Methodological Report

HSMR 2016: Methodological report

Jan van der Laan
Agnes de Bruin
Corine Penning
Janneke van den Akker

Index

1. Introduction

This report presents the methods Statistics Netherlands (CBS) has used to calculate the Hospital Standardised Mortality Ratios (HSMRs) for Dutch hospitals. HSMRs are ratios of observed and expected number of deaths and aim to present comparable hospital mortality figures. This chapter gives a general overview of the HSMR. Chapter 2 presents the changes introduced in the method of calculating the HSMR. The methodological aspects of the model used to calculate the HSMRs are described in chapter 3. The model outcomes are evaluated in chapter 4.

1.1 What is the (H)SMR?

Hospital mortality can be measured as the ratio of the number of hospital deaths to the number of hospital admissions (hospital stays) in the same period. This is generally referred to as the "gross mortality rate". Judging hospital performance on the basis of gross mortality rates is unfair, since one hospital may have had more life-threatening cases than another. For this purpose, it is more appropriate to adjust (i.e. standardise) mortality rates across hospitals as much as possible for differences in characteristics of the patients admitted to these hospitals ("case mix"). To this end, the *SMR* (Standardised Mortality Ratio) of a hospital *h* for diagnosis *d* is defined as

$$SMR_{dh} = 100 \times \frac{\text{Observed mortality}_{dh}}{\text{Expected mortality}_{dh}}.$$

The numerator is the *observed* number of deaths with main diagnosis *d* in hospital *h*. The denominator is the *expected* number of deaths for this type of admission under the assumption that individual mortality probabilities (per admission) do *not* depend on the hospital, i.e. are equal to mortality probabilities of identical cases in other hospitals. The denominator is therefore founded on a model based on data from all hospitals, in which the mortality of an admission is explained by characteristics of the patient, such as age, and characteristics of the admission, such as diagnosis and whether the admission is acute and unplanned versus planned. Characteristics of the hospital, such as the number of doctors per bed, are generally not incorporated in the model, since these can be related to the quality of care in the hospitals, which is the intended outcome of the indicator. The model thus produces an expected (estimated) mortality probability for each admission. Adding up these probabilities per hospital gives the total expected mortality over all admissions of that hospital. For each diagnosis *d*, the average SMR_d across the hospitals equals 100 when each hospital is weighted with its (relative) expected mortality.

The *HSMR* of hospital *h* is defined as

$$HSMR_h = 100 \times \frac{\text{Observed mortality in } h}{\text{Expected mortality in } h}.$$

in which both the numerator and denominator are sums across all admissions for all considered diagnoses. The HSMR thus also has a weighted average of 100. As HSMRs may also deviate from 100 only by chance, confidence intervals of the SMRs and HSMRs are calculated so that hospitals can see whether they have a (statistically) significantly high or low adjusted mortality rate compared with the average of 100.

1.2 Purpose of the HSMR

As in many other countries, there is much interest in measuring the quality of health care in the Netherlands. Hospitals can be assessed using various quality indicators, such as the number of medical staff per bed or the availability of certain facilities. However, these indicators do not measure the outcomes of medical performance. A good indicator for the performance of a hospital is the extent to which its patients recover, given the diagnoses and other important characteristics, such as age, sex and comorbidity, of the patients. Unfortunately, recovery is hard to measure and mostly takes place after patients have been discharged from the hospital. Although hospital mortality is a much more limited quality indicator, it can be measured accurately. That is why this indicator is now used in several countries, using the HSMR and SMRs as defined in section 1.1. If these instruments were totally valid, i.e. the calculations could adjust perfectly for everything that cannot be influenced by the hospital, a value above 100 would always point to inferior care quality, and the difference between numerator and denominator could be considered an estimate of "avoidable mortality". This would only be possible if the measurement was perfect and mortality by unforeseen complications, after adjustment for differences in case mix, was equally distributed across hospitals. However, it is impossible to construct a perfect instrument to measure the quality of health care. A significantly high (H)SMR will at most be an indication of possible shortcomings in hospital care. But the high value may also be caused by coding errors in the data or the lack of essential covariates in the model related to mortality. Still, a significantly high (H)SMR is often seen as a warning sign, a reason for further investigation into the causes.

1.3 History of the HSMR

In 1999 Jarman initiated the calculation of the (H)SMR for hospitals in England (Jarman et al., 1999). In the following years the model for estimating mortality probabilities was improved by incorporating additional covariates into the model. Analogous models were adopted by some other countries.

In 2005, Jarman started to calculate the (H)SMR for the Netherlands. Later on, these Dutch (H)SMRs were calculated by Kiwa Prismant, in collaboration with Jarman and his colleagues of Imperial College London, Dr Foster Intelligence in London and De Praktijk Index in the Netherlands. Their method is described in Jarman et al. (2010) and was slightly adapted by Prismant (Prismant, 2008) up to reporting year 2009. In 2010 Dutch Hospital Data (DHD, Utrecht), the holder of the national hospital discharge data, asked CBS to calculate the (H)SMRs for the period 2008-2010 and for subsequent years. CBS is an independent public body and familiar with the input data for the HSMR, i.e. the hospital discharge register (LMR: Landelijke Medische Registratie, and its successor LBZ: Landelijke Basisregistratie Ziekenhuiszorg), as it uses this data source for a number of health statistics (see www.statline.cbs.nl).

The starting point for CBS was the HSMR methods previously used by Prismant. As a result of progressive insight CBS has introduced changes in the model for the HSMR 2008-2010 and later years, which are described in the yearly methodological reports (CBS, 2011, 2012, 2013, etc.).

1.4 Confidentiality

Under the Statistics Netherlands Act, CBS is required to keep all data about individuals, households, companies or institutions confidential. Therefore it normally does not deliver recognisable data from institutions to third parties, unless the institutions concerned have stated that they do not have any objections to this. For this reason, CBS needs written permission from all hospitals to deliver their hospital specific (H)SMR figures to DHD. CBS only

supplies DHD with (H)SMR outcomes of hospitals that have granted authorisation to do so. In turn DHD sends each hospital its individual outcome report. Publication of (H)SMR data, which has become mandatory in the Netherlands since 2014 by a regulation of the Dutch Healthcare Authority (NZa), is the responsibility of the hospitals themselves. CBS does not publish data on identifiable hospitals.

1.5 CBS output

CBS estimated the models for expected mortality per diagnosis for the most recent three year period. It calculated the HSMRs and SMRs for all hospitals that (1) had authorised CBS, (2) had registered all or a sufficient part of its admissions in the LBZ in the relevant period, and (3) were not excluded on the grounds of criteria for quality and comparability, which means that the hospital's LBZ data were not too deviant in some respects (see section 3.5).

CBS produces the following output:

1. A hospital-specific report for each hospital, sent via DHD, containing the HSMR and the diagnosis-specific SMR figures for the three year period and the individual years. SMRs are also presented for different patient groups (by age, sex and urgency of admission) and diagnosis clusters. Hospitals can see how they compare with the national average, overall, and per diagnosis and patient group. CBS only made reports for hospitals not excluded under the exclusion criteria and that signed the authorisation request.
2. Each hospital not excluded on the grounds of the exclusion criteria and that signed the authorisation request is provided with a dataset with the mortality probabilities for all its admissions. Besides the probability, each record contains the observed mortality (0 or 1) and the scores on the covariates of the HSMR model. The hospital can use these data for internal investigation.
3. A report on the methods used for calculating the HSMR including the model results and parameters (this document; see www.cbs.nl).

1.6 Limitations of the HSMR

In section 1.2 we argued that the HSMR is not the only indicator to measure hospital care quality. Furthermore, the quality and limitations of the HSMR (and the SMR) instrument are under debate. After all it is based on a statistical model (i.e. the denominator), and a model is always a simplification of reality.

Since the very first publication of the HSMR in England, there has been an on-going debate about the quality of the HSMR as an instrument. Supporters and opponents agree that the HSMR is not a unique, ideal measure, but at most a possible indicator for the quality of health care, alongside other possible indicators. But even if HSMR were to be used for a more limited purpose, i.e. standardising hospital mortality rates for unwanted side-effects, the interpretation of HSMRs would present various problems, some of which are described briefly below. See also Van Gestel et al. (2012) for an overview.

- Section 3.4 contains the list of covariates included in the regression model. Hospitals do not always code these variables in the same way. Variables such as Age and Sex do not give any problems, but how aspects like acute admissions, main diagnosis and comorbidity are coded may depend on individual physicians and coders. Lilford and Pronovost (2010) argue that if the quality of the source data is insufficient, the regression model should not adjust for such erroneously coded covariates. Our own investigation (Van der Laan, 2013) shows that comorbidities in particular present a

problem in the Netherlands, as there is not much uniformity in coding this covariate (see also section 4.3). Van den Bosch et al. (2010) refer extensively to the influence of coding errors. Nationwide, the registration of comorbidities in Dutch hospitals has increased strongly up to 2014. From 2015 onwards this has stabilised, although there are still hospitals showing large shifts in the registration of comorbidities. Exclusion criteria for outliers may solve this problem partly but not completely.
Another problem is that some hospitals do not (completely) register whether a comorbidity was a complication or not. As complications are excluded from the HSMR comorbidity covariates, differences in registration can have an undesired effect on the HSMR outcomes. To stimulate correct coding of complications an indicator has been added to the HSMR 2016 reports sent to the hospitals showing the percentage of registered complications of the hospital, and the overall average.

- Some hospitals may have on average more seriously ill patients than others, even if they have the same set of scores on the covariates. University hospitals may, for example, have more serious cases than other hospitals. It is questionable whether the model adjusts satisfactorily for this phenomenon. Some essential covariates related to mortality are then missing. This may be caused by some of the desired covariates not being measured in the LBZ. Some factors will actually even be hard to measure in this type of routinely collected datasets of all hospital discharges.

- The same problem occurs when certain high risk surgical procedures are only performed in certain hospitals. For instance, open heart surgery only occurs in authorised cardiac centres, and these hospitals may have higher SMRs for heart disease because of the more dangerous interventions. This could be solved by including a covariate in the model that indicates whether such a procedure was performed. This has the disadvantage that a method of treatment is used as a covariate, while ideally it should not be part of the model as it is a component of hospital care.

- Hospital admission and discharge policies may differ. For instance, one hospital may admit the same patient more frequently but for shorter stays than another. Or it may discharge a patient earlier than another because there are adequate external terminal care facilities in the neighbourhood. Also, patients may be referred from one hospital to another for further treatment. Obviously, all these situations influence the outcome of the HSMR, as they influence the observed mortality numbers, but these differences in HSMR cannot be translated in terms of quality of care.

- Hospitals can compare their HSMR and SMRs with the national average of 100. The comparison between (H)SMRs of two or more hospitals with each other is more complicated. There is no complete adjustment for differences in case mix between pairs of hospitals. Theoretically, it is even possible that hospital A has higher SMRs than hospital B for all diagnosis groups, but a lower HSMR. Although this is rather theoretical, bilateral comparison of HSMRs should be undertaken with caution (Heijink et al., 2008).

Some issues in the incomplete correction for differences in the case mix between hospitals may be partly addressed by peer group comparison of (H)SMRs. The calculation of H(SMR)s is still based on the model for all hospitals (without correcting for the type of hospital), but a

specialised hospital can then also compare its results with the average for similar hospitals (peer group). For instance, the average HSMR of university hospitals is >100 in the Netherlands, due to insufficient case mix correction, but this still allows these hospitals (and for specific diagnoses also other specialised hospitals) to compare their results with the peer group average.

An indicator including early post-discharge mortality alongside in-hospital mortality could be introduced to tackle the problem of differences in discharge policies (e.g. the availability of terminal care outside hospital), and to some extent referrals between hospitals. Ploemacher *et al*. (2013) saw a decrease in standardised in-hospital mortality in the Netherlands in 2005-2010, which may have been caused by an overall improvement in care quality, but may also be partly explained by substitution of in-hospital mortality by outside-hospital mortality, possibly caused by changes in hospital admission and discharge policies. In cooperation with CBS, Pouw *et al*. (2013) did a retrospective analysis on Dutch hospital data linked to mortality data, and concluded that including early post-discharge mortality is advisable to diminish the effect of discharge bias on the HSMR. In the UK, the SHMI (Summary Hospital-level Mortality Indicator) has been adopted, which includes mortality up to 30 days after discharge (Campbell *et al*., 2011). In 2014, CBS studied the optimal time frame and definition of an indicator including early post-discharge mortality (Van der Laan *et al*., 2015). A fixed period of 45 days after admission in which all mortality is included in the mortality indicator would make the indicator less dependent on hospital discharge policies. A recent French study also recommends fixed post-admission periods of more than 30 days (Lamarche-Vadel et al., 2015).

Although including post-discharge mortality in the indicator would reduce the effect of differences in hospital discharge policies, it would not reduce the effect of differences in admission policies for terminally ill patients. Some hospitals may admit more patients specifically (and sometimes only) for palliative care than other hospitals. As such patients are admitted to die in hospital, not to receive curative care, these admissions may distort HSMR outcomes. Palliative care can be measured in ICD10 (code Z51.5), but this variable should be used with caution, as differences between hospitals in coding practices have been shown in UK and Canada, and adjusting for palliative care may increase the risk of gaming (NHS, 2013; Chong et al., 2012; Bottle et al., 2011). Because of this, and because the LBZ registration does not allow for distinguishing between admissions of terminally ill patients for palliative care only and admissions for curative treatment ending in palliative care, palliative care admissions have not yet been excluded from the calculation of the HSMR in the Netherlands. However, the HSMR reports sent to the hospitals include information on the percentage of the hospital's admissions and deaths related to palliative care as registered in the LBZ compared to the overall average. This may indicate to some extent whether or not palliative care could have biased a hospital's HSMR. However, since the Netherlands also shows a large variation between hospitals in the coding of palliative care, this information should be used with caution.

2. Method changes

This chapter summarizes the changes in the HSMR method (HSMR 2016) compared to the method used last year (HSMR 2015). For previous changes see the respective methodological reports (CBS, 2011, 2012, 2013, 2014, 2015, 2016).

Overall, the method has remained the same. Only the following minor changes have been implemented:

- Recalculation of the covariate *severity of main diagnosis*:
  In 2012 and 2013 hospitals transitioned from using ICD9 to code the diagnoses of admissions to using ICD10. The HSMR covariate *severity of main diagnosis* is determined using historical mortality, based on six years of LBZ data not overlapping with the data used for the HSMR calculation. Previously these historical data consisted of diagnoses coded in ICD9 only, so the severity was determined for ICD9 diagnoses. ICD10 coded main diagnoses were converted to ICD9 to derive the severity class. However, for the HSMR 2014-2016 the historical data used to determine the severity classes (2008-2013) is partially coded in ICD10, and the admissions for which the severity has to be determined are all coded in ICD10. Therefore, it was decided that the severity will be determined for ICD10 diagnoses instead of for ICD9 diagnoses. The historical dataset used now consisted of approximately 1.5 years of data in ICD10, and approximately 4.5 years of data in ICD9. A method was developed to calculate the severity for ICD10 main diagnoses, based on a mixed ICD9/ICD10 dataset, ensuring a gradual shift over time from severities based on ICD9 data to severities based solely on ICD10 data. The method and an investigation of the effects of this change are presented in chapter 5.

- For 2016 an updated version of the classification of socio-economic status (SES) per postal code was used (see section 3.4).

- In the LBZ data of 2016 the *urgency of the admission* variable was set to 'elective' for all liveborn in hospital with ICD10 main diagnosis code Z38. This was done by the holder of the LBZ to improve the data consistency. As these admissions concern normal deliveries that were planned to take place in hospital, they should not be coded as 'acute' according to the LBZ coding rules.

- The routine used to estimate the logistic regression models in previous years (the lrm routine from the rms R-package; CBS, 2016) had in the present models problems with converging for some of the diagnosis groups. Therefore, it was decided to switch to the glm routine in R. Except for the few groups where the other routine did not converge, this did not affect the outcome (SMRs and coefficients). It appeared that the previous routine underestimated the C-statistic in some of the diagnosis groups (see section 4.4 for more information).

3. (H)SMR model

Expected hospital mortality - i.e. the denominator of the SMR - has to be determined for each diagnosis group. To this end we use logistic regression models, with mortality as the target (dependent) variable and various variables available in the LBZ as covariates. The regression models for the (H)SMR of a three year period (year t-2 up to year t), and the (H)SMRs of the individual years t-2, t-1 and t, use LBZ data of four years: year t-3 up to year t. The addition an additional year increases the stability and accuracy of the estimates, while keeping the model up to date.

3.1 Target population and dataset

3.1.1 Hospitals

"Hospital" is the primary observation unit. Hospitals report admission data (hospital stay data) in the LBZ. However, not all hospitals participate in the LB. In principle, the HSMR model includes all short-stay hospitals with inpatient admissions participating in the LBZ in the relevant years. The target population of hospitals that qualify for entry in the HSMR-model thus includes all general hospitals, all university hospitals, and short-stay specialised hospitals with inpatient admissions. For a number of partially non-responding hospitals only the fully registered months were included in the model, as in the other months there were indications that fatal cases were registered completely and non-fatal cases partially. The partially registered months of these hospitals were removed from the model as these would otherwise unjustly influence the estimates.

All the above-mentioned hospitals were included in the model, but (H)SMRs were only calculated for hospitals that met the criteria for LBZ participation, data quality and case mix (see section 3.5).

3.1.2 Admissions

We consider both the population of hospitals and the population of admissions. Our target population of admissions consists of "all hospital stays (inpatient admissions) of Dutch residents in Dutch short-stay hospitals in a certain period". The date of discharge, and not the day of admission, determines the year a record is assigned to. So the population of hospital stays of year *t* comprises all inpatient admissions that ended in year *t*. For the sake of convenience, mostly we call these hospital stays "admissions", thus meaning the hospital stay instead of only its beginning. Day admissions are excluded as these are in principle non-life-threatening cases with hardly any mortality. However, from 2015 onwards the new case type 'prolonged observations, unplanned, without overnight stay' is included in the HSMR. This case-type was introduced by the Dutch Healthcare Authority, and it replaces the majority of one-day inpatient admissions that were formerly registered. This case type involves more mortality than day cases, and it is therefore relevant to include this in the HSMR.

Lastly, admissions of foreigners are excluded from the HSMR model, partly in the context of possible future modifications of the model, when other data can be linked to admissions of Dutch residents. The number of admissions of foreigners is relatively small.

3.2 Target variable (dependent variable)

The target variable for the regression analysis is the "in-hospital mortality". As this variable is binary, logistic regressions were performed.

3.3 Stratification

Instead of performing one logistic regression for all admissions, we performed a separate logistic regression for each of the diagnosis groups *d*. These sub-populations of admissions are more homogeneous than the entire population. Hence, this stratification may improve the precision of the estimated mortality probabilities. As a result of the stratification, covariates are allowed to have different regression coefficients across diagnosis groups.
The diagnosis groups are clusters of ICD codes registered in the LBZ. Here the main diagnosis of the admission is used, i.e. the main reason for the hospital stay, which is determined at discharge. The basis for the clustering is the CCS *(Clinical Classifications Software*[1]), which clusters ICD diagnoses into 259 clinically meaningful categories. For the HSMR, we further clustered these into 157 diagnosis groups, which are partly the same clusters used for the SHMI (Summary Hospital-level Mortality Indicator) in the UK (HSCIC, 2016).

Therefore, the model includes 157 separate logistic regressions, one for each diagnosis group *d* selected.

In the file 'Classification of variables', published together with this report, for each of the 157 diagnosis groups the corresponding CCS group(s) are given, as well as the ICD10 codes of each CCS group.

Apart from the SMRs for each of the 157 diagnosis groups, hospitals also receive SMRs for 17 aggregates of diagnosis groups. This makes it possible to evaluate the SMR outcomes at both the detailed and the aggregated diagnosis level. The 17 main clusters are also given in the 'Classification of variables' file. These were derived from the main clusters in the CCS classification of HCUP (see Appendix C1 in http://www.hcup-us.ahrq.gov/toolssoftware/ccs/CCSUsersGuide.pdf), with the following adaptations:

- HCUP main clusters 17 ("Symptoms; signs; and ill-defined conditions and factors influencing health status") and 18 ("Residual codes; unclassified") were merged into one cluster.
- CCS group 54 ("Gout and other crystal arthropathies") is classified in main cluster "Diseases of the musculoskeletal system and connective tissue", and CCS group 57 ("Immunity disorders") is classified in main cluster "Diseases of the blood and blood-forming organs", whereas in the HCUP classification these groups fall in main cluster "Endocrine, nutritional and metabolic diseases, and immunity disorders".
- CCS group 113 ("Late effects of cerebrovascular disease") is classified in main cluster "Diseases of the nervous system and sense organs", whereas in the HCUP classification this group falls in main cluster "Diseases of the circulatory system".
- CCS group 218 ("Liveborn") is classified in main cluster "Complications of pregnancy, childbirth, and the puerperium; liveborn", whereas in the HCUP classification this group falls in main cluster "Certain conditions originating in the perinatal period".

[1] *See http://www.hcup-us.ahrq.gov/toolssoftware/icd_10/ccs_icd_10.jsp*

These adaptations are in accordance with the diagnosis groups used for the SHMI (Summary Hospital-level Mortality Indicator) in the United Kingdom (HSCIC, 2016).

Although the names of the main clusters have much similarity with the names of the chapters of the ICD10, there is no one-to-one relation between the two. Although most ICD10 codes of a CCS group do fall within one ICD10 chapter, there often are also codes that fall in other chapters. Especially codes from the R chapter of ICD10 are scattered over several HCUP main clusters.

3.4 Covariates (explanatory variables or predictors of in-hospital mortality)

By including covariates of patient and admission characteristics in the model, the in-hospital mortality is adjusted for these characteristics. As a result, the (H)SMRs are adjusted for these covariates as well. Thus, variables (available in the LBZ) associated with patient in-hospital mortality are chosen as covariates. The more the covariates discriminate between hospitals, the larger the effect on the (H)SMR.

The LBZ variables that are included in the model as covariates are age, sex, socio-economic status, severity of main diagnoses, urgency of admission, Charlson comorbidities, source of admission, year of discharge and month of admission. These variables are described below. For the variables socio-economic status, severity of main diagnosis and source of admission the detailed classifications are presented in the file 'Classification of variables', published together with this report.

For the regressions, all categorical covariates are transformed into dummy variables (indicator variables), having scores 0 and 1. A patient scores 1 on a dummy variable if he/she belongs to the corresponding category, and 0 otherwise. As the dummy variables for a covariate are linearly dependent, one dummy variable is left out for each categorical covariate. The corresponding category is the so-called *reference category*. We took the *first* category of each covariate as the reference category.

The general procedure for collapsing categories is described in section 3.6.2. Special (deviant) cases of collapsing are mentioned below.

Age at admission (in years): *0, 1-4, 5-9, 10-14, …, 90-94, 95+*.

Sex of the patient: *male, female.*
If Sex is unknown, "female" was imputed. This is a rare occurrence.

SES (socio-economic status) of the postal area of patient's address: *lowest, below average, average, above average, highest, unknown.*

The SES variable was added to the LBZ dataset on the basis of the postal code of the patient's residence. SES was derived from the Netherlands Institute for Social Research (SCP)[2], which had collected SES data and performed principal component analyses on variables concerning

[2] http://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores

Income, Employment and Education level. Each four-letter postal area was thus assigned a component score. Population-weighted quintiles were calculated from these scores, resulting in the six SES categories mentioned above. Patients for whom the postal area does not exist in the dataset of the SCP (category "unknown"), were added to the category "average" if collapsing was necessary. For 2013, admissions followed the SES classification of 2010, for 2014-2015 admissions followed the SES classification of 2014, and for 2016 the SES classification of 2016 was used.

Severity of main diagnosis groups: *[0-0.01), [0.01-0.02), [0.02-0.05), [0.05-0.1), [0.1-0.2), [0.2-0.3), [0.3-0.4), [0.4-1], Other*.

This is a categorisation of main diagnoses into mortality rates. Each ICD10 main diagnosis code is classified in one of these groups, as explained below.

A separate model was estimated for each of the 157 diagnosis groups. Most groups have many sub-diagnoses (individual ICD codes), which may differ in seriousness (mortality risk). To classify the severity of the sub-diagnosis, we used the method suggested by Van den Bosch et al. (2011), who suggested categorising the ICD codes into mortality rate categories. To this end, we computed inpatient mortality rates for all ICD sub-diagnoses for six historical LMR years, and chose the following boundaries for the mortality rate intervals: 0, .01, .02, .05, .1, .2, .3, .4 and 1. ('0' means 0 percent mortality; '1' means 100 percent mortality). These boundaries are used for all 157 diagnosis groups. The higher severity categories only occur for a few diagnosis groups.

Six historical LMR/LBZ years are used to determine the classification, not overlapping with the years the HSMR is calculated for as otherwise both are using the same mortality data. The period of the historical dataset shifts every year for each new HSMR calculation, to keep it up to date.

Up to the HSMR 2013-2015, the historical LMR dataset consisted of diagnoses coded in ICD9-CM only, and the severities were also determined for ICD9-CM codes. Main diagnoses registered in ICD10 were converted to ICD9-CM to determine the severity covariate. As In 2012-2013 hospitals transitioned from using ICD9-CM to code the diagnoses of admissions to using ICD10, the diagnoses used for the HSMR 2014-2016 calculation are all coded in ICD10, and the historical dataset used to determine the severities also partly consists of ICD10 coded diagnoses. Therefore, for the HSMR 2014-2016 and later HSMR models, the severities are determined for ICD10 diagnoses. For the HSMR 2014-2016 the severity classification was based on the LMR/LBZ of 2008-2013, which consists of a mix of ICD10 and ICD9-CM data. A method was developed to calculate the severity for ICD10 main diagnoses with such historical datasets, ensuring a gradual shift over time from severities based on ICD9 data to severities based solely on ICD10 data. The method and an investigation of the effects of this change are presented in chapter 5.

For the severity classification the Dutch ICD10-ICD9-CM conversion table was used (table 'ICD10 – CvZ80', see http://www.rivm.nl/who-fic/ICD.htm.). As this table had not been updated for recent years, new ICD codes added to the ICD10 in recent years did not have a converted ICD9-CM code. Therefore, in consultation with DHD we added conversions for these codes to the table, to make it complete.

When a ICD10 code and its ICD9-CM equivalent did not occur in the historical dataset, a severity " other" was assigned in the calculation of the (H)SMR. ICD codes that are used by less than four hospitals and/or have less than 20 admissions also receive a severity of "other". The

category ”other” contains diagnoses for which it is not possible to accurately determine the severity. If this category “other” needs to be collapsed however, it does not have a natural nearby category. We decided to collapse “other” with the category with the highest frequency (i.e. the mode), if necessary. In the file with regression coefficients (see section 4.5) this will result in a coefficient for “other” equal to that of the category with which “other” is collapsed. The only exceptions are when Comorbidity_17 (Severe liver disease) is collapsed with Comorbidity 9 (Liver disease), and when Comorbidity_11 (Diabetes complications) is collapsed . with Comorbidity 10 (Diabetes). In these cases the regression coefficient of Comorbdidty 17/11 is set to zero in the coeffients file, and the coefficient of the less severe analogue (Comorbidity 9/11) should be used for Comorbidity 17/11.

The individual ICD10 codes with the corresponding severity category are available in the separate file ‘Classification of variables’, published together with this report.

Urgency of the admission: elective, acute.
The definition of an acute admission is: an admission that cannot be postponed as immediate treatment or aid within 24 hours is necessary. Within 24 hours means 24 hours from the moment the specialist decides an acute admission is necessary.

3.4.1 Comorbidity groups of Charlson index and the corresponding ICD10 codes

No.	Comorbidity groups	ICD10 codes
1	Acute myocardial infarction	I21, I22, I252
2	Congestive heart failure	I50, I110, I130, I132, I255, I42, I43, P290
3	Peripheral vascular disease	I70, I71, I731, I738, I739, I771, I790, I792, K551, K558, K559, Z958, Z959, R02
4	Cerebrovascular disease	G450-G452, G454, G458, G459, G46, I60-I69
5	Dementia	F00-F03, F051, G30, G311
6	Pulmonary disease	J40-J47, J60-J67
7	Connective tissue disorder	M05, M060, M063, M069, M32, M332, M34, M353
8	Peptic ulcer	K25-K28
9	Liver disease	B18, K700-K703, K709, K713-K715, K717, K73, K74, K760, K762-K764, K768, K769, Z944
10	Diabetes	E109, E119, E129, E139, E149
11	Diabetes complications	E100-E108, E110-E118, E120-E128, E130-E138, E140-E148
12	Hemiplegia or paraplegia	G041, G114, G801, G802, G81, G82, G830-G834, G838, G839
13	Renal disease	I120, I131, N01, N03, N052-N057, N18, N19, N25, Z490-Z492, Z940, Z992
14	Cancer	C00-C26, C30-C34, C37-C41, C43, C45-C58, C60-C76, C81-C85, C88, C90-C97
15	HIV	B20-B24
16	Metastatic cancer	C77-C80
17	Severe liver disease	I850, I859, I864, I982, K704, K711, K721, K729, K765, K766, K767

Comorbidity_1 – Comorbidity_17. All these 17 covariates are dummy variables, having categories: *0* (*no*) and *1* (*yes*).

The 17 comorbidity groups are listed in table 3.4.1, with their corresponding ICD10 codes. These are the same comorbidity groups as in the Charlson index. However, separate dummy variable are used for each of the 17 comorbidity groups.
For the few hospitals that still registered in ICD9-CM in 2013 the diagnoses are converted to ICD10 and then classified according to the ICD10 definitions of the Charlson comorbidities.

All secondary diagnoses registered in the LBZ and belonging to the 17 comorbidity groups are used, but if a secondary diagnosis is identical to the main diagnosis, it is not considered a comorbidity. Secondary diagnoses registered as a complication arising during the hospital stay are not counted as a comorbidity either.

In conformity with the collapsing procedure for other covariates, comorbidity groups registered in fewer than 50 admissions or that have no deaths are left out, as the two categories of the dummy variable are then collapsed. An exception was made for Comorbidity_17 (Severe liver disease) and Comorbidity_11 (Diabetes complications). Instead of leaving out these covariates in the case of fewer than 50 admissions or no deaths, they are first added to the less severe analogues Comorbidity_9 (Liver diseases) and Comorbidity_10 (Diabetes), respectively. If the combined comorbidities still have fewer than 50 admissions or no deaths, then these are dropped after all.

The ICD10 definitions listed in table 3.4.1 are mostly identical or nearly identical to those of Quan et al. (2005), with some adaptations, which are described in CBS (2014).

Source of admission: *home, nursing home or other institution, hospital.*
This variable indicates the patient's location before admission.

Year of discharge: *2013, 2014, 2015, 2016*.
Inclusion of the year guarantees the number of observed and expected (predicted) deaths to be equal for that year. As a result the yearly (H)SMRs have an average of 100 when weighting the hospitals proportional to their expected mortality.

Month of admission: *January/February, …, November/December*.
The months of admission are combined into 2-month periods.

3.5 Exclusion criteria

Although all hospitals mentioned in section 3.1.1 are included in the model, HSMR outcome data were not produced for all hospitals. HSMRs were only calculated for hospitals that met the criteria for LBZ participation, data quality and case mix. In addition to this, only HSMRs were calculated for hospitals that had authorised CBS to supply their HSMR figures to DHD.
Criteria used for excluding a hospital from calculating HSMRs were:

No inpatient admissions

0. Hospitals treating only day cases or outpatients are excluded, as calculation of the HSMR is not relevant for them. In fact, these hospitals do not belong to the HSMR population. Therefore, a code "0" was assigned to this criterion.

Insufficient participation in the LBZ

1. From 2014 onwards, hospitals are required to register all inpatient admissions to get HSMR outcomes. From 2015 onwards this also includes the 'prolonged observations, unplanned, without overnight stay'.

Data quality

Hospitals are excluded if*:*

2. ≥2% of inpatient admissions have a vague diagnosis code (ICD10 code R69).
3. ≤30% of inpatient admissions are coded as acute.
4. ≤0.5 secondary diagnoses are registered per inpatient admission, on average per hospital.[3]

Case mix

Hospitals are excluded if:

5. Observed mortality is less than 60 in all registered inpatient admissions.

In addition to the above-mentioned criteria, hospitals are also excluded if they had not authorised CBS to supply their HSMR figures.

3.6 Computation of the model and the (H)SMR

3.6.1 SMR and HSMR

According to the first formula in section 1.1, the SMR of hospital *h* for diagnosis *d* is written as

$$SMR_{dh} = 100\frac{O_{dh}}{E_{dh}} \tag{3.6.1}$$

with O_{dh} the observed number of deaths with diagnosis *d* in hospital *h*, and E_{dh} the expected number of deaths in a certain period. We can denote these respectively as

$$O_{dh} = \sum_i D_{dhi}, \tag{3.6.2}$$

and

$$E_{dh} = \sum_i \hat{p}_{dhi}, \tag{3.6.3}$$

where D_{dhi} denotes the observed mortality for the i^{th} admission of the combination (*d*,*h*), with scores 1 (death) and 0 (survival), and $\hat{p}_{dhi}$ the mortality probability for this admission, as estimated by the logistic regression of "mortality diagnosis *d*" on the set of covariates mentioned in section 3.4 This gives

[3] For this criterion, all secondary diagnoses are considered, even if they do not belong to the 17 comorbidity groups used as covariates. If identical secondary diagnoses (identical ICD10 codes) are registered within one admission, only one is counted. If a secondary diagnosis is identical to the main diagnosis of the admission, it is not counted as a secondary diagnosis.

$$\hat{p}_{dhi} = \text{Prob}(D_{dhi} = 1|X_{dhi}) = \frac{1}{1 + \exp(-\hat{\beta}_d' X_{dhi})}, \tag{3.6.4}$$

with X_{dhi} the scores of admission *i* of hospital *h* on the set of covariates, and $\hat{\beta}_d$ the maximum likelihood estimates of the corresponding regression coefficients, i.e. the so-called log-odds. For the HSMR of hospital *h*, we have accordingly

$$\text{HSMR}_h = 100\frac{O_h}{E_h} = 100\frac{\sum_d O_{dh}}{\sum_d E_{dh}} = 100\frac{\sum_d \sum_i D_{dhi}}{\sum_d \sum_i \hat{p}_{dhi}}. \tag{3.6.5}$$

It follows from the above formulae that:

$$\text{HSMR}_h = 100\frac{\sum_d E_{dh}\frac{O_{dh}}{E_{dh}}}{E_h} = \sum_d \frac{E_{dh}}{E_h}\text{SMR}_{dh}. \tag{3.6.6}$$

Hence, an HSMR is a weighted mean of the SMRs, with the expected mortalities across diagnoses as the weights.

3.6.2 Modelling and model-diagnostics

We estimated a logistic regression model for each of the 157 CCS diagnosis groups, using the categorical covariates mentioned in section 3.4. Computations were performed using the glm routine of the statistical software R (R Core Team, 2015). Categories, including the reference category, are collapsed if the number of admissions is smaller than 50 or when there are no deaths in the category, to prevent standard errors of the regression coefficients becoming too large. This collapsing is performed starting with the smallest category, which is combined with the smallest nearby category, etc. For variables with only two categories collapsing results in dropping the covariate out of the model (except for comorbidities 17 (Severe liver disease) and 11 (Diabetes complications) which are first combined with comorbidity 9 (Liver disease), and comorbidity 10 (Diabetes), respectively; see section 3.4). Non-significant covariates are preserved in the model, unless the number of admissions is smaller than 50 (or if there are no deaths) for all but one category of a covariate. All regression coefficients are presented in a file published together with this report.

The following statistics are presented to evaluate the models:

- *standard errors* for all regression coefficients (published with the regression coefficients);
- *statistical significance* of the covariates with significance level α=.05, i.e. confidence level .95 (see Appendix);
- *Wald statistics* for the overall effect and the significance testing of categorical variables;
- *C-statistics* for the overall fit. The C-statistic is a measure for the predictive validity of, in our case, a logistic regression. Its maximum value of 1 indicates perfect discriminating power and 0.5 discriminating power not better than expected by chance, which will be the case if no appropriate covariates are found. We present the C-statistics as an evaluation criterion for the logistic regressions.

In addition to these diagnostic measures for the regressions, we present the *average shift in HSMR* by inclusion/deletion of the covariate in/from the model (table 4.4.1). This average absolute difference in HSMR is defined as

$$\frac{1}{N}\sum_{h=1}^{N}\left|\mathrm{HSMR}_h - \mathrm{HSMR}_h^{-x_j}\right|, \tag{3.6.7}$$

where $\mathrm{HSMR}_h^{-x_j}$ is the HSMR that would result from deletion of covariate x_j, and N=81 the total number of hospitals for which an HSMR was calculated..

The Wald statistic is used to test whether the covariates have a significant impact on mortality. But it can also be used as a measure of association. A large value of a Wald statistic points to a strong impact of that covariate on mortality, adjusted for the impact of the other covariates. It is a kind of "explained chi-square". As the number of categories may "benefit" covariates with many categories, it is necessary to also take into account the corresponding numbers of degrees of freedom (df), where the df is the number of categories minus 1. As a result of collapsing the categories, the degrees of freedom can be smaller than the original number of categories minus 1.

A high Wald statistic implies that the covariate's categories discriminate in mortality rates. But if the frequency distribution of the covariate is equal for all hospitals, the covariate would not have any impact on the (H)SMRs. Therefore we also present the change in HSMRs resulting from deleting the covariate. Of course, a covariate that only has low Wald statistics has little impact on the (H)SMRs.

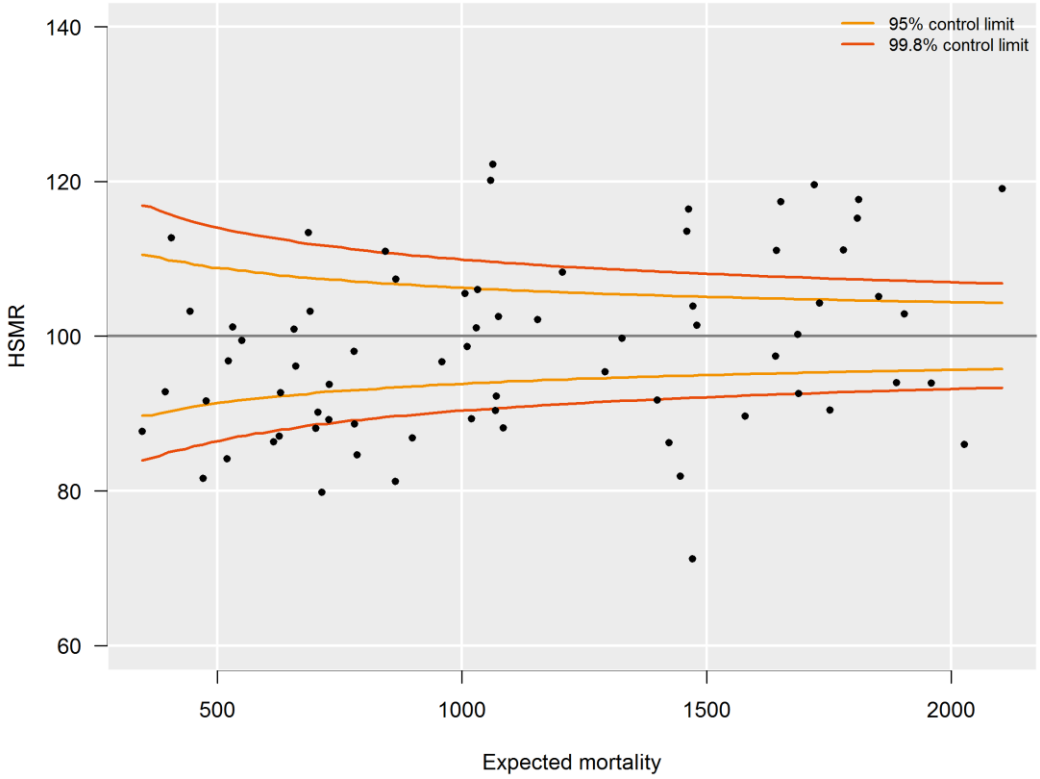
3.6.3 Confidence intervals and control limits

A confidence interval is calculated for each SMR and HSMR, i.e. an upper and lower confidence limit. For the HSMR and most SMRs a confidence level of 95 percent is used, for the SMRs of the 157 diagnosis groups a confidence level of 98 percent is used. These limits are mentioned in the specific reports for the hospitals. A lower limit above 100 indicates a statistically significant high (H)SMR, and an upper limit below 100 a statistically significant low (H)SMR. In the calculation of these confidence intervals, a Poisson distribution is assumed for the numerator of the (H)SMR, while the denominator is assumed to have no variation. This is a good approximation, since the variance of the denominator is small. As a result of these assumptions, we were able to compute exact confidence limits.

HSMRs can be presented in a funnel plot (see figure 3.6.4): a plot of hospitals, where the vertical axis represents the HSMRs and the horizontal axis the expected mortalities. Hospitals located above the horizontal axis (HSMR=100) have a higher than expected mortality. As this might be a non-significant feature, based on chance, control limits are shown in the plot for each possible expected mortality. HSMRs within these control limits do not deviate significantly from 100. In the case of 95 percent control limits, about 2.5 percent of the points would lie above the upper limit if there is no reason for differences between HSMRs, and about 2.5 percent of the points below the lower limit. The same holds, mutatis mutandis, for the 99.8 percent control limits. Here about 0.1 percent of the points would be located above the upper line if there is no reason for differences in standardised mortality rates. Most attention will be paid to this line, as points above this line have a high HSMR that is statistically very significant,

which can hardly be the result of chance alone. These hospitals would be advised to investigate the possible reasons for the significantly high values: coding errors, unmeasured case mix variables and/or suboptimal quality of care.

3.6.4 Funnel plot HSMR (example)



The precision of the HSMR is much greater for a three-year period than for a single year, which is reflected by a smaller range between the control limits. The confidence intervals of the HSMR are also smaller. Of course, drawbacks are that two consecutive three-year figures (e.g. 2013-2015 and 2014-2016) overlap, and that the three-year figure is less up-to-date than the figure of the last year. Therefore we also calculated the figures for the last available year. Observed mortality (numerator) and expected mortality (denominator) are then calculated for this year's admissions, whereas the expected mortality model of the HSMR still uses the four year data. If a hospital has a significantly high HSMR in the last year, but not in the three year period, this is a signal for further investigation, as the quality of care may have deteriorated. On the other hand, if a hospital has a significantly high HSMR for the three year period, but not in the last year, this does not necessarily mean that the situation has improved in the last year, as the one-year figures are less often significant because of the larger margins. In such cases, not only the significance should be taken into account, but also the HSMR levels over the years.

4. Evaluation of the HSMR 2016 model

This chapter presents and evaluates the model results. Some summary measures of the 157 logistic regressions are presented, with inpatient mortality as the dependent variable and the variables mentioned in section 3.4 as explanatory variables. More detailed results are presented in the Appendix, and the regression coefficients and their standard errors are presented in the file "Coefficients HSMR 2016.xls", published together with this report.

4.1 Target population and data set

Table 4.1.1 shows the number of hospitals that were included in the HMSR model. The total number of general hospitals decreased in the period 2013-2016 due to mergers. Some of the merged hospitals requested separate HSMR reports for their pre-merge locations instead of one report for the merger hospital. For these hospitals we have counted the pre-merge locations as separate hospitals in table 4.1.1.

Hospitals that did not register any (complete) inpatient records in the LBZ were not included in the HSMR model.

In 2016 all general and university hospitals could be included in the model. As the HSMR now includes all diagnosis groups, all specialised hospitals that registered complete inpatient records in the HSMR were also included.

4.1.1 Number of hospitals in HSMR model (2013-2016)

		General hospitals[a)]	University hospitals	Short stay specialised hospitals[b)]	Total hospitals
2013	Total number	82	8	4	94
	Used in model	77	8	2	87
2014	Total number	80	8	4	92
	Used in model	78	8	2	88
2015	Total number	75	8	4	87
	Used in model	75	8	3	86
2016	Total number	71	8	4	83
	Used in model	71	8	3	82

a) Excluding military hospital
b) Included are one clinic for lung diseases, one cancer hospital, one clinic for orthopaedics, rheumatic diseases and rehabilitation, and one eye hospital

For hospitals that did not register all its inpatient records (and the 'prolonged observation without overnight stay' records from 2015 onwards) completely, only the completely registered records were included in the model. For some of these hospitals only the fully registered months were included, as in the other months there were indications that fatal cases were registered completely and non-fatal cases partially. The partially registered months of these hospitals were removed from the model as these would otherwise unjustly influence the estimates. For the years 2013 to 2014 this was done for 6 and 1 hospitals, respectively. In 2015 and 2016 this did not occur.

Table 4.1.2 lists some characteristics of the admissions included in the HSMR model. Admissions of foreigners were excluded.

4.1.2 Admissions in HSMR model (2013-2016)

Excluded admissions of foreigners	29 957
Total number of admissions included in model	6 777 557
Number of inpatient admissions	*6 561 531*
Number of observations	*216 026*
Crude mortality (in admissions in model)	1.9%

4.2 Hospital exclusion

In 2016 the total LBZ population comprised 83 hospitals (table 4.1.1). Hospitals were only provided with (H)SMR outcomes if the data fulfilled the criteria stated in paragraph 3.5. In order to qualify for a three-year report (2014-2016) hospitals had to fulfil these criteria for the three consecutive years.

Of the 83 hospitals, the four short stay specialised hospitals and one general hospital had not been asked to grant authorization for providing HSMR numbers because their case mix was very different from that of other hospitals. In fact, four of these hospitals had participated in the LBZ but their data did not meet one or more of the previously stated criteria, such as a minimum number of 60 registered deaths or on average a minimum number of 0.5 comorbidities per admission. All of the 78 hospitals that had granted authorization fulfilled the criteria and were provided with a HSMR figure for 2016.

For these 78 hospitals the data of 2015 and 2014 was additionally scrutinized in order to determine if a three-year report could be provided. Five hospitals did not meet one or more criteria in (one of) those years: no participation (2), partial response (2), <0,5 secondary diagnoses registered per inpatient admission, on average per hospital (1). As a result, the data of the other 73 hospitals met the criteria in all years considered and so these hospitals were provided with three-year HSMR figures.

4.3 Impact of the covariates on mortality and HSMR

The table in the appendix shows which covariates have a statistically significant (95 percent confidence) impact on in-hospital mortality for each of the 157 diagnosis groups: “1” indicates (statistical) significance, and “0” non-significance, while a dash (-) means that the covariate has been dropped as the number of admissions is smaller than 50 (or as there are no deaths) for all but one category of a covariate; see section 3.6.2. The last row of the table in the Appendix gives the numbers of significant results across the diagnosis groups for each covariate. These values are presented again in table 4.3.1 below, as a summary, but ordered by the number of times a covariate is significant. Age, urgency of the admission, and severity of the main diagnosis are significant for the great majority of the diagnosis groups. This is also true for several of the comorbidity groups, especially groups 2 and 13, i.e. for Congestive heart failure and Renal disease. Comorbidity 15 is seldom registered as a comorbidity; most diagnosis groups had fewer than 50 admissions with HIV comorbidity. Compared to last year (CBS, 2016) the number of times year of discharge was significant has dropped from 72 to 43. For the other covariates the changes are small.

4.3.1 Statistical significance of the covariates for the 157 logistic regressions (summary), HSMR 2016 model

Covariate	No. of significant results	Covariate	No. of significant results
Age	144	Sex	51
Comorbidity 2	133	Comorbidity 5	50
Urgency	125	Comorbidity 17	49
Comorbidity 13	113	Comorbidity 10	44
Severity main diagnosis	112	Year of discharge	43
Comorbidity 9	103	Comorbidity 11	32
Comorbidity 16	99	Month of admission	30
Comorbidity 3	91	Comorbidity 7	27
Comorbidity 14	86	Comorbidity 12	24
Source of admission	85	SES	18
Comorbidity 6	82	Comorbidity 8	14
Comorbidity 4	73	Comorbidity 15	1
Comorbidity 1	63		

4.3.2 Wald chi-square statistics for the 157 logistic regressions, HSMR 2016 model

Covariate	Sum of Wald statistics	Sum of df	Covariate	Sum of Wald statistics	Sum of df
Severity main diagnosis	39 860	400	Comorbidity 17	1 418	62
			Month of admission	1 232	782
Age	34 482	2 037	Year of discharge	1 200	471
Urgency	17 772	156	Comorbidity 1	1 154	147
Comorbidity 2	8 998	142	Comorbidity 5	1 049	118
Comorbidity 16	4 463	138	Sex	903	150
Comorbidity 13	3 919	148	SES	821	692
Source of admission	2 676	274	Comorbidity 10	477	153
Comorbidity 3	2 236	145	Comorbidity 12	474	95
Comorbidity 9	2 094	133	Comorbidity 11	401	121
Comorbidity 6	1 978	153	Comorbidity 7	333	123
Comorbidity 14	1 840	145	Comorbidity 8	152	33
Comorbidity 4	1 433	136	Comorbidity 15	17	6

The significance only partially measures the effect of the covariates on mortality. This is better measured using the Wald statistic. Table 4.3.2 presents the total Wald statistics summed over all diagnosis groups for each of the covariates with the respective sum of the degrees of freedom, ordered by value. It shows that severity of main diagnosis has the highest explanatory power. Age and urgency of admission are also important variables. Of the comorbidities comorbidity groups 2, 13 together with 16 are the groups with the most impact on mortality. The explanatory powers of Sex and SES are relatively small. This is also true for a number of

comorbidity groups. Compared to last year (CBS, 2016) the impact of year of discharge has also dropped when looking at the Wald statistic (from 2409 to 1200). This implies that the differences in mortality (corrected for differences in patient characteristics) between the years in the model are decreasing.

As was mentioned in section 3.6.2, when the hospitals differ little on a covariate, the effect of this covariate on the HSMR can still be small even if this covariate is strong predictor for mortality. Table 4.3.3 shows the impact of each covariate on the HSMR, as measured by formula (3.6.7), for the hospitals for which HSMRs are calculated. The comorbidities as a group have the largest effect on the HSMR. This is caused by differences in case mixes between hospitals, but possibly also by differences in coding practice (see Van der Laan, 2013). Notice that we consider the comorbidities as one group here. Deleting Sex hardly has an impact on the HSMRs. Compared to Sex, SES has a reasonable impact on the HSMR 2016. This is because hospitals differ more in terms of SES categories of the postal areas in their vicinity than in terms of the sex distribution of their patients. Overall the differences in the effect of the covariates between this year and the year before are small.

4.3.3 Average shift in HSMR 2016 by inclusion/deletion of covariates

Covariate	Average shift in HSMR	Covariate	Average shift in HSMR
Comorbidity [a)]	6.34	Source of admission	1.12
Age	4.53	SES	0.61
Severity main diagnosis	2.51	Month of admission	0.12
Urgency	2.38	Sex	0.12

a) The comorbidities were deleted as one group and not separately.

4.4 Model evaluation for the 157 regression analyses

Table 4.4.1 presents numbers of admissions and deaths, and C-statistics for the 157 diagnosis groups. The C-statistic is explained in section 3.6.2. Because of an approximation used in the routine used to calculate the C-statistic in previous years, some of the C-statistics reported in previous years were underestimated. This was revealed when we switched from using lrm to glm to estimate the models (see chapter 2). Fortunately, this means that the actual quality of the models was better than previously reported. For the following diagnosis groups the C-statistic has increased more than 0.05 (and ≤ 0.10) compared to last year (this is probably largely caused by the aforementioned problem, but can also be caused by an actual better fit):

- “Other and unspecified benign neoplasm” (37),
- “Mental, affective, anxiety, somatoform, dissociative, and personality disorders” (47),
- “Coma, stupor, and brain damage” (56),
- “Headache and other disorders of the sense organs” (57),
- “Diverticulosis and diverticulitis” (100),
- “Prolapse and other female genital disorders” (117),
- “Other perinatal conditions” (131) ,
- “Joint disorders and dislocations; trauma-related; sprains and strains” (132),
- “Allergic reactions” (155),
- “Rehabilitation and other aftercare, medical examination/evaluation/screening” (156).

And for the following diagnosis groups the increase was more than 0.10:

- "Schizophrenia, mental retardation, preadult disorders and other mental cond" (49),
- "Nonspecific chest pain" (64),
- "Hyperplasia of prostate and other male genital disorders" (115),
- "Nonmalignant breast conditions" (116),
- "Prolapse and other female genital disorders" (117),
- "Complications of pregnancy, childbirth, and the puerperium; liveborn" (118),
- "Osteoarthritis, rheumatoid arthritis, and other musculoskeletal deformities" (122).

Most of the groups above are groups with a relatively low mortality rate. The approximation used by the previously used routine gives lower C-statistics in such cases.

Only two of the 157 diagnosis groups have a C-statistic below 0.70: "Congestive heart failure, nonhypertensive" (70) and "Aspiration pneumonitis; food/vomitus" (84). This was also the case in previous years. In previous years diagnosis group "Complications of pregnancy, childbirth, and the puerperium; liveborn" (118) also had a score below 0.70. However, because of the new routine used to calculate the C-statistic and maybe also partly because the variable *urgency of a admission* was adapted for liveborn infants (see chapter 2), the C-statistic of this group has increased to 0.88.

For the two diagnosis groups with a C-statistic below 0.7, the model is only partially able to explain patient mortality. This increases the risk that the model does not completely correct for population differences between the hospitals. For the highest scoring diagnosis groups (above 0.9) the covariates strongly reduce the uncertainty in predicting patient mortality.

4.4.1 C-statistics for the logistic regressions of the 157 main diagnosis groups, HSMR 2016 model

Diag. group no.	Description diagnosis group	Number of admissions	Number of deaths	C-statistic
1	Tuberculosis	1 697	33	0.90
2	Septicemia (except in labor)*	22 573	5 830	0.75
3	Bacterial infection; unspecified site	6 423	314	0.81
4	Mycoses	2 228	201	0.82
5	HIV infection	1 365	51	0.80
6	Hepatitis, viral and other infections	26 208	240	0.91
7	Cancer of head and neck	14 752	247	0.89
8	Cancer of esophagus*	10 436	554	0.78
9	Cancer of stomach*	12 911	475	0.79
10	Cancer of colon*	50 855	1 342	0.83
11	Cancer of rectum and anus*	25 618	521	0.84
12	Cancer of liver and intrahepatic bile duct	6 743	376	0,79
13	Cancer of pancreas*	16 382	765	0,80
14	Cancer of other GI organs; peritoneum	7 196	353	0,79
15	Cancer of bronchus; lung*	74 692	4 355	0.84
16	Cancer; other respiratory and intrathoracic	3 179	141	0.88
17	Cancer of bone and connective tissue	7 960	93	0.93
18	Melanomas of skin and other non-epithelial cancer of skin	7 399	85	0.92
19	Cancer of breast*	50 789	398	0.95
20	Cancer of uterus	8 094	103	0.93

21	Cancer of cervix and other female genital organs	10 863	95	0.91
22	Cancer of ovary	9 790	256	0.85
23	Cancer of prostate*	24 423	473	0.91
24	Cancer of testis and other male genital organs	5 097	12	0.97
25	Cancer of bladder*	48 741	408	0.91
26	Cancer of kidney, renal pelvis and other urinary organs	14 368	285	0.88
27	Cancer of brain and nervous system	11 280	254	0.75
28	Cancer of thyroid	5 308	39	0.97
29	Hodgkin`s disease	2 418	35	0.86
30	Non-Hodgkin`s lymphoma*	21 049	886	0.83
31	Leukemias*	22 625	1 152	0.79
32	Multiple myeloma	9 867	509	0.78
33	Cancer; other and unspec. primary; maintenance chemotherapy and radioth.	11 391	166	0.94
34	Secondary malignancies*	81 381	4 134	0.78
35	Malignant neoplasm without specification of site	6 538	476	0.84
36	Neoplasms of unspecified nature or uncertain behavior*	13 469	247	0.86
37	Other and unspecified benign neoplasm	73 242	112	0.90
38	Thyroid and other endocrine disorders	24 587	173	0.90
39	Diabetes mellitus without complication	18 332	91	0.91
40	Diabetes mellitus with complications*	24 334	438	0.85
41	Nutritional deficiencies and other nutritional, endocrine, and metabolic disorders	53 825	397	0.94
42	Fluid and electrolyte disorders*	32 587	791	0.85
43	Cystic fibrosis	2 766	25	0.88
44	Immunity and coagulation disorders, hemorrhagic disorders	11 850	188	0.87
45	Deficiency and other anemia*	45 493	385	0.81
46	Diseases of white blood cells	7 629	153	0.79
47	Mental, affective, anxiety, somatoform, dissociative, and personality disorders	33 638	85	0.86
48	Senility and organic mental disorders	11 986	508	0.75
49	Schizophrenia, mental retardation, preadult disorders and other mental cond.	8 292	25	0.92
50	Other psychoses	4 930	61	0.83
51	Meningitis, encephalitis, and other central nervous system infections	9 333	447	0.89
52	Parkinson`s disease	6 637	64	0.85
53	Multiple sclerosis and other degenerative nervous system conditions	15 637	273	0.90
54	Paralysis and late effects of cerebrovascular disease	4 916	76	0.86
55	Epilepsy and convulsions	46 830	515	0.88
56	Coma, stupor, and brain damage*	3 641	487	0.91
57	Headache and other disorders of the sense organs	85 997	48	0.93
58	Other nervous system disorders	121 339	331	0.97
59	Heart valve disorders*	37 908	1 118	0.79
60	Peri-, endo-, myocarditis, and cardiomyopathy	20 821	588	0.87
61	Essential hypertension, hypertension with compl., and secondary hypertension	14 967	132	0.95

62	Acute myocardial infarction*	122 109	4 246	0.81
63	Coronary atherosclerosis and other heart disease*	162 300	1 127	0.86
64	Nonspecific chest pain	167 819	96	0.89
65	Pulmonary heart disease*	33 790	1 126	0.80
66	Other and ill-defined heart disease	3 631	210	0.82
67	Conduction disorders (heart disease)	21 792	319	0.86
68	Cardiac dysrhythmias*	179 095	993	0.89
69	Cardiac arrest and ventricular fibrillation*	11 873	4 358	0.73
70	Congestive heart failure, nonhypertensive*	111 185	9 176	0.68
71	Acute cerebrovascular disease*	126 333	12 316	0.80
72	Transient cerebral ischemia, and other cerebrovascular disease	51 053	529	0.92
73	Peripheral and visceral atherosclerosis*	34 285	1 719	0.90
74	Aortic and other artery aneurysms*	28 483	2 543	0.89
75	Aortic and arterial embolism or thrombosis*	20 727	520	0.88
76	Other circulatory disease*	33 479	599	0.87
77	Phlebitis, varicose veins, and hemorrhoids	17 035	139	0.91
78	Pneumonia*	136 592	10 301	0.77
79	Influenza	8 817	333	0.85
80	Tonsillitis and upper respiratory infections	95 477	121	0.96
81	Acute bronchitis	26 386	107	0.94
82	Chronic obstructive pulmonary disease and bronchiectasis*	131 856	5 897	0.70
83	Asthma	31 885	106	0.90
84	Aspiration pneumonitis; food/vomitus*	6 704	1 549	0.67
85	Pleurisy; pneumothorax; pulmonary collapse*	25 627	654	0.84
86	Respiratory failure; insufficiency; arrest	7 963	2 393	0.76
87	Lung disease due to external agents	1 880	164	0.80
88	Other lower respiratory disease*	29 147	1 032	0.87
89	Other upper respiratory disease	93 359	797	0.91
90	Intestinal infection	43 523	451	0.90
91	Disorders of mouth, teeth, and jaw	21 490	41	0.96
92	Esophageal disorders	16 259	140	0.89
93	Gastroduodenal ulcer	4 739	234	0.92
94	Gastritis, duodenitis, and other disorders of stomach and duodenum	8 983	82	0.89
95	Appendicitis and other appendiceal conditions	61 926	56	0.97
96	Peritonitis and intestinal abscess	4 511	272	0.86
97	Abdominal hernia	52 105	374	0.94
98	Regional enteritis and ulcerative colitis	18 852	54	0.93
99	Intestinal obstruction without hernia*	32 880	1 528	0.83
100	Diverticulosis and diverticulitis*	38 962	488	0.92
101	Anal and rectal conditions	25 930	51	0.95
102	Biliary tract disease*	140 789	763	0.92
103	Liver disease; alcohol-related*	6 574	818	0.73
104	Other liver diseases*	17 942	914	0.81
105	Pancreatic disorders (not diabetes)	29 380	583	0.85
106	Gastrointestinal hemorrhage*	36 428	1 100	0.81
107	Noninfectious gastroenteritis	30 398	340	0.87
108	Other gastrointestinal disorders*	45 911	677	0.94

109	Nephritis; nephrosis; renal sclerosis	15 314	75	0.93
110	Acute and unspecified renal failure*	17 434	1 087	0.78
111	Chronic kidney disease*	15 420	494	0.88
112	Urinary tract infections*	83 576	1 892	0.80
113	Calculus and other diseases of urinary tract	84 511	175	0.94
114	Genitourinary symptoms and ill-defined conditions	32 427	126	0.89
115	Hyperplasia of prostate and other male genital disorders	45 865	59	0.94
116	Nonmalignant breast conditions	18 154	4	0.99
117	Prolapse and other female genital disorders	76 506	47	0.96
118	Complications of pregnancy, childbirth, and the puerperium; liveborn	663 629	40	0.88
119	Skin and subcutaneous tissue infections	54 621	550	0.90
120	Other skin disorders, chronic ulcer of skin	23 503	243	0.92
121	Infective arthritis and osteomyelitis	13 089	237	0.91
122	Osteoarthritis, rheumatoid arthritis, and other musculoskeletal deformities	251 354	238	0.93
123	Other non-traumatic joint disorders	19 569	51	0.94
124	Spondylosis, back problems, and osteoporosis	97 290	145	0.96
125	Pathological fracture	7 477	102	0.80
126	Other connective tissue disease	58 363	232	0.97
127	Cardiac and circulatory congenital anomalies	9 612	154	0.84
128	Noncardiac congenital anomalies	32 561	233	0.93
129	Short gestation; low birth weight; and fetal growth retardation	62 240	633	0.87
130	Intrauterine hypoxia, perinatal asphyxia, and jaundice	55 822	201	0.94
131	Other perinatal conditions	202 350	222	0.92
132	Joint disorders and dislocations; trauma-related; sprains and strains	38 987	32	0.98
133	Fracture of neck of femur (hip)*	74 088	2 379	0.80
134	Skull and face fractures, spinal cord injury	13 048	246	0.89
135	Fracture of upper limb	52 719	141	0.94
136	Fracture of lower limb	54 898	334	0.94
137	Other fractures	46 100	847	0.87
138	Intracranial injury*	50 585	2 055	0.88
139	Crushing injury or internal injury	23 663	298	0.93
140	Open wounds of head; neck; and trunk	7 670	60	0.90
141	Open wounds of extremities	6 589	37	0.95
142	Complication of device, implant or graft*	96 272	1 255	0.87
143	Complications of surgical procedures or medical care*	99 342	955	0.87
144	Superficial injury; contusion	46 117	378	0.90
145	Burns	4 526	78	0.96
146	Poisoning by psychotropic agents, drugs, or other medications	36 576	277	0.84
147	Other injuries and conditions due to external causes	12 031	611	0.89
148	Syncope	55 851	203	0.85
149	Fever of unknown origin	29 502	210	0.83
150	Lymphadenitis and gangrene	7 900	80	0.93
151	Shock*	1 755	727	0.70
152	Nausea and vomiting	15 991	112	0.86
153	Abdominal pain	67 119	271	0.94

154	Malaise and fatigue	15 089	291	0.80
155	Allergic reactions	11 703	29	0.95
156	Rehabilitation and other aftercare, medical examination/evaluation/screening	155 856	364	0.86
157	Residual codes; unclassified	76 133	412	0.95

*Diagnosis groups present in the HSMR models up to 2014. Diagnosis group 45 then only contained CCS group 59 ("Deficiency and other anemia"); from the HSMR 2015 onwards CCS group 60 ("Acute posthemorrhagic anemia") was added to this group.

4.5 Regression coefficients

The file "coefficients HSMR 2016.xls" contains the estimated regression coefficients (columns "Estimate"), also called "log-odds", for each of the 157 logistic regressions, as well as their standard errors (columns "Std. Err."). The estimated regression coefficients are the elements of the vector $\hat{\beta}_d$ in formula (3.6.4), for each diagnosis d. Notice that a β-coefficient has to be interpreted as the difference in log-odds between the category in question and the reference category (first category of the same covariate). For the sake of clarity, the reference categories are given in the first row of the corresponding covariates, and by definition have zero coefficient for each regression.

In many cases categories are collapsed (see section 3.6.2). This results in equal coefficients for the collapsed categories. If all categories were collapsed into one category for a certain variable and for a certain diagnosis group (i.e. if there was only one category with ≥50 admissions and ≥1 death), the variable was dropped from the model and all associated coefficients are set to zero. Therefore, one can directly use the coefficients in the file "coefficients HSMR 2016.xls" to calculate mortality probabilities, with the exception of two of the Charlson comorbidities (Comorbidity 17 and Comorbidity 11). If Charlson comorbidity 17 (Severe liver disease) contains <50 admissions or no mortality, it is collapsed with Charlson comorbidity 9 (Liver disease). In this case the coefficient of Comorbidity 17 is set to zero. When a patient has both comorbidities, it counts as only one comorbidity. Therefore, when the coefficient of Comorbidity 17 is zero in the coefficients file, one should first recode all Charlson 17 comorbidities to Comorbidity 9 and use the coefficient of Comorbidity 9. The same holds for Charlson 11 (Diabetes complications) when it is collapsed with Charlson 10 (Diabetes).

5. Recalculation of the severity of main diagnosis

5.1 Introduction

One the variables used in the model of the HSMR is the severity of the main diagnosis. For each of the 157 main diagnosis groups a separate model is estimated. However, these main diagnosis groups usually contain different main diagnoses. There can be substantial differences in the mortality between the diagnoses within one diagnosis group. In order to correct for these differences, each main diagnosis is assigned a severity, which is a categorical variable with 8 levels of increasing severity with an additional 'Other' level for diagnoses for which the mortality cannot be determined. The severity is determined by calculating the crude mortality for each main diagnosis on historical data. The mortality is then recoded into the 9 severity classes. For determining the severity six years of historical data is used. For the HSMR of 2015 (model 2012-2015) data from 2006-2011 was used. Each year the years used for the historical dataset should shift by one year, therefore for the HSMR of 2016 this would be 2007-2012. However, 2012 is the first year in which part of the hospitals shifted from coding their main diagnoses in ICD9-CM to ICD10. This means we now have a mix of coding systems in the historical dataset.

Since 2013 all hospitals code their diagnoses using ICD10. However, in previous HSMR models (up to HSMR 2015) the severities were determined for ICD9 diagnoses in the historical dataset. In order to link these severities to the ICD10 coded data used in the HSMR model the ICD10 codes were translated to their (default) ICD9 equivalents using a conversion table. The ICD10-ICD9 mapping is however not always one to one: multiple ICD10 codes can belong to one ICD9 code, and multiple ICD9 codes can belong to one ICD10 code. In general ICD10 codes are more detailed, and therefore, it is more frequent that multiple ICD10 codes map onto one ICD9 code.

In the present situation it would be preferable if severity was determined using ICD10, because (1) when converting ICD10 to ICD9 information is lost (because of using a default code and because ICD10 has more differentiated codes), (2) the historical data set now also partly consists of ICD10 data, and (3) a severity classification in ICD10 is easier to use as all admissions are currently coded in ICD10.

Therefore it was decided that the severity will be determined for ICD10 diagnoses for the model of the HSMR 2016 and later years. We developed a method to calculate the ICD10 severities on the basis of a mixed ICD9/ICD10 historical dataset, ensuring a gradual shift over time from severities based on ICD9 data to severities based solely on ICD10 data. This method is described in this chapter. Furthermore, the results of the new method are compared to those of the previous method. To ensure that we had enough diagnoses coded in ICD10, it was decided to calculate the severities for the HSMR 2016 model using data from 2008-2013. This means that we had approximately 1.5 years of data in ICD10 (approximately half of the hospitals coded in ICD10 in 2012, and nearly all hospitals coded in ICD10 in 2013) and 4.5 years of data in ICD9.

Both the new method and the previous method to calculate severities make use of a Dutch ICD10 to ICD9 conversion table (see section 3.4). As this table had not been updated for recent years, new ICD10 codes added to the ICD10 in recent years did not have a converted ICD9 code. For the HSMR 2015 this resulted in a substantial amount of admissions without a severity code (these are classified into severity category 'Other'). As this is undesirable, we decided, as part of this update, to extend the conversion table with (default) ICD9 translations of these new ICD10

codes. The new translations were made in consultation with DHD. Furthermore, for the new ICD10 codes that had been registered in a substantial number of admissions, also the old ICD10 code (that was used prior to the introduction of the new code) was added to the conversion table, as this was relevant for the new method to calculate severities (see section 5.2).

5.2 Methods

We will assume that for each ICD10 diagnosis i there is a single unique ICD9 diagnoses k. An ICD9 diagnosis can belong to multiple ICD10 diagnoses. Let m_k be the number of ICD10 diagnoses that belong to ICD9 diagnosis k. Furthermore, call the number of admissions in k n_k and the number of deaths d_k. Therefore, the data has the following structure (j is another ICD10 code belonging to ICD9 code k):

ICD9			ICD10			
			1	⋮	⋮	⋮
				i	n_i	d_i
k	n_k	d_k	⋮	⋮	⋮	⋮
				j	n_j	d_j
			m_k	⋮	⋮	⋮

The goal is to obtain an estimate of the mortality in i using data from 2012 and 2013 from i and 2008 up until 2012 from k. The reason to combine both is twofold. First, the amount of data in 2012 and 2013 is too small for reliable estimates. Second, this ensures a gradual transition from severities based in ICD9 to severities based solely on ICD10.

One way to look at this problem is to see the mortality following from the number of admissions and deaths coded in ICD9 as a prior estimate for the mortality at the ICD10 level. When one assumes that the prior probability follows a beta distribution with parameters d_k and $n_k - d_k$, the maximum posteriori (MAP) estimate for the mortality of i is given by (Mitchell, 2016)

$$\hat{p}_i^{MAP} = \frac{d_i + d_k - 1}{n_i + n_k - 2}.$$

However, when there are multiple ICD10 diagnoses that belong to one ICD9 diagnosis, the number of admissions for the ICD9 diagnosis will be relatively high compared to the number of admissions for the individual ICD10 diagnoses as the ICD9 diagnosis will also contain all admissions belonging to the other ICD10 diagnoses. The admissions and deaths for the ICD9 diagnosis will, therefore, have a large impact in the formula above. Furthermore, when there are multiple ICD10 diagnoses belonging to one ICD9 diagnosis, it will be more unlikely that the mortality of the ICD9 diagnosis is a strong indicator for the mortality of one individual ICD10 diagnosis. Therefore, it was decided to weigh the deaths and admissions for the ICD9 diagnosis with the number of ICD10 diagnoses belonging to the ICD9 diagnosis:

$$\hat{p}_i^{WMAP} = \frac{d_i + {}^{d_k}/_{m_k}}{n_i + {}^{n_k}/_{m_k}}$$

The differences between these two estimators are investigated further in section 5.3.

New ICD10 diagnoses

The ICD10 coding system is occasionally being updated. This means that new codes can be added and old codes can disappear. In general old codes will be replaced by newer more detailed codes. Therefore, in general the situation will be as below, where the old ICD10 'C' has been replaced by 'D' and 'E'.

ICD9	ICD10 old	ICD10 new
A	B	B
A	C	D
A	C	E

However, more complex situations also occur, such as

ICD9	ICD10 old	ICD10 new
F	H	H
F (default code for I)	I	J
G	I	K

where the old ICD10 code 'I' (to which belongs the default ICD9 code 'F') has been replaced by 'J' and 'K'. In this case 'K' can be translated to the old code 'I' which in turn can be translated to the ICD9 code 'F'. However, the new code 'K' can also be more accurately translated to ICD9 code 'G'. In this case it was decided to use the following formula for the mortality for 'K':

$$\hat{p}_K^{WMAP} = \frac{d_K + d_I + {}^{d_G}/_{m_G}}{n_K + n_I + {}^{n_G}/_{m_G}}.$$

This estimator uses the admissions and deaths from the more accurate ICD9 code 'G' and the prior ICD10 code 'I'. In this case m_G is equal to one as there is only one ICD10 code belonging to 'G' namely 'K'. For 'J' the formula looks like:

$$\hat{p}_J^{WMAP} = \frac{d_J + d_I + {}^{d_F}/_{m_F}}{n_j + n_I + {}^{n_F}/_{m_F}}.$$

where m_F is equal to three as there are three ICD10 codes belonging to 'F' namely 'H', 'I' and 'J'.

Minimal requirements for a severity class

To prevent too much uncertainty in the severity class assigned to a diagnosis and to prevent that a few hospitals determine the severity class, there are restrictions on the number of admissions with a given diagnosis and on the number of hospitals using a given diagnosis code. A diagnosis code should be used by at least 4 hospitals and there should be at least 20 admissions with the given code (Van der Laan *et al.*, 2016). Since the current severity codes are determined using a mixture of ICD9 and ICD10 admissions, we need to specify how these requirements are applied in the new situation.

It was decided to use the same criteria (a minimum of 20 admissions and 4 hospitals). When determining the number of admissions for a given diagnosis both the admissions coded in ICD10 and those coded in ICD9 are taken into account, therefore,

$$n_i + n_k \geq 20,$$

and the number of hospitals is determined by selecting all hospitals using code i and all hospitals using code k, removing duplicate hospitals and counting the number of resulting hospitals.

5.3 Results

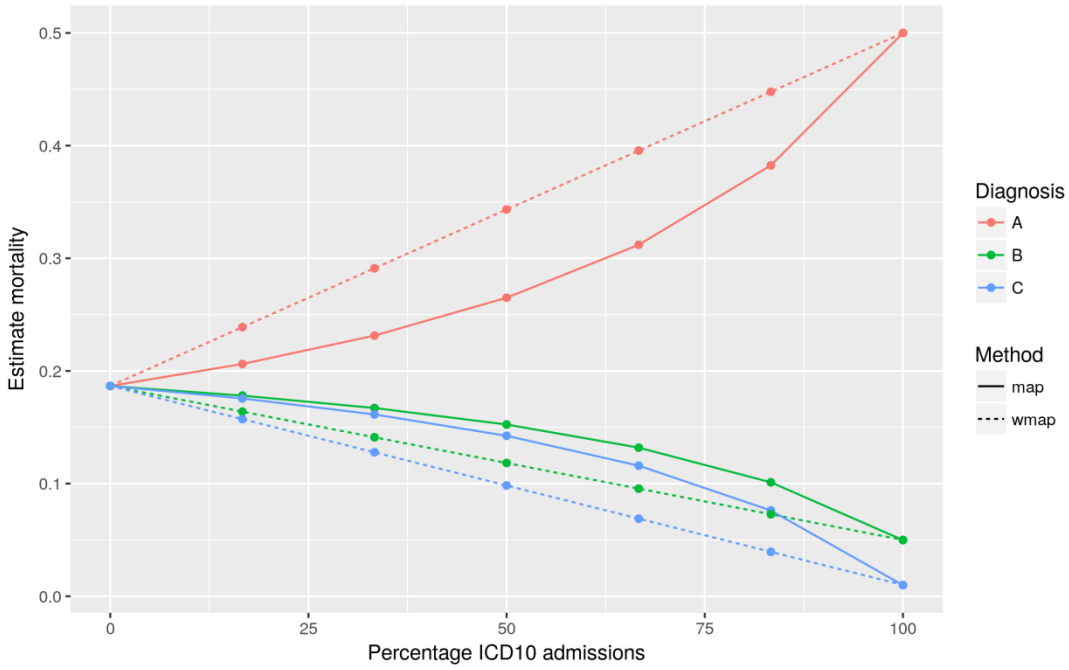
Transition from ICD9 to ICD10

In order to investigate how the severities transition from being completely based on ICD9 admissions to being completely based on ICD10 admissions works out, the two estimators are compared using different fractions of ICD9 and ICD10 admissions. We assume that we have 6 years of data. The number of years with ICD10 data varies from 0 (0% of ICD10 admissions) to 6 (100% of ICD10 admissions). Figures 5.3.1 and 5.3.2 show two examples. Figure 5.3.1 show the situation in which each ICD10 diagnosis has the same number of admissions (1000). Each ICD10 diagnosis has a different mortality rate (0.5, 0.05 and 0.01). When we have one year of ICD10 data, we have, for example, the following numbers:

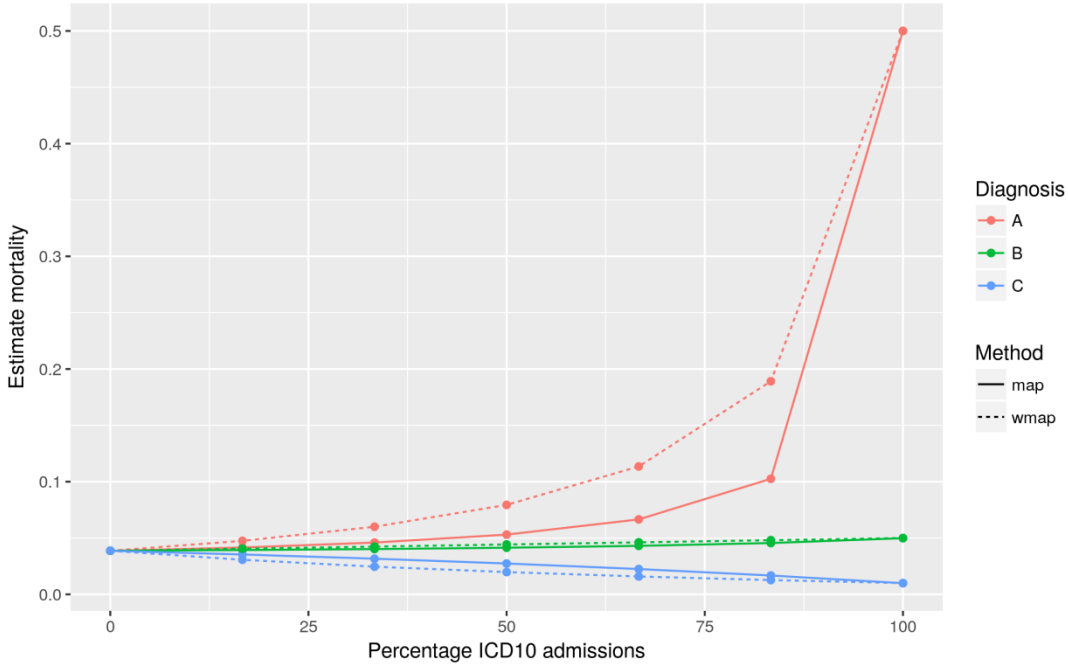
n_k	d_k	m_k		n_i	d_i	$\hat{p}_i^{MAP}$	$\hat{p}_i^{WMAP}$
			A	1 000	500	0.21	0.24
15 000	2 800	3	B	1 000	50	0.18	0.16
			C	1 000	10	0.18	0.16

The figure shows these numbers for different numbers of years coded in ICD10. Figure 5.3.2 shows the same, but for the situation where the numbers of admissions differ for the different ICD10 diagnoses.

5.3.1 Transition of the estimated mortality from ICD9 to ICD10 for a ICD9 class that is converted into three ICD10 classes. All ICD10 classes have the same number of admissions. One has a mortality rate of 0.5, one of 0.05 and of 0.01.



5.3.2 Transition of the estimated mortality from ICD9 to ICD10 for a ICD9 class that is converted into three ICD10 classes. The ICD10 classes have different numbers of admissions. One has a mortality rate of 0.5 (100 admissions per year), one of 0.05 (1000 admissions per year) and of 0.01 (2000 admissions per year).



Both figures show that the MAP-estimator has a tendency to put too much weight on the ICD9 admissions. The estimates tend to stay close to the mortality of the ICD9 diagnoses even when a large number of years are coded in ICD10. The WMAP estimator results in a smooth transition from ICD9 to ICD10. The second figure also shows that in case of an ICD10 code that is relatively little used, the estimates tend to stay closer to the ICD9 mortality, but also here the WMAP estimator gives a smoother transition. Therefore, it was decided to use the WMAP estimator.

Comparison to the severities used for the HSMR of 2015

Table 5.3.3 shows for each ICD10 diagnosis code in the complete ICD10 table, the new 2016 severity classification based on the WMAP estimate (calculated on the historical dataset 2008-2013) and the classification for the HSMR 2015 (calculated on the historical dataset 2006-2011) which was based completely on admissions coded in ICD9. In most cases the severity class remains the same (the admission in the diagonal of the table) or shifts a class up or down. What can also be seen is that the number of ICD10 diagnosis codes with a missing severity class (which is translated into 'Other' when estimating the model) and a severity class of 'Other' is reduced. The large number of ICD10 codes with a severity class 'Missing' in the old severity classification (n=3,813) are predominantly caused by the fact that the ICD10-ICD9 conversion table used for the HSMR 2015 was not yet updated, while for the new classification we used an updated conversion table (see section 5.1). The remaining 'Missing' codes in the new severity classification (n=194) are caused by the fact that some ICD10 codes (and their corresponding ICD9 codes) do not appear in the historical data set used although they are valid codes that could have been used.

5.3.3 Comparison of the 2016 severity classes based on a combination of ICD9 and ICD10 coded admissions, and the 2015 severity classes based completely on ICD9 coded admissions. The cells show the number of ICD10 codes with a given severity class.

Severity 2015	Severity 2016										
	[0,0.01)	[0.01,0.02)	[0.02,0.05)	[0.05,0.1)	[0.1,0.2)	[0.2,0.3)	[0.3,0.4)	[0.4,1]	Other	Missing	Total
[0,0.01)	4 745	325	109	12	2	0	0	0	197	6	5 396
[0.01,0.02)	153	433	178	15	2	0	0	0	4	1	786
[0.02,0.05)	71	113	762	145	7	0	0	0	25	0	1 123
[0.05,0.1)	13	4	67	321	54	0	0	0	13	0	472
[0.1,0.2)	3	0	8	38	112	14	0	0	6	0	181
[0.2,0.3)	0	0	1	1	7	27	0	0	2	0	38
[0.3,0.4)	1	0	1	1	1	1	18	3	0	0	26
[0.4,1]	0	0	0	0	0	0	0	24	0	0	24
Other	126	0	18	19	3	0	0	0	1 537	187	1 890
Missing	970	85	161	77	35	13	3	6	2 463	0	3 813
Total	6 082	960	1 305	629	223	55	21	33	4 247	194	13 749

5.3.4 Comparison of the 2016 severity classes based on a combination of ICD9 and ICD10 coded admission, and the 2015 severity classes based completely on ICD9 coded admissions. The cells show the number of admissions in 2015 for each combination of severity classes.

Severity 2015	Severity 2016									
	[0,0.01)	[0.01,0.02)	[0.02,0.05)	[0.05,0.1)	[0.1,0.2)	[0.2,0.3)	[0.3,0.4)	[0.4,1]	Other	Total
[0,0.01)	993 329	7164	2 096	178	3	0	0	0	284	1 003 054
[0.01,0.02)	55 001	114 480	10 889	55	0	0	0	0	0	180 425
[0.02,0.05)	5 787	20 182	243 407	17 946	96	0	1	0	34	287 453
[0.05,0.1)	151	296	37 812	141 484	825	13	14	0	46	180 641
[0.1,0.2)	30	38	307	34 784	30 431	480	0	0	12	66 082
[0.2,0.3)	0	0	0	0	1 025	6 362	108	0	0	7 495
[0.3,0.4)	0	0	0	0	0	0	7 753	0	0	7 753
[0.4,1]	0	0	0	0	0	0	682	3 493	0	4 175
Other	10 372	81	709	226	148	23	0	0	3 296	14 855
Missing	57 177	3 894	3 037	1 645	206	3	0	205	746	66 913
Total	1 121 847	146 135	298 257	196 318	32 734	6 881	8 558	3 698	4 418	1 818 846

It is more of interest to see the effect in the new classes on the admissions used in the HSMR model. Therefore, table 5.3.4 shows for the admissions of 2015 the old and new classes. Again most admissions are on or around the diagonal: the severity has remained the same or has shifted one class. There seems to be small shift to lower codes. This could be caused by the fact that the hospital mortality rate has been decreasing in the years of the historical dataset. The main thing that can be noticed is that the number of admissions with a missing severity or a severity class equal to 'other' has decreased substantially from 81 768 (4.5% of the admissions)

to 4 418 (0.2% of the admissions). As mentioned earlier, this is mainly caused by the fact that for the 2016 severity classes the ICD10-ICD9 conversion table was extended for new ICD10 codes. When this updated table is also used to link the old severity classes to the 2015 data, the number of admissions with a missing severity class or a severity class equal to 'other' is much smaller: 15 337 (0.8% of the admissions), see table 5.3.5.

5.3.5 Effect of using the new ICD10-ICD9 conversion table also for the 2015 severity classes. Comparison of the 2016 severity classes based on a combination of ICD9 and ICD10 coded admissions, and the 2015 severity classes based completely on ICD9 coded admissions when using the new extended conversion table. The cells show the number of admissions in 2015 for each combination of severity classes.

Severity 2015 (new conversion table)	Severity 2016		
	[0,0.01) - [0.4,1]	Other	Total
[0,0.01) - [0.4,1]	1 802 744	765	1 803 509
Other	11 559	3 366	14 925
Missing	125	287	412
Total	1 814 428	4 418	1 818 846

5.4 Summary and conclusions

In 2012 and 2013 hospitals transitioned from using ICD9 to code the main diagnosis of admissions to ICD10. One of the variables in the calculation in Hospital Standardised Mortality Ratio is the severity of the main diagnosis. This severity is determined using historical mortality. Previously this was done for ICD9 diagnoses, using admissions coded in ICD9 only. However, the data used in estimating the mortality is now partially coded in ICD9 and partially in ICD10, and the admissions for which the severity has to be determined are all coded in ICD10. Therefore, it was decided to determine the severity for ICD10 codes, with a new method that can make use of a mixed ICD10/ICD9 historical dataset. The main result is that the new method brings about a gradual transition over time from severities based on ICD9 data to severities based solely on ICD10 data. Furthermore there are substantially fewer admissions without a severity code in the new severity classification, which is mainly caused by using an updated ICD10 to ICD9 conversion table.

6. References

Bottle, A., Jarman B. and P. Aylin (2011). Hospital standardized mortality ratios: sensitivity analyses on the impact of coding. Health Serv Res 46(6 Pt 1):1741–61.

Campbell, M.J., R.M. Jacques, J. Fotheringham, T. Pearson, R. Maheswaran and J. Nicholl (2011). An evaluation of the Summary Hospital Mortality Index, final report. School of Health and Related Research, The University of Sheffield.

CBS (2011, 2012, 2013,… etc.). HSMR 2010 [2011, 2012, …etc.], Methodological report. Statistics Netherlands, The Hague/Heerlen.
https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/overzicht/methodologische-rapportages-bij-hospital-standardised-mortality-ratios--hsmr--

Chong, C.A.K.Y., Nguyen G.C. and M.E. Wilcox (2012). Trends in Canadian hospital standardised mortality ratios and palliative care coding 2004-2010: a retrospective database analyses.. BMJ Open 2012;2:e001729. doi:10.1136/bmjopen-2012-001729.

Heijink, R., X. Koolman, D. Pieter, A. van der Veen, B. Jarman and G. Westert (2008). Measuring and explaining mortality in Dutch hospitals; The Hospital Standardized Mortality Rate between 2003 and 2005. BMC Health Services Research 8:73.

HSCIC (2016). Indicator Specification: Summary Hospital-level Mortality Indicator. Version 1.22, 24 February 2016. Health & Social Care Information Centre.
www.content.digital.nhs.uk/shmi

Jarman, B., D. Pieter, A.A. van der Veen, R.B. Kool, P. Aylin, A. Bottle and G.P. Westert (2010). The hospital standardised mortality ratio: a powerful tool for Dutch hospitals to assess their quality of care? Qual Saf Healthcare 19, 9-13.

Jarman, B., S. Gault, B. Alves, A. Hider, S. Dolan, A. Cook, B. Hurwitz and L.I. Iezzoni (1999). Explaining differences in English hospital death rates using routinely collected data, BMJ 318, 1515-1519.

Lamarche-Vadel, A., M. Ngantcha, M.-A. Le Pogamm, C. Grenier, L. Meyer and G. Rey (2015). Hospital comparisons based on mortality: revisiting the choice of postadmission timeframe and evaluating the contribution of cause-of-death data, France, 2009. Medical Care 53, 736-742.

Lilford, R. and P. Pronovost (2010). Using hospital mortality rates to judge hospital performance: a bad idea that just won't go away. BMJ 340:c2016.

Mitchell, T.M. (2016). "Estimating Probabilities" additional chapter for the book Mitchell, T.M. (1997). Machine Learning. McGraw-Hill, New York. Online at http://www.cs.cmu.edu/~tom.

NHS - The Health and Social Care Information Centre (2013). The use of palliative care coding in the Summary Hospital-level Mortality Indicator.http://www.hscic.gov.uk/media/11150/Palliative-Care-Coding-report/pdf/Palliative_Care_Coding_Report.pdf

Ploemacher, J, A.Z. Israëls, D.J. van der Laan and A. de Bruin (2013). Gestandaardiseerde ziekenhuissterfte daalt in de tijd. Ned Tijdschr Geneeskd. 157 (22), 1034-1039.

Pouw, M.E., Peelen, L.M., Moons, K.G.M., Kalkman, C.J. and H.F. Lingsma (2013). Including post-discharge mortality in the calculation of hospital standardised mortality ratios: a retrospective analysis of hospital episode statistics. BMJ, in press.

Prismant (2008). De toepasbaarheid van de HSMR in het toezicht van de inspectie voor de gezondheidszorg. Prismant, Utrecht.

Quan H, V. Sundararajan, P. Halfon, A. Fong A, B. Burnand, J.C. Luthi, L.D. Saunders, C.A. Beck, T. E. Feasby, W.A. Ghali (2005). Coding algorithms for defining comorbidities in ICD-9-CM and ICD10 administrative data. Medical Care 43, 1130-1139.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Van den Bosch, W.F., J. Silberbusch, K.J. Roozendaal and C. Wagner (2010). Variatie in codering patiëntengegevens beïnvloedt gestandaardiseerd ziekenhuissterftecijfer. Ned Tijdschr Geneeskd. 154 (2), 1-9.

Van den Bosch, W.F., P. Spreeuwenberg and C. Wagner (2011). Gestandaardiseerd ziekenhuissterftecijfer (HSMR): correctie voor ernst hoofddiagnose kan beter. Ned Tijdschr Geneeskd. 155:A3299, 66-75.

Van der Laan, D.J. (2013). Quality of the Dutch Medical Registration (LMR) for the calculation of the Hospital Standardised Mortality Ratio. Statistics Netherlands, The Hague/Heerlen. https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/aanvullende%20onderzoeksbeschrijvingen/quality-of-the-dutch-medical-registration--lmr---for-the-calculation-of-the-hospital-standardised-mortality-ratio--hsmr--

Van der Laan, D.J., A. de Bruin, J. van den Akker-Ploemacher and Frank Pijpers (2015). Post-discharge mortality in the Hospital Standardized Mortality Ratio. Discussion Papers, Statistics Netherlands, The Hague/Heerlen. https://www.cbs.nl/nl-nl/achtergrond/2015/23/operationalisatie-sterfte-na-ontslag-in-de-hsmr

Van der Laan, D.J. and A. de Bruin (2016). Uitbreiding hoofddiagnosegroepen in de HSMR. Discussion Paper, Statistics Netherlands, The Hague/Heerlen. https://www.cbs.nl/nl-nl/achtergrond/2016/41/uitbreiding-hoofddiagnosegroepen-in-de-hsmr

Van Gestel, Y.R.B.M., V.E.P.P. Lemmens, H.F. Lingsma, I.H.J.T. de Hingh, H.J.T Rutten.and J.W.W Coebergh (2012). The hospital standardized mortality ratio fallacy; a narrative review. Medical Care 50 (8), 662-667.

Appendix. Statistical significance of covariates, HSMR 2016 model

Statistical significance (95% confidence) of the covariates for the 157 logistic regressions (1=significant; 0=non-significant; "-"=variable dropped because all categories are collapsed, due to < 50 admissions or no deaths in all but one category)

No. diagnosis group	Severity main diagnosis	Sex	Age	Urgency	Comorbidity_1	Comorbidity_2	Comorbidity_3	Comorbidity_4	Comorbidity_5	Comorbidity_6	Comorbidity_7	Comorbidity_8	Comorbidity_9	Comorbidity_10	Comorbidity_11	Comorbidity_12	Comorbidity_13	Comorbidity_14	Comorbidity_15	Comorbidity_16	Comorbidity_17	SES	Month admission	Year discharge	Source admission
1	0	0	1	1	-	-	-	-	-	0	-	-	0	0	-	-	1	-	-	-	-	0	0	0	0
2	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	1	1	-	1	1	1	1	1	1
3	1	0	1	0	0	1	1	0	0	0	0	-	1	0	0	0	1	0	-	1	1	1	0	0	0
4	1	1	1	1	1	1	0	-	-	0	0	-	1	0	-	-	0	1	-	1	-	0	0	0	1
5	1	0	0	0	-	-	-	-	-	0	-	-	0	0	-	-	0	0	1	-	-	0	0	0	1
6	1	0	1	1	0	1	0	0	0	1	0	-	1	0	0	0	1	1	-	0	1	0	0	0	1
7	1	0	1	1	0	0	1	1	0	0	0	-	1	0	1	-	0	1	-	1	-	0	0	0	0
8	0	0	0	1	0	1	1	0	-	0	0	-	1	0	0	-	1	0	-	1	-	0	0	0	1
9	0	0	0	1	0	1	1	0	0	1	0	1	0	0	0	-	1	0	-	1	-	0	0	1	0
10	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	-	1	1	0	0	1	1
11	-	1	1	1	1	1	1	1	0	1	0	-	1	0	0	0	1	0	-	1	-	0	0	0	0
12	0	1	1	1	1	1	1	1	-	0	-	-	0	0	0	-	1	0	-	1	1	0	0	0	0
13	0	0	1	1	0	1	1	1	0	1	0	-	1	0	0	-	1	1	-	1	1	0	0	1	1
14	1	0	1	1	0	1	1	0	-	0	0	-	1	0	1	-	1	0	-	1	-	1	0	0	0
15	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	0	-	1	1	0	0	0	1

16	-	0	0	1	1	-	0	-	-	1	-	-	-	0	-	-	0	0	-	1	-	0	0	0	1
17	1	0	1	1	0	-	0	-	-	0	-	-	-	1	-	-	0	0	-	1	-	0	0	0	1
18	0	0	0	1	0	1	0	0	-	0	-	-	-	0	-	-	0	0	-	1	-	0	0	0	0
19	1	0	1	1	0	1	1	0	0	0	1	-	1	0	0	0	1	0	-	1	-	0	0	0	0
20	0	-	0	1	0	0	-	-	-	0	-	-	-	1	-	-	1	0	-	1	-	0	0	0	0
21	0	-	1	1	0	-	0	0	-	1	0	-	-	1	-	-	1	0	-	1	-	0	0	1	1
22	-	-	1	1	0	0	1	0	-	1	-	-	1	1	-	-	1	0	-	1	-	0	0	0	0
23	-	-	1	1	0	1	1	0	0	0	0	-	1	0	0	1	1	1	-	1	-	0	0	0	0
24	0	-	0	1	0	-	-	-	-	0	-	-	-	0	-	-	-	-	-	1	-	1	0	0	0
25	1	0	1	1	0	1	0	1	0	1	0	-	1	0	0	-	1	1	-	1	-	0	1	0	0
26	1	0	1	1	0	1	0	1	-	0	0	-	1	0	0	-	1	0	-	1	-	0	1	0	0
27	0	0	1	1	0	-	0	1	-	0	0	-	-	1	-	0	0	0	-	1	-	0	0	1	1
28	-	1	1	1	0	-	-	-	-	0	-	-	-	0	-	-	-	0	-	1	-	0	1	0	-
29	0	0	1	1	-	-	-	-	-	0	-	-	-	0	-	-	-	-	-	-	-	0	0	0	-
30	1	0	1	1	1	1	1	1	0	1	0	0	1	0	0	1	1	1	0	1	1	0	0	0	1
31	1	1	1	1	0	1	0	1	1	0	0	-	1	0	0	0	1	1	-	0	1	1	0	1	1
32	0	0	1	1	0	1	0	1	-	1	0	-	1	0	1	0	1	1	-	0	-	0	0	0	0
33	1	0	1	1	1	0	0	0	-	0	-	-	-	0	-	-	1	0	-	1	-	0	0	0	1
34	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	-	1	1	1	1	1	1
35	0	1	0	1	0	1	1	1	-	1	1	-	0	0	-	-	1	1	-	1	-	0	0	0	1
36	1	0	1	1	0	1	1	0	0	1	0	-	1	0	0	0	1	0	-	0	-	0	0	0	0
37	1	1	1	1	0	1	0	1	0	1	0	-	1	0	0	0	0	0	-	1	1	0	0	0	1
38	1	0	1	1	0	1	1	0	0	0	0	-	1	0	0	0	1	0	-	0	-	0	1	0	1
39	1	0	1	0	0	1	1	1	1	0	0	-	0	0	0	-	0	1	-	0	0	1	0	0	0
40	1	0	1	1	0	1	1	1	0	0	0	-	1	0	0	0	1	1	-	1	-	0	1	0	1
41	1	1	1	1	0	1	0	0	0	0	0	-	0	1	0	0	1	1	-	1	1	0	0	0	1

42	1	1	1	1	0	1	1	1	1	1	0	-	1	0	0	0	0	1	-	1	1	0	1	0	0
43	0	0	0	1	-	-	-	-	-	0	-	-	1	0	-	-	0	-	-	-	-	0	0	1	1
44	1	0	1	1	0	1	1	1	0	1	0	-	0	1	0	-	0	1	-	0	-	0	1	0	1
45	1	0	1	1	0	1	1	0	1	1	0	0	1	0	0	0	1	0	-	1	0	0	0	0	1
46	-	0	1	0	1	0	0	-	-	1	0	-	1	0	-	-	1	0	-	1	-	0	0	0	0
47	1	1	1	0	1	1	0	0	1	0	-	-	1	0	-	-	0	1	-	0	1	0	0	0	0
48	1	1	1	1	1	1	1	0	0	1	1	-	1	0	0	0	1	1	-	1	-	0	0	0	0
49	1	0	1	0	1	-	-	0	-	0	-	-	1	1	-	-	-	0	-	-	-	0	0	0	0
50	0	0	1	1	0	1	0	0	0	1	0	-	1	0	1	-	1	0	-	0	-	0	0	0	0
51	1	0	1	1	0	1	0	1	1	1	0	-	1	0	0	0	1	1	-	1	-	0	0	0	1
52	0	0	1	1	0	1	0	0	1	0	-	-	-	0	0	-	0	0	-	-	-	0	0	0	0
53	1	0	1	1	0	0	0	0	0	0	1	-	0	1	1	0	0	0	-	0	-	0	0	1	1
54	1	0	1	1	0	1	0	0	0	0	-	-	-	0	0	1	0	0	-	0	-	0	0	0	1
55	1	0	1	0	1	1	1	1	1	1	0	-	1	1	0	0	1	1	-	1	1	0	0	0	1
56	1	0	1	1	0	1	1	0	0	1	-	-	1	0	0	0	1	0	-	1	-	0	0	0	1
57	0	0	1	0	0	1	0	0	0	0	0	-	0	0	0	0	0	1	-	0	-	0	0	0	0
58	1	0	1	1	1	1	1	1	1	0	0	-	0	0	0	0	1	1	-	1	1	0	1	0	1
59	1	0	1	1	1	1	1	1	1	1	0	-	1	0	1	0	1	1	-	0	-	1	0	1	1
60	1	1	1	1	1	1	1	1	-	0	0	-	1	0	0	0	1	0	-	1	1	0	0	1	1
61	1	0	1	1	0	1	0	1	0	0	0	-	0	0	1	-	0	0	-	0	-	0	1	0	1
62	1	0	1	1	0	1	1	1	1	1	0	0	1	1	1	0	1	1	-	1	1	0	1	1	1
63	1	0	1	1	1	1	1	1	0	1	0	1	1	1	0	0	1	1	-	1	1	0	1	1	1
64	-	1	1	0	0	0	0	1	0	0	0	-	1	0	0	-	1	1	-	1	-	1	0	1	-
65	1	1	1	1	1	1	1	1	1	1	0	-	1	1	1	1	1	1	-	1	1	0	1	0	1
66	1	0	1	1	1	1	0	0	-	0	-	-	-	0	0	-	1	1	-	-	-	0	0	1	1
67	1	0	1	1	1	1	0	1	1	1	0	-	1	1	0	-	1	1	-	0	-	1	0	0	0

68	1	0	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	-	1	1	1	1	1	1
69	1	1	1	1	1	1	1	1	1	1	1	-	1	1	1	0	1	1	-	1	1	0	1	1	1
70	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	0	1	1	1
71	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-	1	1	0	1	1	1
72	1	0	1	1	1	1	0	1	0	1	0	-	1	0	1	0	1	1	-	1	-	0	0	1	0
73	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1	-	1	1	0	0	0	1
74	1	1	1	1	1	1	1	1	1	1	0	-	1	0	1	0	1	0	-	1	-	1	0	1	0
75	1	0	1	1	1	1	1	1	1	1	1	-	1	0	0	0	1	1	-	1	-	0	0	0	1
76	1	0	1	1	1	1	1	1	0	1	1	-	1	0	0	0	1	1	-	1	1	0	1	1	1
77	1	1	1	1	0	1	1	0	0	1	0	-	1	0	0	0	1	1	-	1	0	0	0	0	0
78	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1
79	1	0	1	0	1	1	1	0	0	1	0	-	1	0	1	0	1	1	-	0	-	0	1	1	1
80	1	0	1	1	0	1	0	0	0	1	0	-	0	0	-	0	1	0	-	0	-	0	0	0	1
81	0	1	1	0	0	1	0	1	0	0	0	-	0	0	0	0	1	0	-	0	-	0	0	0	1
82	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	1
83	-	0	1	1	0	1	0	0	0	1	0	-	0	0	0	-	1	1	-	0	-	0	0	1	1
84	-	1	1	0	1	1	1	0	1	0	0	-	1	0	0	0	1	1	-	1	-	0	0	1	0
85	1	0	1	1	1	1	0	0	1	1	0	-	1	1	0	1	1	1	-	1	1	0	0	0	1
86	1	1	1	1	1	0	1	0	1	0	1	-	1	0	0	0	1	1	-	1	-	0	0	0	0
87	1	0	1	1	0	1	0	-	-	0	0	-	-	0	-	-	0	1	-	0	-	0	1	0	1
88	1	1	1	1	1	1	1	1	0	0	0	-	1	0	0	0	1	1	-	1	1	0	0	0	1
89	1	1	1	1	0	1	0	1	1	1	0	-	1	0	1	0	0	1	-	1	0	0	1	1	1
90	1	0	1	0	0	1	1	1	0	1	0	-	1	0	0	0	1	1	-	1	1	0	0	0	1
91	0	0	1	1	1	1	0	0	0	0	0	-	1	0	-	-	1	1	-	1	-	1	1	0	0
92	1	0	1	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	-	1	0	0	0	0	0
93	1	0	1	0	0	1	1	0	0	1	-	0	1	0	-	-	1	1	-	0	-	0	0	0	0

94	1	0	1	1	0	1	1	0	0	0	0	0	0	0	0	-	0	1	-	0	0	1	0	0	0
95	-	1	1	0	0	1	0	0	0	1	1	-	1	1	-	-	0	0	-	0	-	0	0	1	1
96	1	1	1	1	1	1	0	0	-	1	-	-	1	0	0	-	1	0	-	1	1	0	0	0	1
97	1	1	1	1	1	1	1	1	0	1	0	0	0	1	0	0	1	0	-	1	1	0	0	0	0
98	-	0	1	0	0	1	0	0	-	0	0	-	1	0	0	-	0	0	-	0	-	0	0	0	0
99	1	1	1	1	0	1	1	1	1	1	0	-	1	1	1	0	1	1	-	1	1	0	0	1	1
100	1	0	1	1	0	1	1	1	1	1	1	-	1	0	0	0	1	1	-	1	-	0	0	0	1
101	0	0	1	0	1	1	0	0	0	0	0	-	1	0	0	-	1	0	-	1	-	0	0	1	0
102	1	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0	1	1	-	1	1	0	0	0	1
103	1	1	1	1	0	1	1	0	-	0	-	0	1	1	0	-	1	0	-	-	1	0	0	0	1
104	1	0	1	1	0	1	1	0	0	0	0	0	0	1	1	-	1	1	-	1	1	0	0	0	1
105	1	0	1	1	1	1	1	0	1	1	1	1	1	0	0	0	1	1	-	1	1	0	0	0	1
106	1	0	1	1	1	1	1	1	1	0	0	0	1	0	0	0	1	1	-	1	1	0	0	1	1
107	0	1	1	1	0	1	1	0	1	1	0	-	0	0	0	0	1	1	-	1	0	1	0	0	0
108	1	1	1	1	0	1	1	0	1	0	0	1	1	1	1	0	1	1	-	1	0	0	0	0	1
109	1	0	1	1	0	1	1	1	0	0	1	-	0	1	0	-	1	0	-	1	-	0	0	0	0
110	1	0	1	1	1	1	1	1	1	1	0	1	1	0	0	0	0	1	-	1	1	0	0	0	1
111	1	0	1	1	0	1	1	0	1	1	0	-	1	0	1	0	0	1	-	1	-	1	0	0	1
112	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	-	1	1	0	1	1	1
113	1	1	1	1	1	1	0	1	0	1	0	-	1	1	0	1	1	1	-	1	-	0	0	0	0
114	1	0	1	1	0	1	0	0	1	1	0	-	0	0	0	-	1	1	-	0	-	0	0	0	1
115	1	-	1	1	0	1	0	0	0	1	0	-	1	0	0	0	0	1	-	1	-	0	0	0	0
116	-	-	0	-	-	-	-	-	-	-	-	-	-	1	-	-	-	0	-	0	-	0	0	0	-
117	1	0	1	1	0	1	1	-	0	0	1	-	-	1	0	-	1	1	-	1	-	0	0	0	0
118	1	0	1	1	-	1	1	1	-	0	-	-	1	-	-	-	0	0	-	-	-	0	0	0	0
119	1	0	1	1	1	1	1	1	0	1	0	-	1	1	0	0	1	1	-	1	1	0	1	0	1

120	1	0	1	1	1	1	1	0	0	0	0	-	1	1	0	0	1	0	-	0	-	0	0	0	1
121	1	0	1	1	0	1	1	0	0	0	0	-	1	0	0	0	1	0	-	0	-	0	0	0	1
122	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	-	0	1	0	0	0	1
123	0	0	1	1	1	1	0	1	0	0	0	-	1	0	0	0	1	0	-	1	-	0	0	0	-
124	1	0	1	1	0	1	0	1	1	0	0	-	0	1	1	1	0	0	-	1	-	0	0	0	1
125	0	1	1	0	0	1	0	0	0	1	0	-	0	0	0	0	1	0	-	0	-	0	0	0	1
126	1	0	1	1	1	1	1	1	0	0	1	-	1	0	0	0	1	1	-	0	-	0	0	0	0
127	1	0	1	1	0	1	0	1	-	0	-	-	-	0	-	-	1	-	-	-	-	0	0	0	0
128	1	0	1	0	0	1	1	0	-	0	-	-	1	0	-	1	0	-	-	-	-	0	0	0	0
129	1	1	-	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	1	1
130	1	0	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	1
131	1	0	0	0	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	0	0	1
132	0	1	1	1	0	1	1	0	0	0	-	-	-	0	0	-	0	-	-	-	-	0	0	0	0
133	-	1	1	0	1	1	1	1	1	1	1	-	1	1	1	0	1	1	-	1	1	0	1	0	0
134	1	1	1	1	0	0	0	0	1	0	-	-	0	0	-	0	1	1	-	-	-	0	0	1	0
135	0	1	1	1	0	1	1	1	0	1	1	-	1	0	0	0	1	1	-	0	-	0	0	0	0
136	1	1	1	1	1	1	1	1	0	1	0	-	1	1	0	0	1	1	-	0	-	0	0	0	0
137	1	1	1	1	0	1	1	0	1	1	0	-	1	1	0	1	1	1	-	1	-	0	0	0	1
138	1	1	1	1	1	1	1	1	1	1	0	-	1	0	0	1	1	1	-	1	-	0	0	0	1
139	1	0	1	1	1	1	1	0	0	1	0	-	1	1	1	1	1	1	-	0	-	0	0	0	0
140	1	0	1	0	1	1	0	0	0	0	-	-	1	0	-	-	1	1	-	-	-	0	1	0	0
141	1	0	1	1	0	1	0	-	0	1	0	-	-	0	0	-	0	-	-	-	-	0	0	0	-
142	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	1
143	1	0	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	0	1	1
144	1	1	1	0	0	1	1	1	1	1	0	-	1	0	1	0	1	1	-	1	-	0	1	0	0
145	1	0	1	1	-	-	1	-	-	0	-	-	-	0	-	-	-	-	-	-	-	0	0	1	0

146	1	0	1	0	0	1	0	0	0	1	0	-	1	0	0	0	0	0	-	0	1	0	0	0	1
147	1	1	1	0	0	1	0	1	1	1	0	-	1	1	0	0	0	1	-	0	-	1	0	0	0
148	-	0	1	0	0	1	1	0	0	0	0	-	0	0	0	1	1	0	-	1	-	0	0	0	0
149	0	0	1	0	1	1	1	1	0	0	0	-	1	0	0	0	1	0	-	1	0	0	0	0	0
150	1	0	1	1	0	1	1	0	0	1	0	-	-	0	0	-	1	0	-	0	-	0	0	0	0
151	-	1	1	0	1	1	1	-	-	0	-	1	1	0	-	-	0	0	-	1	1	0	0	1	0
152	-	0	1	1	0	1	0	0	0	0	0	-	0	0	0	-	1	0	-	1	0	0	0	0	0
153	-	0	1	1	0	1	1	0	1	0	0	0	0	0	0	0	1	1	-	1	1	0	0	1	0
154	-	0	1	1	0	1	0	0	1	1	0	-	1	0	0	1	1	1	-	0	0	0	0	1	0
155	1	0	1	0	0	1	0	-	-	1	0	-	0	0	-	-	1	0	-	0	-	0	0	0	0
156	1	0	1	1	1	1	1	1	0	1	0	0	0	0	0	1	1	0	-	1	0	0	0	1	1
157	1	1	1	1	1	1	1	1	0	0	0	-	0	0	0	0	1	1	-	1	0	0	0	1	1
total	112	51	144	125	63	133	91	73	50	82	27	14	103	44	32	24	113	86	1	99	49	18	30	43	85

The numbers of the comorbidity groups in the header of the table above are the following comorbidities:

Comorbidity_1 - Acute myocardial infarction
Comorbidity_2 - Congestive heart failure
Comorbidity_3 - Peripheral vascular disease
Comorbidity_4 - Cerebral vascular accident
Comorbidity_5 - Dementia
Comorbidity_6 - Pulmonary disease
Comorbidity_7 - Connective tissue disorder
Comorbidity_8 - Peptic ulcer
Comorbidity_9 - Liver disease / Severe liver disease

Comorbidity_9 - Liver disease / Severe liver disease
Comorbidity_10 - Diabetes / Diabetes complications
Comorbidity_11 - Diabetes complications
Comorbidity_12 - Paraplegia
Comorbidity_13 - Renal disease
Comorbidity_14 - Cancer
Comorbidity_15 - HIV
Comorbidity_16 - Metastatic cancer
Comorbidity_17 - Severe liver disease

The names of the diagnosis groups are described in Table 4.4.1, and the corresponding CCS groups and ICD10 codes are given in the file 'Classification of variables', published together with this report.

Explanation of figures

Empty cell	Figure not applicable
.	Figure is unknown, insufficiently reliable or confidential
*	Provisional figure
**	Revised provisional figure
2016–2017	2016 to 2017 inclusive
2016/2017	Average for 2016 to 2017 inclusive
2016/'17	Crop year, financial year, school year, etc., beginning in 2016 and ending in 2017
2014/'15–2016/'17	Crop year, financial year, etc., 2014/'15 to 2016/'17 inclusive

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

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress
CCN Creatie, The Hague

Design
Edenspiekermann

Information
Telephone +31 88 570 7070
Via contact form: www.cbs.nl/information