals can differ significantly in their patient populations and that it is therefore not fair to compare hospitals based on their crude mortality rates. Therefore, mortality is standardised to remove differences between hospitals caused by differences in their patient populations. As can be seen from the results from section 3.3 adding comorbidities and urgency to the model increases the differences between the hospitals. Therefore, one could question whether or not these variables should be used to standardise mortality. On the other hand, both the comorbidities and urgency are very strong predictors for mortality (Israëls et al., 2011, 2012). They are also specifically variables expected to be discriminatory for identifying differences in patient population between hospitals. Hospitals specialising in care for more complex patients (such as university hospitals) will probably receive more patients that score on e.g. the comorbidities variables (and severity of the main diagnosis), while on other variables such as age and gender hospitals will be much more uniform. Therefore, by not including these variables in the model, one runs the risk of missing important differences in patient population. Furthermore, it is impossible to

check the assumptions made in this research for estimating the magnitude of the coding differences. Therefore, without further research we do not recommend removing the variables from the model for the HSMR. However, steps should be taken to ensure that the coding differences between hospitals decrease substantially. Dutch Hospital Data, the organisation responsible for the collection of the LMR, is working on clear coding standards for these variables and training of the coders. Both of which should help in obtaining less coding variation.

From the results in section 3.3, it seems that the number of hospitals outside the control limits is too large. This was also found for the Summary Hospital-level Mortality Indicator (SHMI) in England (Campbell et al., 2011). There it was decided to correct the control limits for unmodelled hospital heterogeneity using a multi-level model (National Health Service, 2012; Campbell et al., 2011). A similar approach could be taken for the HSMR. This has the advantage that the HSMRs remain comparable in time, while not overestimating the number of hospitals with an extraordinary high or low HSMR. However, since the number of hospitals in the LMR from which the additional spread has to be estimated is quite small (approx. 80), this has the disadvantage that the additional variation in the HSMR caused by unmodelled hospital heterogeneity is overestimated. This will lead to overestimation of the control limits, reducing the signal function of the HSMR. Further investigation should lead to a decision as to whether the control limits should be corrected for unmodelled hospital heterogeneity. The best way to solve the problem of coding differences, however, is to improve the registration in the hospitals.

References

Campbell, M. J., R. M. Jacques, J. Fotheringham, T. Pearson, R. Maheswaran, and J. Nicholl (2011). An evaluation of the summary hospital mortality index. Technical report, The University of Sheffield.

Charlson, M. E., P. Pompei, K. L. Ales, and C. R. MacKenzie (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseased 40*(5), 373–383.

Hosmer, D. W. and S. Lemeshow (2000). *Applied logistic regression* (2nd ed.). New York: John Wiley & Sons.

Israëls, A., J. van der Laan, J. van den Akker-Ploemacher, and A. de Bruin (2011). HSMR 2010: Methodological report. Technical report, Statistics Netherlands.

Israëls, A., J. van der Laan, J. van den Akker-Ploemacher, and A. de Bruin (2012). HSMR 2011: Methodological report. Technical report, Statistics Netherlands.

Jarman, B. (2008). In defence of the hospital standardized mortality ratio. *Healthcare Papers 8*(4), 37–41.

National Health Service (2012). Indicator specification: summary hospital-level mortality indicator. Technical report. Version 1.12, indicator reference no. 100699.

Pieter, D., R. B. Kool, and G. P. Westert (2010). Beperkte invloed gegevensregistratie op gestandaardiseerd ziekenhuissterftecijfer (hsmr). *Nederlands Tijdschrift voor de Geneeskunde 154*, A2186.

van den Bosch, W. F., J. Silberbusch, K. J. Roozendaal, and C. Wagner (2010). Variatie in codering patiëntgegevens beïnvloedt gestandardiseerd ziekenhuissterftecijfer (hsmr). *Nederlands Tijdschrift voor de Geneeskunde 154*, A1189.