Methodological paper

HSMR 2021:

Methodological report

Jan van der Laan
Daisy Kolk
Corine Witvliet-Penning
Agnes de Bruin

2022

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 numbers of deaths and aim to present in-hospital mortality figures in comparison to the national average. 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. In chapter 5 the development of an SMR model for COVID-19 admissions is described, and in chapter 6 analyses are presented, investigating the possible impact of the COVID-19 pandemic on the (non-COVID-19) HSMR outcomes of 2021. Conclusions are given in chapter 7.

1.1 What is the (H)SMR?

In-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. Therefore it is more appropriate to adjust (i.e. standardise) mortality rates for differences in patient characteristics ("case mix") across hospitals as much as possible. 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 determined using a model based on data from all national hospitals, in which the in-hospital mortality is explained by patient characteristics, such as age and diagnosis, and characteristics of the admission, such as whether the admission is acute or not. 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 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 SMRd across all 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 by chance only, confidence intervals are calculated for the SMRs and HSMRs to inform

hospitals whether they have a (statistically) significantly high or low adjusted mortality rate compared with the average value 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. Various quality indicators are available, 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 patient characteristics, such as age, sex and comorbidity. Unfortunately, recovery is hard to measure and mostly occurs after patients have been discharged from the hospital. Although in-hospital mortality is a much more limited quality indicator, it can be measured accurately and is therefore used as a quality indicator in several countries, using the HSMR and SMRs as defined in section 1.1. If these instruments were totally valid, i.e. the calculations would adjust perfectly for everything that cannot be influenced by the hospital, a value above 100 would always indicate inferior quality of care, 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 unforeseeable complications was equally distributed across hospitals, after adjustment for differences in case mix. However, it is impossible to construct such a perfect instrument to measure the quality of health care; the outcome of the indicator will to some extent always be partially influenced by differences between hospitals with regard to case mix, availability of highly specialized treatment options, etc. A significantly high (H)SMR will at most be an indication of possible shortcomings in hospital care, but a high value may also be caused by coding errors in the data or a lack in the model of essential covariates related to mortality. Still, a significantly high (H)SMR is often seen as a warning sign and a reason for further investigation by the hospital.

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. 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, described in Jarman et al. (2010), was slightly adapted by Prismant (Prismant, 2008) up to reporting year 2009. In 2010 DHD (Dutch Hospital Data, Utrecht), the registry 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 is 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 https://opendata.cbs.nl/statline/#/CBS/nl/).

The starting point for CBS was the HSMR method previously used by Prismant. As a result of progressive insight, over the years CBS has introduced changes in the model for the HSMR, which are described in the annual 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 explicitly agree. For this reason, CBS needs written consent 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 annually estimates the models for expected mortality per diagnosis for the most recent three-year period. It calculates the HSMRs and SMRs for all hospitals that (1) had authorised CBS, (2) had registered all of its admissions in the LBZ in the relevant period, and (3) were not excluded on the grounds of criteria for data 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. Individual hospital reports, containing their HSMR and diagnosis-specific SMR figures for the most recent reporting year and the three-year period. SMRs are also presented for different patient groups (by age, sex and urgency of admission) and for clusters of diagnoses. Hospitals can compare their outcome with the national average: overall, and per diagnosis and patient group.
2. A dataset for each hospital with the mortality probabilities for all its individual admissions. Besides the mortality probability, each admission record contains the observed mortality (0 or 1) and the scores on the covariates of the HSMR model. Hospitals 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 quality of hospital care. 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), which is always a simplification of reality.

Since the very first publication on the HSMR in the United Kingdom, there has been an on-going debate about the value 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 of the quality of health care, alongside other possible indicators. But even if HSMRs were to be used for a more limited purpose, i.e. standardising in-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 are

registered uniformly, but the registration of whether an admission was acute or not, the main discharge diagnosis or comorbidity 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 research (Van der Laan, 2013) shows that comorbidities in particular present a problem in the Netherlands, as there is large variation in coding this covariate (see also section 4.3). Van den Bosch *et al.* (2010) refer extensively to the influence of coding errors, also Van Erven *et al.* (2018) report underreporting of comorbidities. Nationwide, the registration of comorbidities in Dutch hospitals has increased strongly up to 2014. From 2015 onwards the yearly increase is smaller, and in 2021 the registration of comorbidities seemed to stabilize. But there are still hospitals showing annual 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 sufficiently register whether a comorbidity was a complication or not. As complications are excluded from the HSMR comorbidity covariates, underreporting complications might falsely lead to a higher comorbidity rate, thus influencing the HSMR outcomes. To stimulate correct coding of complications, an indicator has been added to the hospital HSMR reports showing the percentage of registered complications of the hospital, and the overall average. The introduction of this indicator has led to less underreporting of complications, though there are still differences in the number of complications registered by hospitals.

- On average, some hospitals may treat more seriously ill patients than others, even if those patients have the same set of scores on the covariates. University hospitals may, for example, have more complicated cases than other hospitals, while regional hospitals are generally more involved in end of life care. It is questionable whether the model sufficiently adjusts for this phenomenon, since some essential information regarding severity or complexity of disease is then missing. Some of the desired covariates are not registered in the LBZ and some will actually even be hard to measure at all in this type of registry with routinely collected hospital discharge data.

- The same problem occurs when certain high-risk surgical procedures are only performed in a selection of 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. The downside however, of using a treatment method as a covariate, is that ideally it should not be part of the model as it is a component of hospital care.

- Hospital admission and discharge policies may differ between hospitals. For instance, one hospital may admit the same patient more frequently but for shorter stays than another. Or it may discharge a patient earlier due to higher availability of external terminal care facilities in the neighbourhood. Also, patients may be referred from one hospital to another for further treatment. Obviously, all these situations that mostly are unrelated to quality of care also influence the outcome of the HSMR, as they influence the observed mortality numbers.

– Hospitals can compare their HSMR and SMRs with the national average value of 100. The comparison of (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)SMRs is still based on the model for all hospitals (without correcting for the type of hospital), but peer group comparison allows a specialised hospital to compare its results with the average for similar hospitals. For instance, the average HSMR of university hospitals is >100 in the Netherlands due to insufficient case mix correction, but comparing their results with a peer group average allows these hospitals (and for specific diagnoses also other specialised hospitals) to better interpret their own scores.

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) observed a decrease in standardised in-hospital mortality in the Netherlands in 2005-2010, which may have been caused by an overall improvement in quality of care, 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. Pouw *et al*. (2013) performed a retrospective analysis on Dutch hospital data linked to mortality data, and concluded that including early post-discharge mortality diminishes 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). Including all mortality within a 45-day period after admission was advised to reduce the influence of hospital discharge policies on the HSMR. In addition, a French study also recommended fixed post-admission periods of more than 30 days (Lamarche-Vadel *et al*., 2015).

However, including post-discharge mortality in the indicator will not reduce the effect of differences in admission policies for terminally ill patients. Some hospitals may admit more terminally ill patients to provide terminal palliative care than other hospitals and those admissions may distort HSMR outcomes. Palliative care in general can be measured in ICD-10 (code Z51.5), but this variable should be used with caution, as differences between hospitals in coding practices have been shown in e.g. the UK and Canada, and adjusting the HSMR for palliative care may increase the risk of gaming (NHS, 2013; Chong *et al.*, 2012; Bottle *et al.*, 2011). Because of this, and because ICD-10 code Z51.5 does not distinguish between early and terminal palliative care, palliative care admissions have not yet been excluded from the calculation of the HSMR in the Netherlands. However, the hospital HSMR reports 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.

Despite of the above-mentioned limitations and the ongoing debate on the validity and reliability of mortality-based indicators like the HSMR, other studies suggest that mortality monitoring can be indicative of failings in quality of care. Cecil *et al.* (2020) found that in an English hospital setting mortality alerts, based on higher than expected mortality in 122 diagnosis and procedure groups, were associated with structural indicators of lower quality of care (e.g. lower nurse-to-bed ratio, overcrowding and financial pressures) and outcome indicators like lower patient and trainee satisfaction. They conclude that a mortality alerting system might be valuable in highlighting poor quality of care.

2. Method changes

This chapter summarizes the changes in the HSMR method (HSMR 2021) compared to the method used last year (HSMR 2020). For previous changes see the respective methodological reports (CBS, 2011, …, 2021). Overall, the method has remained the same. Only a few changes related to selection of the data and related to the covariates have been implemented. These changes are described below.

2.1 Exclusion of COVID-19 admissions

Both in 2020 and in 2021 the COVID-19 pandemic has significantly impacted hospital care. For the HSMR 2020, in agreement with the Dutch Healthcare Authority (NZa) and the Health and Youth Care Inspectorate (IGJ), it was decided to omit the 2020 admissions with COVID-19 as main diagnosis (ICD-10 codes U07.1 (COVID-19, virus identified (lab confirmed)) and U07.2 (COVID-19, virus not identified (clinically diagnosed))) from the HSMR model, due to the different nature of these admissions and since specific expertise for treating COVID-19 was lacking during the first months of the pandemic. For the HSMR 2021 these admissions were excluded again, while admissions with the new ICD-10 code U10.9 (Multisystem inflammatory syndrome associated with COVID-19, unspecified) were additionally excluded. Admissions with COVID-19 as a comorbidity only, have not been excluded. The admissions with COVID-19 as main diagnosis were also excluded from the dataset that was used to evaluate data quality and case mix (see section 3.5). Apart from ICD-10 code U10.9, two additional ICD-10 codes have been introduced in 2021 that are related to a previous episode of COVID-19. Admissions with these two codes as a main diagnoses (U08.9, personal history of COVID-19, unspecified; U09.9, post COVID-19 condition, unspecified) have been included in the HSMR model.

Thus, the 2021 HSMR values are based on regular (non-COVID-19 related) care and the included admissions in 2020 and 2021 have largely similar main diagnoses compared to earlier years. Due to downscaling of regular hospital care as a result of the COVID-19 pandemic, and to a smaller extent also due to the exclusion of the COVID-19 admissions, the number of admissions of 2021 included in the model, was on average 13% lower than the number of 2019 admissions, while in 2020 it was 14% lower compared to 2019.

Although admissions with COVID-19 as main diagnosis have been excluded from the HSMR model, CBS has repeated some of the analyses done last year to determine whether and to what extent the COVID-19 pandemic might have influenced the hospital care not related to COVID-19 in 2021. The outcome of these analyses is presented in Chapter 6.

This year CBS also developed a separate SMR model for COVID-19 admissions of 2021. This COVID-19 2021 model is described in Chapter 5.

2.2 Updated socioeconomic status (SES) scores

Since 2010 the HSMR model has included a variable that measures the socioeconomic status of patients based on the four-digit postal code of the patient's home address (CBS, 2011, …, 2021). This SES score was derived from the Netherlands Institute for Social Research (SCP), which had

collected SES data on income, employment and educational level per four-digit postal area on which they performed principal component analyses to get a SES score for each four-digit postal area. Population-weighted quintiles were calculated from these scores, resulting in six SES categories that are used in the HSMR model. Postal codes with less than 100 residential households did not have a component score and were assigned to the sixth SES category 'unknown'.

However, SCP no longer calculates these SES scores and the latest version is from 2017. In its place CBS has derived a new score, called the SES-WOA score (Arts *et al*., 2022), using 2019 data. This score is based on household data concerning welfare (a combination of income and wealth), level of education and recent labour participation. The scores are determined at the household level and then averaged to scores per four-digit postal code. The final SES categories were based on household-weighted quintiles of these averaged scores per postal code. By weighting with the number of households the average SES-WOA scores of the postal codes in the middle SES category (the third quintile) includes zero (the average SES-score over all households in the Netherlands is zero). Similar to the previous SCP scores, postal codes with less than 100 residential households were assigned to the SES category 'unknown'.

The new SES variable based on 2019 data has been used in the present HSMR model for the LBZ 2021 data. The new SES variable, which will be regularly updated by CBS, will also be used for later years of LBZ data in future HSMR models.

Since the method for determining the SES scores has changed, the scores will not be fully comparable to the scores used for previous LBZ years. Table 2.2.1 compares the new SES-WOA score based on 2019 data (calculated by CBS) to the SES score based on 2017 data (calculated by SCP). We found that the SES categories of these scores show a strong correlation ($r_s = 0.72$, $p < 0.001$).

2.2.1 Comparison of the new 2019 SES score (CBS) to that of 2017 (SCP). Number of postal codes per category.

	SES score 2019 (CBS)					
SES score 2017 (SCP)	**lowest**	**below average**	**average**	**above average**	**highest**	**unknown**
lowest	302	176	46	13	27	1
below average	57	199	226	118	48	3
average	18	73	204	286	132	3
above average	8	27	114	316	375	4
highest	4	7	26	119	624	4
unknown	1	5	3	3	11	472

We may compare this to the correlation between the SES categories of 2016 and 2014 (both calculated by SCP) which is the same time interval, see table 0. These scores were more strongly correlated, ($r_s = 0.92$, $p < 0.001$). We found the same correlation coefficient for a three-year interval (2014-2017), showing that the changes from the SCP calculated scores over 2017 to the CBS calculated scores over 2019 are larger than the changes over time within SCP calculated scores. In general, many postal codes tend to score a bit higher in the new SES-WOA score.

2.2.2 Comparison of the 2016 SES score to that of 2014 (both SCP). Number of postal codes per category.

	SES score 2016 (SCP)					
SES score 2014 (SCP)	**lowest**	**below average**	**average**	**above average**	**highest**	**unknown**
lowest	483	61	0	0	0	0
below average	71	440	117	7	0	0
average	1	131	463	185	2	1
above average	1	13	136	559	131	1
highest	0	0	1	87	645	5
unknown	3	1	0	1	7	12

Note: the number of postal codes in category 'unknown' is less than in table 2.2.1 because the SCP classification excludes postal codes of post office boxes and non-residential areas (e.g. industrial areas).

We also compared the HSMR 2021 outcomes per hospital when using the SCP scores (2017) and when using the CBS scores (2019) in the 2021 data of the HSMR model. The HSMR outcomes differed less than 2 points. So although there are differences between the two scores, the impact on the HSMRs is modest. However, this is also influenced by the fact that the CBS scores are only used in one year (2021) of the four years (2018-2021) in the HSMR model.

2.3 Additional Charlson comorbidity code

ICD-10 code G83.5 (Locked-in syndrome), which can be registered in the LBZ from 2021, was added to Charlson comorbidity group 12 (Hemiplegia or paraplegia); see table 3.4.1.

3. (H)SMR model

Expected in-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 to calculate 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, are based on LBZ data of four years: year t-3 up to year t. The addition of an additional year (t-3) increases the stability and accuracy of the estimates, while the moving four-year period up to year t keeps 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 LBZ. 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 that participate in the LBZ. In case of partial non-response by hospitals, only the fully registered months are included in the model, as in the other months fatal cases might be registered completely and non-fatal cases partially. The partially registered months of those hospitals are removed from the model as these might otherwise unjustly influence the estimates. In addition, if for any reason registered data of hospitals in a specific LBZ year had not been not validated by DHD, that year of data is not included in the HSMR model.

All of the above-mentioned hospitals were included in the model, with the exception of the unvalidated 2018 data of two hospitals that had closed in 2018. Data of the short-stay specialised hospital that had started operating as an independent hospital in 2018, were also included. (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 (i.e. inpatient admissions, and prolonged observations, unplanned, without overnight stay) of Dutch residents in Dutch hospitals in a certain period, except admissions that do not meet the billing criteria of the Dutch Healthcare Authority for inpatient admissions and prolonged observations. 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 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 (NZa), and it replaces the majority of the acute one-day inpatient admissions that had formerly been registered. It involves more mortality than day cases, and it is therefore relevant to include this case type in the HSMR.

Admissions that do not meet the billing criteria of the Dutch Healthcare Authority are removed from the data in all consecutive model years. This primarily concerns one-day inpatient admissions where the patient returned home after discharge. Also, about 100 in-hospital deaths where the patient was admitted after 20:00 hrs. and died before 24:00 hrs. on the same day, were removed from the dataset.
For the years 2020 and 2021 all admissions with COVID-19 as the principal diagnosis (ICD-10 codes U07.1 (COVID-19, virus identified (lab confirmed)), U07.2 (COVID-19, virus not identified (clinically diagnosed)) and U10.9 (Multisystem inflammatory syndrome associated with COVID-19, unspecified)) were removed from the dataset (see section 2.1).

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 259 categories 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 ICD-10 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 allows the evaluation of 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, with the following adaptations:

[1] *Clinical Classifications Software is developed for the Healthcare Cost and Utilization Project (HCUP) with the purpose of clustering ICD-codes into clinically meaningful categories, see http://www.hcup-us.ahrq.gov/toolssoftware/ccs10/ccs10.jsp*

- 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”.

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 are quite similar to the names of the chapters of the ICD-10, there is no one-to-one relation between the two. Although most ICD-10 codes of a CCS group do fall within one ICD-10 chapter, often some of the codes are categorised in other chapters. Especially the codes from the R chapter of ICD-10 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 and thus the (H)SMRs are adjusted for these characteristics. 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, socioeconomic status, severity of the main diagnosis, urgency of admission, Charlson comorbidities, source of admission, year of discharge and month of admission. These variables are described below. For the variables socioeconomic status, severity of the 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 of 0 or 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 (socioeconomic status) of the postal area of patient's home 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, as registered in the LBZ. For LBZ years up to 2020 the SES-variable developed by the Netherlands Institute for Social Research (SCP) was used. SCP had collected data on income, employment and educational level per four-digit postal area on which they performed principal component analyses to get a SES score for each four-digit postal area. Population-weighted quintiles were calculated from these SCP scores, resulting in the six SES categories mentioned above. For 2021 and later the SES classification is derived from scores calculated by Statistics Netherlands (Arts *et al.*, 2022). This SES score is based on household data concerning welfare (a combination of income and wealth), level of education and recent labour participation. The scores are determined at the household level and then averaged to scores per four-digit postal code. Household-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 (category "unknown"), were added to the category "average" if collapsing was necessary.
For 2018-2020 the 2017 SCP classification was used and for 2021 the 2019 classification calculated by Statistics Netherlands was used in the HSMR model.

Severity of the 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 ICD-10 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 subdiagnoses (individual ICD codes), which may differ in severity (mortality risk). To classify the severity of the subdiagnoses, 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 occurring ICD subdiagnoses of the admissions in the current model years, using data of six historical LMR/LBZ 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 individual ICD codes. The higher severity categories only occur in a few diagnosis groups.

Six historical LMR/LBZ years are used to determine this 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 ICD-9-CM only, and the severities were also determined for ICD-9-CM codes. Main diagnoses registered in ICD-10 were converted to ICD-9-CM to determine the severity covariate. As In 2012-2013 hospitals transitioned from using ICD-9-CM to code the diagnoses of admissions to using ICD-10, the diagnoses used for the HSMR 2014-2016 calculation are all coded in ICD-10, and the historical dataset used to determine the severities also partly consists of ICD-10 coded diagnoses. Therefore, for the HSMR 2014-2016 and later HSMR models, the severities are

determined for ICD-10 diagnoses. For the HSMR 2019-2021 the severity classification was based on the LMR/LBZ of 2013-2018, which mostly consists of ICD-10 data. A method was developed to calculate the severity for ICD-10 main diagnoses with such historical datasets, ensuring a gradual shift over time from severities based on ICD-9 data to severities based solely on ICD-10 data. The method and an investigation of the effects of this change are described in the HSMR 2017 methodological report (CBS, 2018).

For the severity classification the Dutch ICD-10-ICD-9-CM conversion table was used (table "ICD-10 – CvZ80", see https://www.whofic.nl/downloads-en-links/icd-10). As this table had not been updated for recent years, newly added ICD-10 codes in recent years did not have a converted ICD-9-CM code or a converted "old" ICD-10 code (used prior to the introduction of the new ICD-10 code). For these new codes, in consultation with DHD, we added conversions (default counterpart codes in ICD-9-CM and ICD-10) to the conversion table, to make it complete. When an ICD-10 code and its ICD-9-CM equivalent did not occur in the historical dataset, a severity of "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 Comorbidity 17/11 is set to zero in the coefficients file, and the coefficient of the less severe analogue (Comorbidity 9/11) should be used for Comorbidity 17/11.

The individual ICD-10 codes with the corresponding severity categories 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 medical treatment or aid within 24 hours is necessary. Within 24 hours means 24 hours from the moment the specialist decides that acute admission is necessary.

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 ICD-10 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.

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 ICD-10 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). Furthermore, newly introduced ICD-10 codes have been added by CBS if the corresponding 'old' ICD-10 code is part of a comorbidity group.

3.4.1 Comorbidity groups of Charlson index and the corresponding ICD-10 codes

No.	Comorbidity groups	ICD-10 codes
1	Myocardial infarction	I21, I22, I25.2
2	Congestive heart failure and cardiomyopathy	I50, I11.0, I13.0, I13.2, I25.5, I42, I43, P29.0
3	Peripheral vascular disease	I70, I71, I731, I73.8, I73.9, I77.1, I79.0, I79.2, K55.1, K55.8, K55.9, Z95.8, Z95.9, R02, Z99.4
4	Cerebrovascular disease	G45.0-G45.2, G45.4, G45.8, G45.9, G46, I60-I69
5	Dementia	F00-F03, F05.1, G30, G31.1
6	Pulmonary disease	J40-J47, J60-J67
7	Connective tissue disorder	M05, M06.0, M06.3, M06.9, M32, M33.2, M34, M35.3
8	Peptic ulcer	K25-K28
9	Liver disease	B18, K70.0-K70.3, K70.9, K71.3-K71.5, K71.7, K73, K74, K76.0, K76.2-K76.4, K76.8, K76.9, Z94.4
10	Diabetes	E10.9, E11.9, E12.9, E13.9, E14.9
11	Diabetes complications	E10.0-E10.8, E11.0-E11.8, E12.0-E12.8, E13.0-E13.8, E14.0-E14.8
12	Hemiplegia or paraplegia	G04.1, G11.4, G80.1, G80.2, G81, G82, G83.0-G83.5, G83.8, G83.9
13	Renal disease	I12.0, I13.1, N01, N03, N05.2-N05.7, N18, N19, N25, Z49.0-Z49.2, Z94.0, Z99.2
14	Cancer	C00-C26, C30-C34, C37-C41, C43, C45-C58, C60-C76, C81-C85, C86.0-C86.6, C88, C90-C97, D47.5
15	HIV	B20-B24, O98.7
16	Metastatic cancer	C77-C80
17	Severe liver disease	I85.0, I85.9, I86.4, I98.2, I98.3, K70.4, K71.1, K72.1, K72.9, K76.5, K76.6, K76.7

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

Year of discharge: *2018, 2019, 2020, 2021*.
Inclusion of the year of discharge 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.
The criteria for excluding a hospital from calculating HSMRs, based on the characteristics of the registered inpatient admissions and prolonged observations without overnight stay, were:

Insufficient participation in the LBZ

- Hospitals are excluded if they do not register all inpatient admissions and "prolonged observations, unplanned, without overnight stay" that meet the billing criteria of the Dutch Healthcare Authority (NZa) in the LBZ.

Data quality

Hospitals are excluded if*:*

- ≤30% of admissions are coded as acute.
- ≤1.5 secondary diagnoses are registered per admission, on average per hospital.[2]

Case mix

Hospitals are excluded if:

- Observed mortality is less than 60 in all registered admissions.

Admissions with COVID-19 as main diagnosis were excluded from the dataset that was used to calculate the outcomes of the data quality and case mix criteria.

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

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

[2] 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 ICD-10 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.

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

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

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

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

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 one. As a result of collapsing the categories, the degrees of freedom can be smaller than the original number of categories minus one.

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

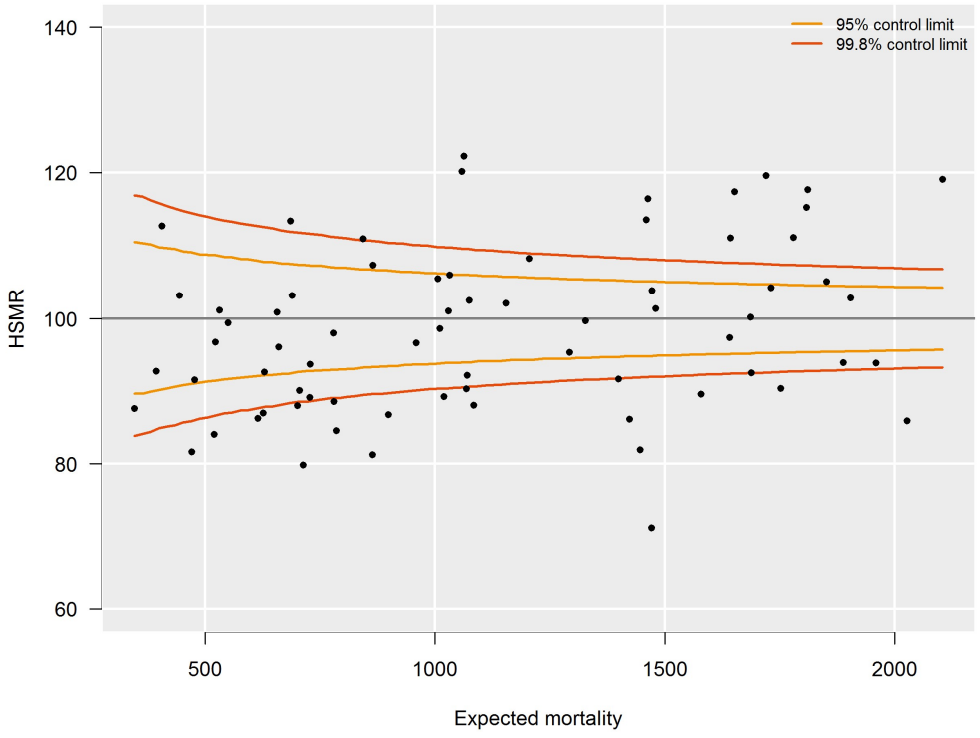
A confidence interval, i.e. an upper and lower confidence limit, is calculated for each SMR and HSMR. For the HSMR and most SMRs a confidence level of 95 percent is used, while for the SMRs of the 157 diagnosis groups a confidence level of 98 percent is used to reduce the number of undue statistically significant SMRs as a result of the large number of comparisons made when evaluating 157 diagnosis groups. 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 be 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.

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. 2015-2017 and 2016-2018) overlap, and that the three-year figure is less up-to-date than the figure of the last year. Therefore we also calculated the HSMR figures for the most recent 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 most recent year, but not in the three-year period, this is a signal for further investigation, as the quality of care may have deteriorated. There can also be other reasons for this, e.g. differences in registration practices over the years, but it is a signal for the hospital for further investigation.

3.6.4 Funnel plot HSMR (example).



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.

3.6.5 P-Values

From 2017 onwards, it was decided to also calculate p-values for the SMRs of the 157 diagnosis groups. The reason is that high SMRs for diagnosis groups are often an important starting point for further research and hospitals might need an extra tool for prioritizing such research, in case of multiple high SMRs. The lower the p-value, the more the observed mortality deviates from the expected mortality. Also, because of the large number of diagnosis groups, there is a risk of incorrectly labelling SMRs as significantly high of low (so called type I errors, or false-positives, due to multi-testing). The p-values can also be used to correct for this issue. The p-values are not included in the reports sent to the hospitals, but hospitals can request them from Dutch Hospital Data.

Separate p-values are given for the alternative hypotheses: "the observed mortality (O_{dh}) is higher than the expected mortality (E_{dh})" and "the observed mortality (O_{dh}) is lower than the expected mortality (E_{dh})." The p-values belonging to the these hypotheses are denoted by $p_{high}(O_{dh})$ and $p_{low}(O_{dh})$ respectively. The main reason for calculating two separate p-values is that by using a confidence of 99 percent for each of the two tests results in the same significant SMRs as found with the 98 percent confidence interval of the SMRs. Another reason is that often the main interest is $p_{high}(O_{dh})$.

The p-value of null-hypothesis "the observed mortality is lower or equal to the expected mortality" is given by the probability of observing a mortality equal to or higher than the observed mortality given the expected mortality:

$$p_{high}(O_{dh}) = \Pr(X \geq O_{dh}|E_{dh}) = 1 - \Pr(X < O_{dh}|E_{dh}). \quad (3.6.8)$$

Assuming that the observed mortality follows a Poisson-distribution with and expected value equal to the expected mortality this is equal to

$$p_{high}(O_{dh}) = 1 - \mathrm{P}_{E_{dh}}(X \leq O_{dh}) + \mathrm{P}_{E_{dh}}(X = O_{dh}), \quad (3.6.9)$$

With $\mathrm{P}_{E_{dh}}(X \leq O_{dh})$ the cumulative distribution function and $\mathrm{P}_{E_{dh}}(X = O_{dh})$ the probability distribution function of the Poisson-distribution with an expected value of E_{dh}.

Likewise the p-value of the null-hypothesis "the observed mortality is higher or equal to the expected mortality" is given by:

$$p_{low}(O_{dh}) = \mathrm{P}_{E_{dh}}(X \leq O_{dh}). \quad (3.6.10)$$

4. Evaluation of the HSMR 2021 model

This chapter presents and evaluates the model results. Summary outcomes of the 157 logistic regressions are presented, with in-hospital 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 2021.xlsx", published together with this report.

4.1 Target population and data set

All hospitals that register complete records of inpatient admissions and prolonged observations without overnight stay in the LBZ are included in the HSMR model. In 2019 and 2020 all general hospitals and university hospitals were included in the model, as well as three short-stay specialised hospitals (two cancer hospitals and an eye hospital). In 2018 two general hospitals were excluded because they had closed in 2018 and the registered data of that year was incomplete and unvalidated. In the period 2018-2021 all of the other hospitals had completely registered admission data in the LBZ.
Based on the hospital units in 2021 (counting merged hospitals as one unit in previous years also), plus the closed hospitals that were included in the model in previous years, the total number of hospitals included in the HSMR model of 2018-2021 is 75 and includes 64 general hospitals, 8 university hospitals and 3 short stay specialised hospitals.

4.1.1 Admissions in HSMR 2021 model.

	2018	2019	2020	2021	total
Excluded admissions not meeting the NZa criteria*	87 723	138 310	123 889	131 521	481 443
Excluded admissions of foreigners	9 221	9 825	6 231	5 993	31 270
Excluded admissions due to COVID-19**			40 430	60 228	100 658
Total number of admissions included in model	1 647 563	1 636 522	1 413 747	1 426 945	6 124 777
Number of inpatient admissions	*1 541 992*	*1 523 148*	*1 306 395*	*1 314 945*	*5 685 884*
Number of observations	*105 571*	*113 374*	*107 352*	*112 596*	*438 893*
Number of deaths included in model	33 417	32 647	29 151	29 779	124 994
Crude mortality (in admissions in model)	2.0%	2.0%	2.1%	2.1%	2.0%

**Admissions that do not meet the billing criteria of the Dutch Healthcare Authority (NZa) for inpatient admissions, and for prolonged observations, unplanned, without overnight stay. The number of these admissions in the LBZ varies over the years, due to different registration instructions of DHD.*
***Admissions with COVID-19 as the main diagnosis (ICD-10 codes U07.1 (COVID-19, virus identified (lab confirmed)), U07.2 (COVID-19, virus not identified (clinically diagnosed)) and, from 2021 onwards, U10.9 (Multisystem inflammatory syndrome associated with COVID-19, unspecified)).*

Table 4.1.1 lists some characteristics of the admissions included in the HSMR model by model year. Admissions not meeting the criteria of the Dutch Healthcare Authority and admissions of foreigners were excluded. For the years 2020 and 2021, admissions with COVID-19 as the principal diagnosis were additionally excluded. Due to that and due to the overall impact of the COVID-19 pandemic on regular hospital care, the number of admissions included in 2021 was on average still 13% lower compared to that of 2019, while in 2020 it was on average 14% lower. As a result, the total number of admissions in the 2021 model (6 124 777)

was 4.5% lower compared to the 2020 model (6 416 367), which was 5.3% lower compared to the 2019 model (6 777 534). This decrease of admissions due to the COVID-19 pandemic was higher than the 2.8% reduction of admissions between the 2019 (6 777 534) and the 2018 model (6 974 721).
The crude mortality of the non-COVID-19 admissions included in the HSMR model stabilizes around 2.0% in the period 2018-2021. The COVID-19 admissions (not included in the model) had a much higher in-hospital mortality rate (15.5% in 2020 and 13.8% in 2021).

4.2 Hospital exclusion

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 (2019-2021) hospitals had to fulfil these criteria for the three consecutive years.
Of the 75 hospitals included in the model, all had registered (valid) data over 2021. The three 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 the other hospitals. In fact, all of these hospitals had participated in the LBZ but the data of these hospitals 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 other 71 hospitals that had granted authorization fulfilled the criteria and were provided with a HSMR figure for 2021.
For these 71 hospitals the data of 2020 and 2019 was additionally investigated in order to determine if a three-year report could be provided. The data of all 71 hospitals met the criteria in all years considered and so all 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, 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 a large majority of the diagnosis groups. This is also true for several of the comorbidity groups, especially for group 2 (Congestive heart failure and cardiomyopathy). Comorbidity 15 (HIV) is only rarely registered as a comorbidity; most diagnosis groups had fewer than 50 admissions with HIV comorbidity. In only one of the models comorbidity 15 was statistically significant. The number of times month of admission is significant varies over the years (CBS, 2018, 2019, 2020, 2021): it increased from 27 in 2017 to 43 in 2018, but in 2019 it decreased to 37, to 35 in 2020 and to 33 in 2021. The number of models in which year of discharge was significant, has dropped over the years: from 72 in the HSMR 2012-2015 model (CBS, 2016) to 22 times in the HSMR 2018-2021 model. Compared to the HSMR 2020 model, the largest changes were observed for the covariates comorbidity 1 (Myocardial infarction, no longer significant in 11 models) and comorbidity 5 (Dementia, no longer significant in 7 models). For the other covariates the changes are smaller. The total number of significant covariates decreased from 1 666 in 2020 to 1 632 in 2021.

4.3.1 Statistical significance of the covariates for the 157 logistic regressions (summary), HSMR 2021 model.

Covariate	No. of significant results	Covariate	No. of significant results
Age	136	Comorbidity 17	50
Comorbidity 2	126	Comorbidity 1	46
Urgency	126	Sex	45
Severity main diagnosis	117	Comorbidity 10	33
Comorbidity 13	112	Month of admission	33
Comorbidity 16	104	Comorbidity 12	32
Source of admission	96	Comorbidity 11	25
Comorbidity 3	95	Comorbidity 7	23
Comorbidity 6	90	Year of discharge	22
Comorbidity 9	88	SES	22
Comorbidity 14	69	Comorbidity 8	14
Comorbidity 4	68	Comorbidity 15	1
Comorbidity 5	59		

Comorbidity groups: 1 Myocardial infarction, 2 Congestive heart failure and cardiomyopathy, 3 Peripheral vascular disease, 4 Cerebrovascular disease , 5 Dementia, 6 Pulmonary disease, 7 Connective tissue disorder, 8 Peptic ulcer, 9 Liver disease, 10 Diabetes, 11 Diabetes complications, 12 Hemiplegia or paraplegia, 13 Renal disease, 14 Cancer, 15 HIV, 16 Metastatic cancer, 17 Severe liver disease.

The significance only partially measures the effect of the covariates on mortality. This is better measured using the Wald statistic. Table 0 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 the main diagnosis has the highest explanatory power. Age and urgency of admission are also important variables. Of the comorbidities in the model, comorbidity groups 2, 16, and 13 are the groups with the most impact on mortality. Compared to the outcome of the 2020 model (CBS, 2021), the order of the covariates is almost identical. Note that the values of the Wald statistics themselves cannot be compared directly as these values depend on the number of admissions used in the models. This also explains the smaller values in the model of 2020 and 2021 compared to the 2019 model.

When comparing the outcomes of the past seven models (HSMR 2015 up to HSMR 2021) the explanatory power of year of discharge has decreased with 68%. This implies that the differences in mortality between the years in the model (corrected for differences in patient characteristics) have decreased. In the HSMR models of 2019 and 2020 the Wald statistics of year of discharge has however stabilized, and it is showing a slight increase again in the HSMR model of 2021. The impact of comorbidity 1 (myocardial infarction) is still decreasing, with in total a 46% decrease over the past seven models. The impact of comorbidity 12 (hemiplegia or paraplegia) has increased considerably (86%) over the past seven HSMR models. An increased impact of a comorbidity could reflect an increased effect of the comorbidity (e.g. the likelihood of dying in hospital when having this condition as comorbidity has increased), and/or an increased number of patients with this comorbidity resulting in more accurate estimates of the effect of this comorbidity (which also increases the Wald statistic), and vice versa applies in the case of a decreased impact of a variable. The impact of the SES variable, which was relatively stable in previous years, has increased by 12% compared to the 2020 model, which might be a

result of the use of updated values for the most recent year in the current model (see section 2.2 for more details on the update of the SES variable in 2021).

4.3.2 Wald chi-square statistics for the 157 logistic regressions, HSMR 2021 model.

Covariate	Sum of Wald statistics	Sum of df	Covariate	Sum of Wald statistics	Sum of df
Severity main diagnosis	39 137	398	Comorbidity 4	1 246	126
Age	29 158	1 979	Comorbidity 5	1 224	120
Urgency	16 325	155	SES	931	693
Comorbidity 2	9 540	145	Comorbidity 12	832	101
Comorbidity 16	4 480	137	Sex	809	148
Comorbidity 13	3 857	151	Year of discharge	779	469
Source of admission	3 134	285	Comorbidity 1	692	147
Comorbidity 3	2 616	145	Comorbidity 10	447	151
Comorbidity 6	2 300	152	Comorbidity 11	348	118
Comorbidity 9	1 795	125	Comorbidity 7	285	127
Comorbidity 17	1 568	61	Comorbidity 8	131	32
Month of admission	1 375	782	Comorbidity 15	17	11
Comorbidity 14	1 268	143			

Comorbidity groups: 1 Myocardial infarction, 2 Congestive heart failure and cardiomyopathy, 3 Peripheral vascular disease, 4 Cerebrovascular disease , 5 Dementia, 6 Pulmonary disease, 7 Connective tissue disorder, 8 Peptic ulcer, 9 Liver disease, 10 Diabetes, 11 Diabetes complications, 12 Hemiplegia or paraplegia, 13 Renal disease, 14 Cancer, 15 HIV, 16 Metastatic cancer, 17 Severe liver disease.

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 a 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 2021, because hospitals probably 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 and the order of the covariates in table 4.3.3 (ranked by magnitude of the effect) remains unchanged.

4.3.3 Average shift in HSMR 2021 by inclusion/deletion of covariates.

Covariate	Average shift in HSMR	Covariate	Average shift in HSMR
Comorbidity*	5.08	Source of admission	1.15
Age	3.78	SES	0.76
Severity main diagnosis	2.76	Month of admission	0.16
Urgency	2.34	Seks	0.09

**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. Overall the C-statistics have changed little compared to the previous model. Only two C-statistics showed a moderate increase: after last year's increase from 0.89 to 0.94, the C-statistic of "Cancer of testis and other male genital organs" increased further to 0.98. In addition, the C-statistic of "Schizophrenia, mental retardation, preadult disorders and other mental conditions" increased from 0.87 to 0.92. All other changes are smaller than 0.04 with most of them below 0.02. For 70 diagnosis groups the C-statistic did not change compared to last year.

Only three of the 157 diagnosis groups have a C-statistic below 0.70: "Congestive heart failure, nonhypertensive" (70), "Chronic obstructive pulmonary disease and bronchiectasis" (82) and "Aspiration pneumonitis; food/vomitus" (84). For the 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. In 2021, 59 diagnosis groups had a C-statistic above 0.9.

4.4.1 C-statistics for the logistic regressions of the 157 main diagnosis groups, HSMR 2021 model.

Diag. group no.	Description of diagnosis group	Number of admissions	Number of deaths	C-statistic
1	Tuberculosis	1 605	43	0.88
2	Septicemia (except in labor)	13 973	3 799	0.73
3	Bacterial infection; unspecified site	9 500	544	0.81
4	Mycoses	2 371	234	0.81
5	HIV infection	752	31	0.84
6	Hepatitis, viral and other infections	24 483	244	0.94
7	Cancer of head and neck	15 159	229	0.89
8	Cancer of esophagus	10 956	576	0.80
9	Cancer of stomach	12 278	472	0.82
10	Cancer of colon	44 628	1 143	0.84
11	Cancer of rectum and anus	19 272	402	0.86
12	Cancer of liver and intrahepatic bile duct	7 693	435	0.82
13	Cancer of pancreas	18 242	859	0.82
14	Cancer of other GI organs; peritoneum	8 664	399	0.79
15	Cancer of bronchus; lung	57 486	4 220	0.82
16	Cancer; other respiratory and intrathoracic	2 725	139	0.84
17	Cancer of bone and connective tissue	8 224	88	0.92
18	Melanomas of skin and other non-epithelial cancer of skin	6 925	119	0.94
19	Cancer of breast	41 071	404	0.97
20	Cancer of uterus	8 976	124	0.92
21	Cancer of cervix and other female genital organs	11 340	98	0.93
22	Cancer of ovary	8 661	262	0.85
23	Cancer of prostate	25 668	343	0.93
24	Cancer of testis and other male genital organs	6 481	10	0.98
25	Cancer of bladder	54 772	475	0.92
26	Cancer of kidney, renal pelvis and other urinary organs	15 773	281	0.89

27	Cancer of brain and nervous system	12 034	259	0.79
28	Cancer of thyroid	5 956	54	0.97
29	Hodgkin`s disease	2 020	33	0.94
30	Non-Hodgkin`s lymphoma	22 951	961	0.82
31	Leukemias	21 952	1 202	0.81
32	Multiple myeloma	10 793	464	0.80
33	Cancer; other and unspec. primary; maintenance chemotherapy and radioth.	5 057	128	0.91
34	Secondary malignancies	81 040	4 190	0.77
35	Malignant neoplasm without specification of site	2 809	333	0.83
36	Neoplasms of unspecified nature or uncertain behavior	12 082	210	0.90
37	Other and unspecified benign neoplasm	67 515	90	0.88
38	Thyroid and other endocrine disorders	23 411	219	0.91
39	Diabetes mellitus without complication	14 463	116	0.88
40	Diabetes mellitus with complications	23 875	507	0.85
41	Nutritional deficiencies and other nutritional, endocrine, and metabolic disorders	56 173	475	0.94
42	Fluid and electrolyte disorders	34 936	896	0.84
43	Cystic fibrosis	2 245	9	0.91
44	Immunity and coagulation disorders, hemorrhagic disorders	10 834	185	0.87
45	Deficiency and other anemia	47 295	449	0.80
46	Diseases of white blood cells	8 881	214	0.77
47	Mental, affective, anxiety, somatoform, dissociative, and personality disorders	21 336	64	0.88
48	Senility and organic mental disorders	10 182	497	0.71
49	Schizophrenia, mental retardation, preadult disorders and other mental cond.	6 585	20	0.92
50	Other psychoses	2 974	28	0.85
51	Meningitis, encephalitis, and other central nervous system infections	9 488	511	0.87
52	Parkinson`s disease	5 614	92	0.86
53	Multiple sclerosis and other degenerative nervous system conditions	12 512	257	0.91
54	Paralysis and late effects of cerebrovascular disease	3 849	59	0.88
55	Epilepsy and convulsions	43 641	552	0.89
56	Coma, stupor, and brain damage	2 262	231	0.91
57	Headache and other disorders of the sense organs	63 786	42	0.94
58	Other nervous system disorders	62 529	338	0.95
59	Heart valve disorders	36 603	845	0.78
60	Peri-, endo-, myocarditis, and cardiomyopathy	21 790	654	0.88
61	Essential hypertension, hypertension with compl., and secondary hypertension	13 297	100	0.94
62	Acute myocardial infarction	136 537	3 483	0.85
63	Coronary atherosclerosis and other heart disease	114 840	839	0.85
64	Nonspecific chest pain	153 818	37	0.93
65	Pulmonary heart disease	31 283	1 061	0.81
66	Other and ill-defined heart disease	1 663	113	0.90
67	Conduction disorders (heart disease)	25 120	340	0.89

68	Cardiac dysrhythmias	184 102	766	0.90
69	Cardiac arrest and ventricular fibrillation	15 230	5 578	0.75
70	Congestive heart failure, nonhypertensive	123 091	9 672	0.67
71	Acute cerebrovascular disease	151 557	12 739	0.80
72	Transient cerebral ischemia, and other cerebrovascular disease	45 486	322	0.89
73	Peripheral and visceral atherosclerosis	44 723	1 883	0.91
74	Aortic and other artery aneurysms	27 109	2 453	0.89
75	Aortic and arterial embolism or thrombosis	11 990	451	0.87
76	Other circulatory disease	29 747	563	0.87
77	Phlebitis, varicose veins, and hemorrhoids	11 439	143	0.87
78	Pneumonia	116 964	9 751	0.76
79	Influenza	19 266	1 037	0.80
80	Tonsillitis and upper respiratory infections	57 305	103	0.95
81	Acute bronchitis	25 256	63	0.96
82	Chronic obstructive pulmonary disease and bronchiectasis	110 277	6 304	0.69
83	Asthma	25 736	99	0.91
84	Aspiration pneumonitis; food/vomitus	8 365	1 708	0.65
85	Pleurisy; pneumothorax; pulmonary collapse	23 170	618	0.84
86	Respiratory failure; insufficiency; arrest	5 835	1 768	0.74
87	Lung disease due to external agents	1 777	168	0.79
88	Other lower respiratory disease	24 509	967	0.88
89	Other upper respiratory disease	56 405	331	0.90
90	Intestinal infection	47 615	562	0.89
91	Disorders of mouth, teeth, and jaw	20 867	48	0.96
92	Esophageal disorders	12 885	112	0.92
93	Gastroduodenal ulcer	4 953	242	0.91
94	Gastritis, duodenitis, and other disorders of stomach and duodenum	7 494	86	0.89
95	Appendicitis and other appendiceal conditions	69 813	69	0.96
96	Peritonitis and intestinal abscess	5 041	346	0.81
97	Abdominal hernia	44 997	448	0.92
98	Regional enteritis and ulcerative colitis	17 800	63	0.95
99	Intestinal obstruction without hernia	33 236	1 506	0.83
100	Diverticulosis and diverticulitis	35 625	508	0.91
101	Anal and rectal conditions	22 475	42	0.94
102	Biliary tract disease	119 397	1 020	0.90
103	Liver disease; alcohol-related	7 191	895	0.71
104	Other liver diseases	17 616	1 020	0.83
105	Pancreatic disorders (not diabetes)	35 481	785	0.83
106	Gastrointestinal hemorrhage	37 289	1 039	0.81
107	Noninfectious gastroenteritis	11 884	195	0.81
108	Other gastrointestinal disorders	37 070	772	0.95
109	Nephritis; nephrosis; renal sclerosis	14 585	106	0.91
110	Acute and unspecified renal failure	15 989	1 015	0.77
111	Chronic kidney disease	14 048	444	0.89
112	Urinary tract infections	98 045	2 665	0.77
113	Calculus and other diseases of urinary tract	82 585	200	0.91
114	Genitourinary symptoms and ill-defined conditions	26 505	104	0.88

115	Hyperplasia of prostate and other male genital disorders	41 233	44	0.91
116	Non-neoplastic breast conditions	14 480	3	0.98
117	Prolapse and other female genital disorders	61 907	31	0.99
118	Complications of pregnancy, childbirth, and the puerperium; liveborn	584 859	14	0.85
119	Skin and subcutaneous tissue infections	52 821	723	0.90
120	Other skin disorders, chronic ulcer of skin	17 563	242	0.91
121	Infective arthritis and osteomyelitis	13 473	286	0.89
122	Osteoarthritis, rheumatoid arthritis, and other musculoskeletal deformities	228 199	186	0.92
123	Other non-traumatic joint disorders	11 582	34	0.92
124	Spondylosis, back problems, and osteoporosis	84 408	196	0.97
125	Pathological fracture	5 267	81	0.81
126	Other connective tissue disease	35 022	271	0.97
127	Cardiac and circulatory congenital anomalies	9 129	184	0.87
128	Noncardiac congenital anomalies	27 323	219	0.95
129	Short gestation; low birth weight; and fetal growth retardation	73 614	473	0.89
130	Intrauterine hypoxia, perinatal asphyxia, and jaundice	61 351	215	0.95
131	Other perinatal conditions	223 277	237	0.95
132	Joint disorders and dislocations; trauma-related; sprains and strains	23 906	33	0.98
133	Fracture of neck of femur (hip)	85 487	2 712	0.80
134	Skull and face fractures, spinal cord injury	11 662	222	0.90
135	Fracture of upper limb	43 314	141	0.95
136	Fracture of lower limb	49 933	275	0.94
137	Other fractures	44 303	1 012	0.86
138	Intracranial injury	38 756	2 523	0.78
139	Crushing injury or internal injury	20 746	455	0.91
140	Open wounds of head; neck; and trunk	5 911	47	0.86
141	Open wounds of extremities	5 305	37	0.95
142	Complication of device, implant or graft	96 462	1 330	0.88
143	Complications of surgical procedures or medical care	100 112	846	0.86
144	Superficial injury; contusion	59 456	476	0.92
145	Burns	3 856	95	0.92
146	Poisoning by psychotropic agents, drugs, or other medications	36 954	342	0.87
147	Other injuries and conditions due to external causes	12 913	655	0.89
148	Syncope	43 984	126	0.85
149	Fever of other and unknown origin	21 367	117	0.83
150	Lymphadenitis and gangrene	4 160	26	0.96
151	Shock	1 218	455	0.72
152	Nausea and vomiting	12 352	70	0.85
153	Abdominal pain	42 345	118	0.95
154	Malaise and fatigue	9 475	143	0.83
155	Allergic reactions	11 709	31	0.96
156	Rehabilitation and other aftercare, medical examination/evaluation/screening	83 026	207	0.86
157	Residual codes; unclassified	67 023	152	0.98

4.5 Regression coefficients

The file “coefficients HSMR 2021.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 2020.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. COVID-19 2021 model

For admissions in 2021 with COVID-19 as main diagnosis, which were excluded from the HSMR 2021 model, a separate model was developed to calculate the SMR for individual hospitals for COVID-19 admissions. This chapter describes the methods used to develop the COVID-19 2021 model, and presents and evaluates the results. Mostly the same methods are used as for the HSMR 2021 model. Method differences and specific aspects of the COVID-19 model are described below.

5.1 Target population

In paragraph 3.1 and 4.1 the target population and dataset of the HSMR 2021 model are described, including the criteria for hospitals and admissions to be included in the model. For the COVID-19 2021 model the same target population and dataset were used. A total number of 74 hospitals are included in the COVID-19 model, including 64 general hospitals, 8 university hospitals and 2 short stay specialised hospitals. One short stay specialised hospital which is included in the HSMR model did not have any COVID- 19 admissions and was, therefore, not included in the COVID-19 model.

For admissions, the same criteria were applied as for the HSMR model with the exception of the selection of the ICD-10 main diagnoses. In the COVID-19 model only admissions with a main diagnosis of COVID-19 (ICD-10 codes U07.1 (COVID-19, virus identified (lab confirmed)), U07.2 (COVID-19, virus not identified (clinically diagnosed)) and U10.9 (Multisystem inflammatory syndrome associated with COVID-19, unspecified)) are included. A total number of 60.228 COVID-19 admissions were included in the model. Table 5.1.1 lists the numbers of admissions and deaths per COVID-19 subdiagnosis.

5.1.1 Admissions in the COVID-19 2021 model for each subdiagnosis*

	U07.1	U07.2	U10.9	Total
Excluded admissions not meeting the NZa criteria**	170	4	2	176
Excluded admissions of foreigners	204	2	5	211
Total number of admissions included in model	58 675	838	715	60 228
Number of inpatient admissions	*57 564*	*819*	*714*	*59 097*
Number of observations	*1 111*	*19*	*1*	*1 131*
Number of deaths included in model	8 078	141	112	8 331
Crude mortality (in admissions in model)	13.8%	16.8%	15.7%	13.8%

**Admissions with COVID-19 as the main diagnosis included the ICD-10 codes U07.1 (COVID-19, virus identified (lab confirmed)), U07.2 (COVID-19, virus not identified (clinically diagnosed)) and U10.9 (Multisystem inflammatory syndrome associated with COVID-19, unspecified)).*
***Admissions that do not meet the billing criteria of the Dutch Healthcare Authority (NZa) for inpatient admissions, and for prolonged observations, unplanned, without overnight stay.*

The in-hospital crude mortality rate of the admissions in the COVID-19 2021 model was 13.8%, which is much higher than the crude mortality rate of the non-COVID-19 admissions in the HSMR 2021 model (2.1% in 2021; see Table 4.1.1). During the tail of the second COVID-19-wave (up to January 2021, only including patients that were discharged in 2021) the crude mortality rate of the COVID-19 admissions was 18.0%. The crude mortality rate decreased to 6.9% in June

2021 and increased again to 16.2% in the third COVID-19-wave (from September 2021 onwards, only including patients that were discharged in 2021).

5.2 Selection of covariates

To adjust for differences in the case mix between hospitals, covariates of patient and admission characteristics that are associated with in-hospital mortality are included in the model. In the COVID-19 model the same covariates were included as in the HSMR model (see section 3.4), with the following exceptions or modifications:

Year of discharge was not included as a covariate because the model was only developed for admissions with a discharge date in 2021.

Severity of the main diagnosis has been adjusted for the COVID-19 model. The model only includes COVID-19 admissions, which comprise three ICD-10 subdiagnoses. As no multiple year in-hospital mortality data is available for these COVID-19 subdiagnoses, we were not able to calculate the severity of each diagnosis in the same way as done in the HSMR model. Therefore all three COVID-19 subdiagnoses are included in the COVID-19 model as separate categories: *U07.1 (reference), U07.2 and U10.9.*

Month of admission is a variable that has six 2-month categories in the HSMR model. For the COVID-19 model this variable was tested against a new, more detailed variable **'month of admission 2021'**. This new variable includes all months of the year as separate 1-month categories and includes an additional category 'before 2021' for admissions that started before 2021. This new variable was introduced to be able to follow in more detail outcome of the different COVID-19 waves. By adding the category 'before 2021' admissions that had started in e.g. December 2020 are separated from admissions that had started in December 2021 (category 'December'). In the month of admission variable used in the HSMR these admissions are all included in one category ('November-December'). The new 'month of admission 2021' variable for the COVID-19 model has therefore 13 categories: *0-before 2021 (reference), 1-January, 2-February, …, 12-December.*

Vaccination rate is an additional covariate that was tested in the COVID-19 model, as vaccination rate may differ between regions and vaccination status can be associated with mortality (RIVM and CBS, 2022; Haas *et al.*, 2021). As data on individual vaccination status were not available for this purpose, admissions were assigned to a vaccination rate category on the basis of the age and postal code of the patient. The vaccination rate was defined as the percentage of the population that had received at least one COVID-19 vaccination as per December 31 2021, per four-digit postal code and per age group (18-29, 30-59 and ≥60 yrs). These data were received from the National Institute for Public Health and the Environment (RIVM). For the age group below 18 yrs no vaccination data were available. Admissions of patients <18 yrs, together with a small group of patients with postal codes for which no vaccination rates were available either, were therefore placed in a category 'Unknown'. The vaccination rate (%) was categorized into the following groups: *[0-40) , [40-60), [60-70), [70-75), [75-80), [80-85), [85-90), [90-100] (reference), Unknown.*

The interaction between vaccination rate and 'month of admission 2021' was also investigated as the vaccination rate was determined in December 2021 while the actual vaccination rate varied strongly during the year (vaccination started in January 2021). Furthermore, the

interaction between age and 'month of admission 2021' was investigated as age groups may have been affected differently in different COVID-19 waves.

5.3 Development of the model

To develop the COVID-19 model first a basic model was fitted, including the HSMR covariates age, sex, socioeconomic status, urgency of admission, Charlson comorbidities, source of admission and month of admission (six 2 month categories). This model had a C-statistic of 0.8019. The additional value of the covariates month of admission 2021, severity of the main diagnosis, and vaccination rate were subsequently evaluated using a likelihood-ratio test. Each covariate was separately added to the model and covariates with a *p*-value < .05 were retained.

Replacing covariate 'month of admission' (six 2 months categories) by the adjusted covariate 'month of admission 2021' (13 categories) resulted in a significantly improved fit of the model ($p < 0.01$; C-statistic = 0.8023) and this variable is therefore included in the COVID-19 model instead of the covariate 'month of admission'. Subsequently, the adjusted covariate 'severity of main diagnosis' (also called 'COVID-19 subdiagnosis' in this chapter) was added, which also improved the model fit ($p < .001$; C-statistic = 0.8025), and therefore this covariate is included as well.

The subsequent addition of the vaccination rate did not show to improve the fit of the model significantly ($p = 0.422$; C-statistic = 0.8026), nor did the interaction between vaccination rate and 'month of admission 2021' ($p = 0.321$; C-statistic = 0.8039). Consequently, the vaccination rate was not included in the COVID-19 model. The non-significant effect of vaccination rate on the model fit is reflected by the low correlation between the average vaccination rate and the SMR per hospital ($r = 0.08$) and the similarity in vaccination rate between hospitals. As shown in table 5.3.1, most hospitals had a mean vaccination rate of 85-90% (n = 42) or 90-100% (n=14). Note that the vaccination rate per hospital was here calculated as the average vaccination rate of the hospital's COVID-19 admissions, based on the postal code and age group of the patients. This is therefore not based on the vaccination status of the individual patients. The vaccination rate variable tested in the COVID-19 model was not based on the vaccination status of individual patients either, which may partly explain the non-significant explanatory power of this variable.

5.3.1 Average vaccination rate in hospitals (n=74) in COVID-19 2021 model

Average vaccination rate	0-70%	70-75%	75-80%	80-85%	85-90%	90-100%	Unknown
Number of hospitals	0	1	7	9	42	14	1

Adding the interaction between age and 'month of admission 2021' did not significantly improve the fit of the COVID-19 model either ($p = 0.755$; C-statistic = 0.8047) and therefore this interaction was not included in the model.

The final COVID-19 2021 model therefore includes the following covariates: age, sex, socioeconomic status, urgency of admission, Charlson comorbidities, source of admission, month of admission 2021 (13 categories) and severity of main diagnosis (three COVID-19 subdiagnosis categories). This final model has a C-statistic of 0.803.

5.4 Impact of the covariates

Table 5.4.1 shows the effect of the covariates on mortality, presenting the Wald statistic and whether the covariates had a statistically significant (95 percent confidence) impact on in-

hospital mortality: “1” indicates (statistical) significance, and “0” non-significance. The covariates are ordered by the value of the Wald statistic. It shows that age has the highest explanatory power, followed by comorbidity 2 (congestive heart failure and cardiomyopathy), comorbidity 13 (renal disease), and comorbidity 14 (cancer). All, but one (comorbidity 15: HIV), covariates showed to have a significant impact on in-hospital mortality, which is also seen in the 157 logistic regressions in the HSMR 2021 model. When comparing the Wald statistics of the COVID-19 model and the HSMR 2021 model, the order of covariates is different. The severity of the main diagnosis has no high explanatory power in the COVID-19 model in contrast to the HSMR 2021 model. This may be explained by the fact that the ICD-10 categories U07.2 and U10.9 contain relatively few admissions. Urgency of admission also has a lower explanatory power in the COVID-19 model compared to the HSMR 2021 model. This may be explained by the fact that 94% of the COVID-19 admissions were acute. In contrast, sex and month of admission have a higher impact on in-hospital mortality in the COVID-19 model than in the HSMR 2021 model. The latter may be due to the fact that the refinement of the variable in individual months, including a ‘pre 2021’ category, allows better differentiation between the COVID-19 waves (see section 5.2).

5.4.1 Wald chi-square statistic and statistical significance of the covariates, COVID-19 2021 model

Covariate	Wald	Significant	Covariate	Wald	Significant
Age	4087	1	Comorbidity 4	28	1
Comorbidity 2	248	1	SES	19	1
Comorbidity 13	233	1	Comorbidity 7	19	1
Comorbidity 14	153	1	Severity diagnosis	19	1
Source of admission	146	1	Comorbidity 11	18	1
seks	78	1	Urgency	14	1
Month of admission	76	1	Comorbidity 9	13	1
Comorbidity 3	56	1	Comorbidity 17	12	1
Comorbidity 1	46	1	Comorbidity 10	12	1
Comorbidity 5	37	1	Comorbidity 12	7	1
Comorbidity 16	30	1	Comorbidity 15	2	0
Comorbidity 6	29	1			

Comorbidity groups: 1 Myocardial infarction, 2 Congestive heart failure and cardiomyopathy, 3 Peripheral vascular disease, 4 Cerebrovascular disease , 5 Dementia, 6 Pulmonary disease, 7 Connective tissue disorder, 9 Liver disease, 10 Diabetes, 11 Diabetes complications, 12 Hemiplegia or paraplegia, 13 Renal disease, 14 Cancer, 15 HIV, 16 Metastatic cancer, 17 Severe liver disease.
Comorbidity 8 (peptic ulcer) is not shown as it was dropped from the model due to collapsing.

As is explained in section 4.3, even if a covariate is a strong predictor for mortality, the effect of this covariate on the SMR can still be small when hospitals differ little on the covariate. Table 5.4.2 shows the shift in the COVID-19 SMR by the inclusion/deletion of each covariate, indicating the impact of each covariate on the SMR. Age shows to have the highest impact on the COVID-19 SMR, followed by comorbidity as a group. These covariates also have the highest impact on the HSMR, but there the comorbidities have the highest impact (see table 4.3.3). Severity of the main diagnosis and urgency have a lower impact on the COVID-19 SMR than on the HSMR, while source of admission and SES have a higher impact. Month of admission and sex also have a somewhat higher impact on the COVID-19 SMR than on the HSMR, but the shift in SMR is small. So although month of admission and sex have a high impact on in-hospital COVID-

19 mortality, these covariates do not differ much between hospitals and thus will not contribute much to differences in SMR values.

5.4.2 Average shift in SMR by inclusion/deletion of covariates, COVID-19 2021 model

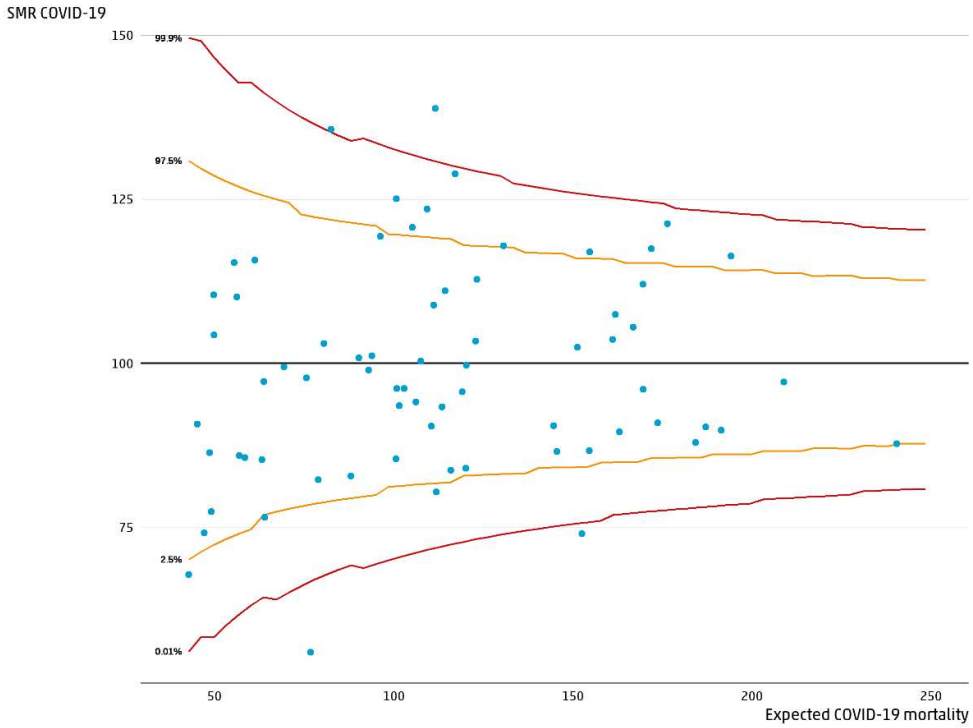
Covariate	Shift in SMR	Covariate	Shift in SMR
Age	5.21	Severity main diagnosis	0.79
Comorbidity*	3.61	Urgency	0.69
Source of admission	2.30	Month of admission	0.42
SES	1.06	Sex	0.35

**The comorbidities were deleted as one group and not separately.*

5.5 Model evaluation

The COVID-19 model has a C-statistic of 0.803, which is lower than the median C-statistic of the 157 diagnosis groups in the HSMR 2021 model. In the HSMR 2021 model the C-statistics ranged from 0.65 to 0.99, with a median of 0.88 (interquartile range: 0.83 – 0.92). A value of 0.80 indicates a model with good explanatory power. However, it does leave room for possible uncorrected differences in case mix between the hospitals.

5.5.1 Funnelplot of SMRs 2021 for the COVID 19 diagnosis group*



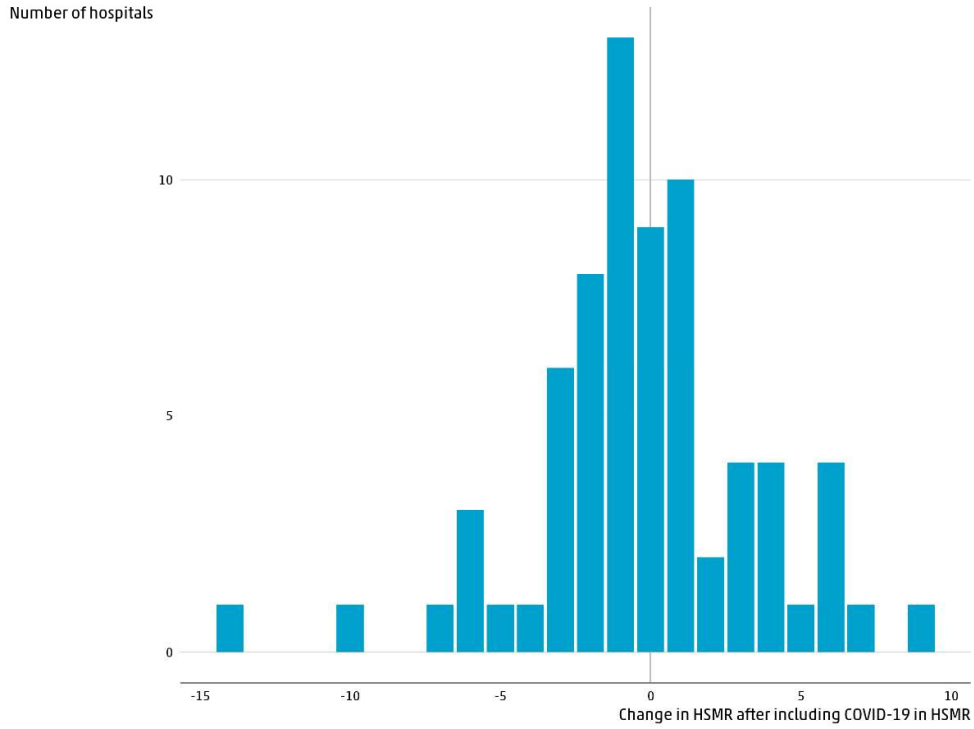
** Included are the 71 hospitals for which HSMR 2021 outcomes were calculated, except one hospital that is omitted from the figure because of disclosure control.*

Using the model it is possible to calculate SMRs for the COVID-19 diagnosis group (i.e. the COVID-19 admissions) for each of the hospitals. Figure 5.5.1 shows the funnel plot of these SMRs. This figure is similar to that of the HSMR: more hospitals are outside of the 95% and 99.8% expectation intervals than would be expected based on chance. This indicates that there are significant differences between the hospitals with respect to their mortality rates for COVID-19. It is of course possible that these differences are partly caused by unexplained population differences not accounted for by the COVID-19 model.

When comparing the SMRs for the COVID 19 diagnosis group to the HSMRs (not including COVID 19) there is a significant positive correlation of 0.33 ($p < 0.01$). Hospitals that score high on their HSMR thus tend to some degree to also score high on the SMR for COVID 19 and vice versa. The correlation is however not very strong.

Using the SMRs for COVID-19 and the HSMRs without COVID-19, it is also possible to calculate what the HSMR would become when COVID-19 would be included in the HSMR. Figure 5.5.2 shows that for most hospitals the HSMR 2021 would only change by a few points when COVID-19 would be included into the HSMR. However, there are also hospitals that show a larger change in their HSMR. Table 5.5.3 shows that there are eight hospitals for which the HSMR changes from significant low, high or not-significant (95% confidence) to another category, of which seven hospitals shift from a not-significant HSMR to a significantly high (n=2) or low (n=5) HSMR.

5.5.2 Distribution of the changes in the HSMRs 2021when COVID-19 would be included into the HSMR.



5.5.3 Number of hospitals with a significantly high of low HSMR (95% confidence) for the HSMR 2021 and the HSMR 2021 including COVID-19.

	HSMR including COVID-19			
HSMR	**Significantly low**	**Not significant**	**Significantly high**	**Total**
Significantly low	11	0	0	**11**
Not significant	5	42	2	**49**
Significantly high	0	1	10	**11**
Total	**16**	**43**	**12**	**71**

5.6 Regression coefficients

Table 5.6.1 shows the estimated regression coefficients of the logistic regression of the COVID-19 2021 model, as well as the standard errors. 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). The reference category is given in the first row of the corresponding covariates, and by definition has a coefficient of zero.

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 (see section 3.6.2). In the COVID-19 2021 model categories 'age = 0 years' and 'age = 1-5 years' were collapsed, resulting in equal coefficients for the collapsed categories. Comorbidity 8 (peptic ulcer) was also collapsed and therefore dropped from the model because the covariate only has two categories. Comorbidity 8 is therefore not shown in table 5.6.1.

Looking at specific beta's, it can be concluded that the COVID-19 subdiagnosis 'U07.2' and 'U10.9' have a higher mortality risk than 'U07.1'. Furthermore, females have a lower mortality risk than males, and overall a higher age is associated with a higher mortality risk. Lower SES categories have a higher mortality risk than SES categories average and above. Admissions that had started before 2021 (end of 2020) have a higher mortality risk than admissions that had started in 2021, and patients being admitted from nursing homes or from other hospitals also have a higher mortality risk than patients being admitted from home.

5.6.1 Regression coefficients of the covariates, COVID-19 2021 model

Coefficient	Beta	Standard error
Intercept	-6.459	1.003
COVID-19 subdiagnosis = U07.1 (reference)	0.000	0.000
COVID-19 subdiagnosis = U07.2	0.282	0.103
COVID-19 subdiagnosis = U10.9	0.417	0.119
Sex = Male (reference)	0.000	0.000
Sex = Female	-0.238	0.027
Age = 0-1 + 1-5 (reference)	0.000	0.000
Age = 5-10	1.163	1.419
Age = 10-15	0.981	1.417
Age = 15-20	1.767	1.227
Age = 20-25	1.070	1.225

Age = 25-30	1.763	1.069
Age = 30-35	1.653	1.044
Age = 35-40	1.493	1.040
Age = 40-45	2.269	1.015
Age = 45-50	2.070	1.011
Age = 50-55	2.836	1.002
Age = 55-60	3.169	1.001
Age = 60-65	3.720	1.000
Age = 65-70	4.223	1.000
Age = 70-75	4.595	0.999
Age = 75-80	4.887	0.999
Age = 80-85	5.216	0.999
Age = 85-90	5.348	1.000
Age = 90-95	5.441	1.001
Age = 95+	5.478	1.012
Urgency of admission = Elective (reference)	0.000	0.000
Urgency of admission = Acute	0.263	0.072
Comorbidity 1 = 0 (reference)	0.000	0.000
Comorbidity 1 = 1	0.266	0.039
Comorbidity 2 = 0 (reference)	0.000	0.000
Comorbidity 2 = 1	0.716	0.044
Comorbidity 3 = 0 (reference)	0.000	0.000
Comorbidity 3 = 1	0.367	0.048
Comorbidity 4 = 0 (reference)	0.000	0.000
Comorbidity 4 = 1	0.482	0.090
Comorbidity 5 = 0 (reference)	0.000	0.000
Comorbidity 5 = 1	0.383	0.062
Comorbidity 6 = 0 (reference)	0.000	0.000
Comorbidity 6 = 1	0.166	0.030
Comorbidity 7 = 0 (reference)	0.000	0.000
Comorbidity 7 = 1	0.288	0.065
Comorbidity 9 = 0 (reference)	0.000	0.000
Comorbidity 9 = 1	0.348	0.094
Comorbidity 10 = 0 (reference)	0.000	0.000
Comorbidity 10 = 1	0.099	0.029
Comorbidity 11 = 0 (reference)	0.000	0.000
Comorbidity 11 = 1	0.332	0.076
Comorbidity 12 = 0 (reference)	0.000	0.000
Comorbidity 12 = 1	0.374	0.140
Comorbidity 13 = 0 (reference)	0.000	0.000
Comorbidity 13 = 1	0.516	0.033
Comorbidity 14 = 0 (reference)	0.000	0.000
Comorbidity 14 = 1	0.702	0.055
Comorbidity 15 = 0 (reference)	0.000	0.000
Comorbidity 15 = 1	0.598	0.403
Comorbidity 16 = 0 (reference)	0.000	0.000
Comorbidity 16 = 1	0.489	0.089
Comorbidity 17 = 0 (reference)	0.000	0.000

Comorbidity 17 = 1	0.752	0.205
SES = Lowest (reference)	0.000	0.000
SES = Below average	0.011	0.038
SES = Average	-0.049	0.039
SES = Above average	-0.094	0.040
SES = Highest	-0.148	0.043
SES = Unknown	-0.067	0.205
Month of admission = before 2021 (reference)	0.000	0.000
Month of admission = January	-0.141	0.064
Month of admission = February	-0.144	0.068
Month of admission = March	-0.230	0.066
Month of admission = April	-0.289	0.066
Month of admission = May	-0.496	0.077
Month of admission = June	-0.476	0.149
Month of admission = July	-0.335	0.115
Month of admission = August	-0.215	0.094
Month of admission = September	-0.245	0.105
Month of admission = October	-0.158	0.079
Month of admission = November	-0.087	0.063
Month of admission = December	-0.246	0.067
Source of admission = Home (reference)	0.000	0.000
Source of admission = Nursing home or other institution	0.616	0.069
Source of admission = (Other) hospital	0.390	0.043

Comorbidity groups: 1 Myocardial infarction, 2 Congestive heart failure and cardiomyopathy, 3 Peripheral vascular disease, 4 Cerebrovascular disease , 5 Dementia, 6 Pulmonary disease, 7 Connective tissue disorder, 9 Liver disease, 10 Diabetes, 11 Diabetes complications, 12 Hemiplegia or paraplegia, 13 Renal disease, 14 Cancer, 15 HIV, 16 Metastatic cancer, 17 Severe liver disease.
Comorbidity 8 (peptic ulcer) is not shown as it was dropped from the model due to collapsing.

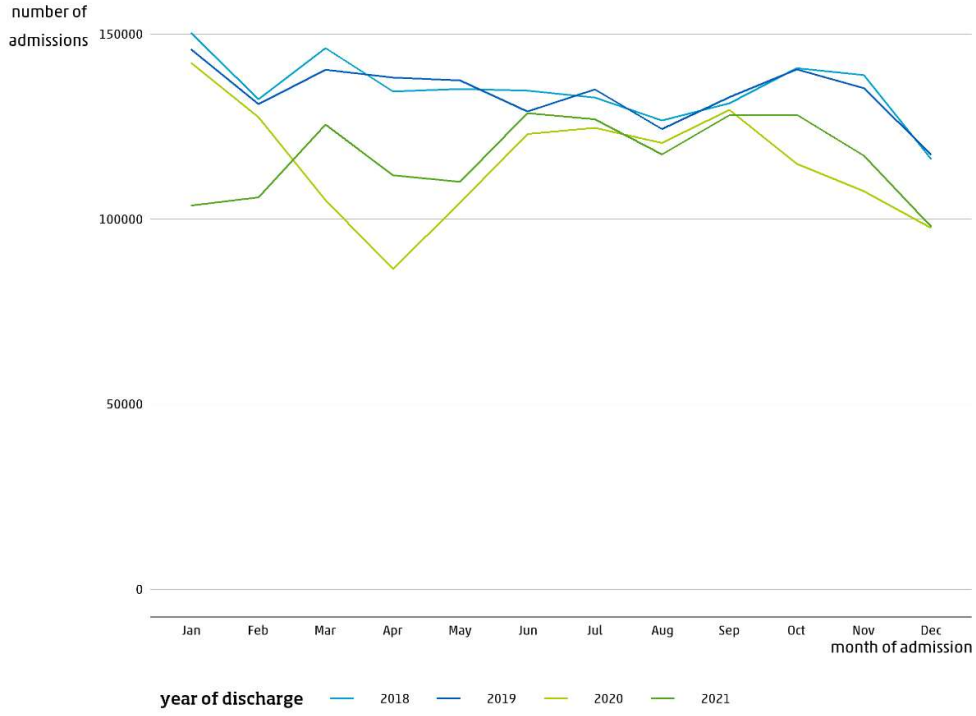
6. Effect of the COVID-19 pandemic on the HSMR

Although this year a model was estimated for admissions with COVID-19 as the main diagnosis (see chapter 5), it was decided, like last year, not to take these admissions into account in the current (H)SMR figures. This was done because at the start of the production of the HSMR 2021 outcomes it was not yet clear whether the HSMR model can sufficiently correct for differences in patient populations between hospitals for these admissions.
Last year (CBS, 2021) it was investigated whether or not the HSMR and SMR figures for the non-COVID-19 admissions for 2020 were affected by the COVID-19 pandemic. Overall all effects found were small and therefore it was concluded that the HSMR of 2020 was not strongly affected by the COVID-19 pandemic in general and that results were in general comparable to those of previous years. It is expected that also for 2021 the effects of the pandemic on the HSMR will be small or absent. In order to check if this is indeed the case, a limited set of the analyses done previous year have been repeated for 2021. The results of these analyses are described in this chapter.

6.1 Number of admissions and deaths

The annual number of hospital admissions selected for the HSMR model (i.e. inpatient admissions and prolonged observations without overnight stay) was low in 2020: the number of admissions had decreased by 14% when compared to 2019. For 2021 the number of admissions included in the HSMR model was reduced by 13% when compared to pre-pandemic year 2019

6.1.1 Number of hospital admissions by month of admission and year of discharge.[a]

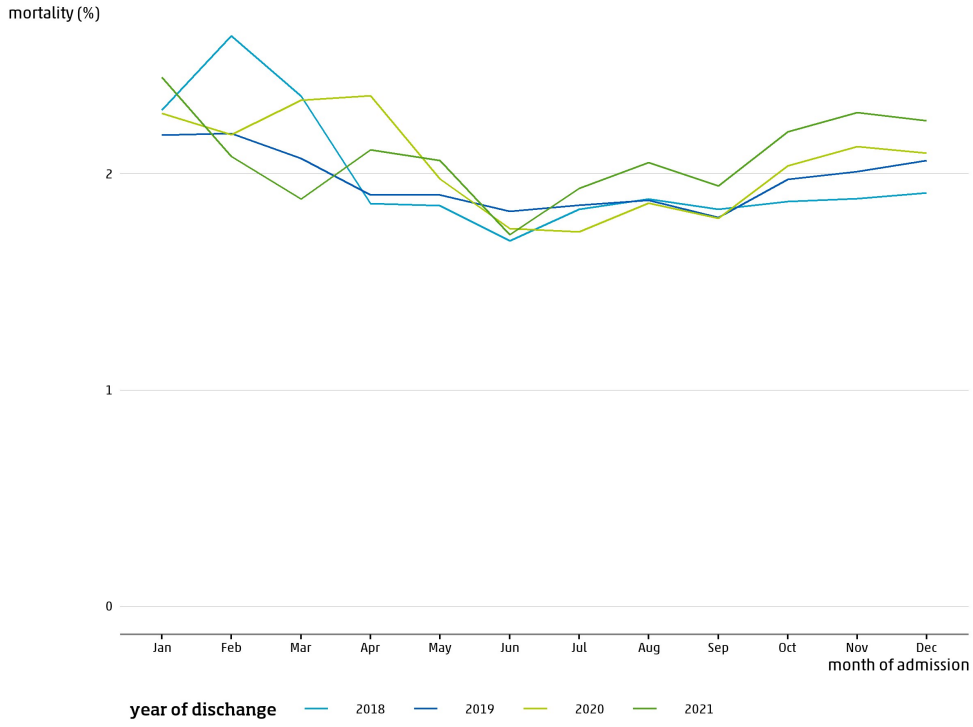


[a] *Admissions that had started in the year prior to the year of discharge are not included. Therefore, the displayed numbers of admissions in the last months of the year do not correctly reflect the total number of admissions in those months.*

(see table 4.1.1). This reduction of admissions is not only due to the exclusion of the COVID-19 admissions, but is for a larger part also due to downscaling of regular hospital care because of the pandemic.
Figure 6.1.1 shows the number of non-COVID-19 admissions per month of admission for 2021 and for earlier years. In pre-pandemic years the number of admissions is typically a little lower during the summer months. For most of the year 2021 the number of admissions is lower than in the two non-pandemic years 2018 and 2019. The number of admissions is especially lower during the first half of 2021 which corresponds to the tail of the second COVID-19 wave in the Netherlands. The number of admissions was however not as low as in April 2020 when the first COVID-19 wave caused a large drop in the number of non-COVID-19 admissions. Note that the number of admissions in the last months of the year (particularly December) is artificially lower in figure 6.1.1 for all years, as the admissions with an admission date in the end of the year and a discharge date in the next year are not shown.

Figure 6.1.2 shows that the in-hospital mortality rate is slightly higher in the second half of 2021 compared to the same months in previous years. This corresponds to the third corona wave starting around August/September 2021. In February 2018 the mortality rate was also higher, probably caused by the more severe influenza epidemic in that year, which probably led to more hospital admissions and mortality (included in the HSMR model). The next section will look at the mortality rates corrected for the patient population.

6.1.2 Crude in-hospital mortality rate by month of admission and year of discharge.[a]



[a] *Admissions that had started in the year prior to the year of discharge are not included.*

6.2 HSMR in relation to the COVID-19 admissions

In this paragraph we will explore whether there is a relationship between the HSMRs, based on the non-COVID-19 hospital admissions, and the pressure on hospitals caused by the COVID-19

pandemic in 2021. If the pandemic affects the HSMR, we would expect that this depends on the degree to which hospitals are affected by the pandemic. We will use the fraction of admissions with main diagnosis COVID-19 in the hospital as a measure for how severely the pandemic affected care in that hospital. Note that care can also be affected in other ways not visible in the fraction of admissions with main diagnosis COVID-19.

We analysed the relationship between the HSMRs 2021 of the regular hospital care versus the fraction of COVID-19 admissions per hospital. It appears that there is no statistically significant relationship (r=-0.07; p-value=0.54) between the two outcomes. We also analysed the relationship between the HSMR 2021 and the fraction of COVID-19 admissions for three types of hospitals (university hospitals, general hospitals with some specializations, and other general hospitals) and by size of the hospitals (number of admissions). This was tested by linear regression of the fraction of COVID-19 admissions on the HSMR by adding type of hospital or size as an interaction in the model. No statistically significant relationships were found here either.

However, also in previous years (before the pandemic) the HSMRs show quite a large inter-hospital variation. It could be that this regular variation in the HSMRs dominates the effect of the pandemic. However, hospitals that score high or low on their HSMR in one year are in general likely to score high or low again the next year. Therefore, any change in the HSMR caused by the pandemic, can be more clearly seen by comparing HSMR values of 2021 to that of 2019 (last pre-pandemic year) for each hospital. In that way we take into account the correlation between the yearly HSMRs. If the pandemic affected the HSMRs, we would expect systemic intra-hospital changes in the HSMRs that are related to the level of the pressure on the hospitals caused by the pandemic. We therefore analysed the relation between the difference between the HSMR in 2021 and 2019 for each individual hospital and the fraction of COVID-19 admissions in 2021. Again, there is no relation between the fraction of patients with COVID-19 and the change in the HSMR (r=0.006, p-value=0.959). We also analysed the relationship between the HSMR 2021 and HSMR 2019 difference and the fraction of COVID-19 admissions by type and size of hospital. Again, no relationships were found.

In the previous analyses we have not found an effect of the pandemic on the HSMRs. This is also confirmed by table 6.2.1 that shows the number of significantly high or low HSMRs and the variation in the HSMRs. These are also similar to those of previous years.

6.2.1 Number of HSMRs in 2019, 2020 and 2021 that are significantly different from 100 (α=0.05) and the variation in HSMRs per year.

Year of discharge	Number of HSMRs that are significantly lower than 100	Number of HSMRs that are significantly higher than 100	Standard deviation of the HSMRs
2019	12	10	11.24
2020	11	10	11.36
2021	11	11	11.83

7. Conclusions

Sections 4.4 and 4.5 evaluate the quality of the HSMR model of 2018-2021. No significant differences were found compared to the models of previous years. The importance of the variables in the model and the predictive power have remained largely the same even though nearly half of the admissions in the model are from years where the COVID-19 pandemic was present. Therefore, it seems that the present HSMR model, from which the COVID-19 admissions in 2020 and 2021 have been excluded, is not significantly affected by the COVID-19 pandemic. Last year (CBS, 2021) this was also extensively investigated for the HSMR of 2020 by comparing SMRs and HSMRs to the fraction of patients admitted with COVID-19, by comparing the SMRs and HSMRs between the different COVID-19 waves and by comparing different regions (which were differently affected by the pandemic). Only few small effects were found, and these were of the same order as fluctuations seen in years prior to the pandemic. Part of these analyses were also performed for the HSMR of 2021 (see chapter 6). Again no significant effects of the pandemic on the HSMR were found. We, therefore, again conclude that the usability of the regular (non-COVID-19) HSMR is not affected by the pandemic.

This year also a separate model for COVID-19 admissions (in 2021) was developed (see chapter 5). The developed COVID-19 model is similar to the model used for the regular HSMR. There are three differences in the used covariates:

1. **Year of discharge** was not included as a covariate because the model was only developed for admissions in 2021.
2. **Severity of the main diagnosis** has been adjusted for the COVID-19 model. The model includes the three ICD-10 subdiagnoses of COVID-19 as separate categories: *U07.1 (reference), U07.2 and U10.9.*
3. **Month of admission** includes six 2-month categories in the HSMR, but was adjusted to a new variable that includes all months of the year as separate 1-month categories and an additional category for admissions that started before 2021. This results in a total of 13 categories: *0-before 2021 (reference), 1-January, 2-February, …, 12-December.*

Vaccination rate per age group and postal area did not improve the model and was, therefore, not included in the model.

The COVID-19 2021 model has good predictive power (C-statistic of 0.80). Because of this, and the fact that hospitals have now gained experience in treating COVID-19 patients, it is possible to include COVID-19 in future HSMR models. We investigated the effect of including the COVID-19 2021 SMR into the HSMR 2021. Most hospitals showed a modest shift in their HSMR (less than 2 points). However, there are also hospitals that showed a larger shift in their HSMR. In general hospitals with a low regular HSMR (not including COVID-19) tend to score lower on the SMR for COVID-19 and vice versa, though the correlation is not very strong. Out of 71 hospitals there were five hospitals who shifted from a not significant HSMR to a significantly low HSMR when including COVID-19 admissions, two hospitals from a not significant HSMR to a significantly high HSMR, and one hospital shifted from significantly high to not significant.

When the COVID-19 admissions would be included into future HSMR models, the COVID-19 model may have to be adapted slightly. First, with multiple years of COVID-19 admissions, 'year of discharge' will have to be included into the model. Second, we now used a more

detailed encoding of month of admission in the COVID-19 model. In order to keep unity between the HSMR models, we could consider using the original HSMR month of admission variable (with six 2-month categories) again, also because we then have multiple years in the COVID-19 model and month of admission can then not be linked to a specific COVID-19 wave in a certain year anyway. Alternatively, we could consider to include the interaction between month of admission and year of discharge in the model. For 'severity of main diagnosis' the different subdiagnoses for COVID-19 will probably have to be kept as separate categories as there are not yet multiple years of historic COVID-19 mortality data to calculate the regular HSMR severity categories, and mortality may be different for different variants of the COVID-19 virus over the years.

8. References

Arts, K., R. van Gaalen, J. van der Laan, F. Linder, J. Mol, J. van Rooijen and C. Siermann (2022). Berekenwijze Sociaal Economische Status scores, Statistics Netherlands, The Hague/Heerlen. https://www.cbs.nl/nl-nl/maatwerk/2021/45/berekenwijze-ses-score-per-wijk-buurt

Bottle, A., B. Jarman 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--

Cecil, E., A. Bottle, A. Esmail, C. Vincent and P. Aylin (2020). What is the relationship between mortality alerts and other indicators of quality of care? A national cross-sectional study. Journal of Health Services Research & Policy 25(1): 13-21.

Chong, C.A.K.Y., G.C. Nguyen 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.

Haas, E.J., F.J. Angulo, J.M. McLaughlin, E. Anis, S.R. Singer, F. Khan, ... & S. Alroy-Preis (2021). Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. The Lancet, 397(10287), 1819-1829.

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., L.M. Peelen, K.G.M. Moons, C.J. Kalkman 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 347:f5913.

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 ICD-10 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/.

RIVM and CBS (2022). Sterfte en oversterfte in 2020 en 2021: Onderzoek door het CBS en het RIVM, in het kader van het ZonMw onderzoeksprogramma Oversterfte. National Institute for Public Health and the Environment (RIVM), Bilthoven, and Statistics Netherlands (CBS), The Hague/Heerlen.
https://www.rivm.nl/sites/default/files/2022-06/Eindrapport%20CBS%20en%20RIVM%20-%20Sterfte%20en%20oversterfte%20in%202020%20en%202021_def.pdf

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/2016/41/uitbreiding-hoofddiagnosegroepen-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 Erven, J.A., L.S. van Galen, A.A. Hettinga-Roest, E.P.J. Claessens, J.C. Roos, M.H.H. Kramer and P.W.B Nanayakkara (2018). Hospital standardised mortality ratio: A reliable indicator of quality of care?. Neth J Med 76(2): 72-77.

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 2021 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	1	0	1	1	-	-	-	-	-	1	-	-	0	0	-	-	0	-	-	-	-	0	0	1	1
2	1	1	1	1	0	1	1	1	0	1	0	1	1	0	1	0	1	1	-	1	1	0	1	1	1
3	1	1	1	0	1	1	0	1	0	1	0	-	0	0	1	1	1	1	-	0	1	0	0	0	0
4	1	0	1	1	0	1	0	-	-	0	1	-	0	0	0	-	1	0	-	1	-	0	0	0	0
5	0	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-	-	1	0	-	-	0	0	0	1
6	1	0	1	1	0	1	0	1	1	1	0	-	0	0	0	1	1	1	-	0	1	0	0	0	1
7	1	0	1	1	0	1	0	0	-	0	0	-	0	0	0	-	0	0	-	1	0	1	1	0	0
8	0	0	0	1	0	1	1	1	0	0	1	-	0	1	0	-	1	0	-	1	1	0	0	0	1
9	0	0	1	1	0	1	1	0	0	0	1	1	1	0	0	-	1	0	-	1	-	0	0	0	0
10	1	0	1	1	0	1	1	1	1	1	0	0	1	1	1	0	1	1	-	1	1	0	0	0	0
11	0	0	1	1	1	0	1	0	0	1	0	-	1	0	0	-	0	0	-	1	-	0	0	0	0
12	1	0	1	1	0	0	0	-	-	0	0	-	1	0	0	-	1	0	-	1	1	0	1	0	1
13	0	1	1	1	0	1	1	1	0	1	0	1	1	0	1	-	1	1	-	1	1	0	1	0	1
14	1	0	1	1	0	0	0	0	0	1	0	0	1	0	0	-	1	0	-	1	1	0	0	0	0
15	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	1	-	1	1	0	1	0	1

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

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

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

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

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

146	1	0	1	1	0	1	1	1	0	1	0	-	1	0	1	0	1	0	-	1	1	0	0	0	1
147	1	0	1	1	0	1	1	1	0	0	0	-	1	0	0	0	1	0	-	0	-	0	0	0	0
148	-	0	1	0	0	1	0	0	0	1	0	-	0	0	1	0	1	0	-	0	-	0	1	0	1
149	-	0	1	0	0	1	1	0	0	0	1	-	0	0	0	0	1	1	-	1	-	0	0	0	0
150	0	0	1	1	0	0	0	-	-	0	-	-	-	0	0	-	1	0	-	0	-	0	0	0	0
151	1	1	1	0	1	0	1	-	-	0	-	-	0	0	-	-	0	0	-	1	-	0	0	0	0
152	-	0	1	0	0	1	0	0	0	0	0	-	0	0	0	0	0	0	-	1	-	0	0	0	0
153	1	0	1	1	0	1	1	0	1	1	0	-	0	0	0	0	1	1	-	0	-	0	0	0	0
154	-	0	1	0	0	1	0	0	0	0	0	-	1	0	1	1	0	0	-	1	-	0	0	0	1
155	1	0	1	0	1	1	0	0	-	0	0	-	0	0	-	-	1	1	-	0	-	0	0	0	0
156	-	0	1	1	0	1	1	1	0	1	0	-	1	0	0	0	1	1	-	0	1	0	0	0	1
157	1	0	1	1	0	1	0	1	0	0	0	-	0	1	0	1	0	0	-	1	1	0	0	0	1
total	117	45	136	126	46	126	95	68	59	90	23	14	88	33	25	32	112	69	1	104	50	22	33	22	96

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

Comorbidity 1	- Myocardial infarction
Comorbidity 2	- Congestive heart failure and cardiomyopathy
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
Comorbidity 10	- Diabetes
Comorbidity 11	- Diabetes complications
Comorbidity 12	- Hemiplegia or 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 ICD-10 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
–	(between two numbers) inclusive
0 (0.0)	Less than half of unit concerned
2020–2021	2020 to 2021 inclusive
2020/2021	Average for 2020 up to and including 2021
2020/'21	Crop year, financial year, school year, etc., beginning in 2020 and ending in 2021
2018/'19–2020/'21	Crop year, etc., 2018/'19 to 2020/'21 inclusive

Because of rounding, some totals may not correspond to the sum of the separate cells.
Revised figures are not marked as such.

Colophon

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

Prepress: Statistics Netherlands
Design: Edenspiekermann

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