

BMJ Open Standardised mortality ratios as a user-friendly performance metric and trigger for quality improvement in a Flemish hospital network: multicentre retrospective study

Wim Tambreur,¹ Pieter Stijnen,² Guy Vanden Boer,² Pieter Maertens,² Caroline Weltens,³ Frank Rademakers,¹ Dirk De Ridder,⁴ Kris Vanhaecht,^{5,6} Luk Bruyneel^{5,6}

To cite: Tambreur W, Stijnen P, Vanden Boer G, *et al*. Standardised mortality ratios as a user-friendly performance metric and trigger for quality improvement in a Flemish hospital network: multicentre retrospective study. *BMJ Open* 2019;**9**:e029857. doi:10.1136/bmjopen-2019-029857

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2019-029857>).

Received 14 February 2019
Revised 03 July 2019
Accepted 31 July 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Wim Tambreur;
wim.tambreur@uzleuven.be

ABSTRACT

Objective To illustrate the development and use of standardised mortality rates (SMRs) as a trigger for quality improvement in a network of 27 hospitals.

Design This research was a retrospective observational study. The primary outcome was in-hospital mortality. SMRs were calculated for All Patient Refined—Diagnosis-Related Groups (APR-DRGs) that reflect 80% of the Flemish hospital network mortality. Hospital mortality was modelled using logistic regression. The metrics were communicated to the member hospitals using a custom-made R-Shiny web application showing results at the level of the hospital, patient groups and individual patients. Experiences with the metric and strategies for improvement were shared in chief medical officer meetings organised by the Flemish hospital network.

Setting 27 Belgian hospitals.

Participants 1 198 717 hospital admissions for registration years 2009–2016.

Results Patient gender, age, comorbidity as well as admission source and type were important predictors of mortality. Altogether the SMR models had a C-statistic of 88%, indicating good discriminatory capability. Seven out of ten APR-DRGs with the highest percentage of hospitals statistically significantly deviating from the benchmark involved malignancy. The custom-built web application and the trusted environment of the Flemish hospital network created an interoperable strategy to get to work with SMR findings. Use of the web application increased over time, with peaks before and after key discussion meetings within the Flemish hospital network. A concomitant reduction in crude mortality for the selected APR-DRGs from 6.7% in 2009 to 5.9% in 2016 was observed.

Conclusions This study reported on the phased approach for introducing SMR reporting to trigger quality improvement. Prerequisites for the successful use of quality metrics in hospital benchmarks are a collaborative approach based on trust among the participants and a reporting platform that allows stakeholders to interpret and analyse the results at multiple levels.

Strengths and limitations of this study

- Building on a long tradition of using hospital discharge datasets for financing mechanisms, this is the first study in Belgium illustrating the potential value of these data for collaborative quality improvement.
- Our standardised mortality ratio model showed good discriminatory capability, overall as well as for the included All Patient Refined—Diagnosis-Related Groups (APR-DRGs) separately.
- We custom built a web application to streamline the process from knowledge creation to knowledge sharing.
- While our models include APR-DRGs that account for about 80% of hospital mortality, these only reflect 30% of hospital stays, implying that caution is warranted when drawing conclusions about quality of care at the hospital level.
- To allow fair comparisons across hospitals, analyses will need to include postdischarge mortality, increased attention to readmission rates, close follow-up of palliative coding and disease-specific refinement of patient-mix adjustment.

INTRODUCTION

Identifying, monitoring and explaining variation in patient mortality is an intuitively appealing strategy for hospitals to increase insight into the quality of care delivered to their patients. Standardised mortality rates (SMRs) and hospital standardised mortality ratio (HSMR) are performance metrics that are used in health systems across the globe to make inferences about mortality. HSMR originated in the UK almost two decades ago.¹ This relatively simple expression of observed (ie, crude mortality) over expected mortality adjusted for differences in patient

mix has shown that variation in mortality across hospitals is substantial and persists over time.

Researchers in North America, Europe and Asia have reported statistically and clinically relevant decreases in HSMRs. In the Netherlands, a decrease in crude mortality and a constant decreasing trend of 8% in HSMR per year was noticed between 2003 and 2005, while the relative position of hospitals remained stable over the years.² In Canada, where HSMR findings are publicly reported nationwide since 2007,³ a 22% decrease in HSMR was observed between years 2006 and 2007 and 2012 and 2013. In the same period, crude mortality declined from 8.7% to 7.3%.⁴ Similar observations were made in other countries.^{5–7}

All authors from the above-mentioned studies fairly concluded that these decreases might reflect improvements in the quality of care, but at least in some part could be explained by changes in coding whether or not reflecting adverse behaviour to manipulate the system. For example, in countries like Canada^{4,8} and the UK⁹ where palliative coding is accounted for in the HSMR modelling, frequency of palliative care coding has been reported to be increasing over time. The HSMR methodology and the publication of HSMR in the form of league tables¹⁰ or in formats that intuitively make the user compare individual hospitals have been extensively criticised.^{11–13} As a consequence, other statistical approaches were tested but did not provide better solutions for examining variation in hospital mortality.^{13–16}

Such extensive methodological discussions about fair comparisons of hospital mortality are necessary to obtain an acceptable and meaningful indicator of quality of care. At the same time however, several large studies have found that there remains an undeniably large and unexplained variation in risk-adjusted hospital mortality between and within countries, and patient safety experts have called for national and international strategies to urgently tackle this variation.^{17–19}

In Flanders, Belgium, the Flemish hospital network KU Leuven (further referred as '*the Flemish hospital network*'), a not-for-profit association of 27 hospitals that aims to optimise quality and efficiency of patient care, embraced SMR as one approach to measure and compare quality of care among its members. SMR, like HSMR, is an expression of observed over expected patient-mix adjusted mortality, but for a homogenous patient group in terms of pathology. *The Flemish hospital network* acknowledged the known methodological issues regarding HSMR and SMR methodology and developed SMR models and a reporting framework enabling full transparency regarding the properties of the statistical model. This is facilitated by an easy to use web-accessible graphical user interface that can be further customised to include other quality or hospital performance indicators.

The purpose of this study is twofold. First, we describe methodological aspects regarding the calculation of the SMR using routinely collected hospital administrative data in Flanders, Belgium. Second, we illustrate

the implementation and application of a user-friendly web-based tool that supports a collaborative approach for chief medical officers (CMOs), clinicians and hospital management to study in detail their position to the reference population for the included SMRs. In addition, we illustrate that SMR reporting within a trusted collaborative environment could trigger awareness and quality improvement initiatives and might be helpful to understand and potentially lower in-hospital mortality rates.

METHODS AND ANALYSIS

Data collection

In Belgium, the data registered in the hospital discharge datasets (HDDs) are delivered every semester to the federal health authorities passing extensive quality checks. The Flemish hospital network members agreed to benchmark these data to share insights on hospital performance and quality of care. For this purpose, HDDs are securely sent to the central data management unit of the network and stored on a server with limited access. All analyses and the used methodology are shared with the members of *the Flemish hospital network* on a secure online platform. Written consent for these analyses was obtained from all members. All patient identifiers were pseudonymised and data were processed in accordance with article 6 of the General Data Protection Regulation (European Regulation 2016/679) as Belgian hospitals have legal obligation to support quality-of-care improvement initiatives.

Data validity, inclusion and exclusion criteria

The current study includes discharge data from 27 hospitals (26 regional and 1 academic hospital) for discharge years 2009–2016, with the exception of 2015 when the registration of International Classification of Diseases 10—Clinical Modification (ICD10-CM) diagnoses was not mandatory in Belgium. *The Flemish hospital network* members receive a data validation report containing descriptive feedback which allows each hospital to check the data that were submitted. The All Patient Refined—Diagnosis-Related Groups (APR-DRG) 31.0 (3M) grouping system was used to group hospital stays in homogenous groups. All hospitalised stays from general care hospitals were included in the analyses, with the exclusion of psychiatric stays and admissions grouped within APR-DRGs 950, 951, 952, 953, 954, 955, 956 and APR-Major Diagnostic Categories (MDCs) 14, 15, 22, 24, SS, hereafter referred to as *excluded pathology groups* (online supplementary table 1). These were identified by an expert panel of CMOs of *the Flemish hospital network*. Reasons for exclusion were irrelevance of the pathology for hospital mortality, vague description of APR-DRGs or APR-DRG's with ungroupable hospital stays. For hospital stays prior to 2015, diagnoses and procedures were coded in ICD-9-CM. Starting from 2016, coding was done in ICD-10-CM. Sixty-one APR-DRGs that accounted for 80% of the in-hospital mortality in the Flemish hospital network were retained (figure 1).¹

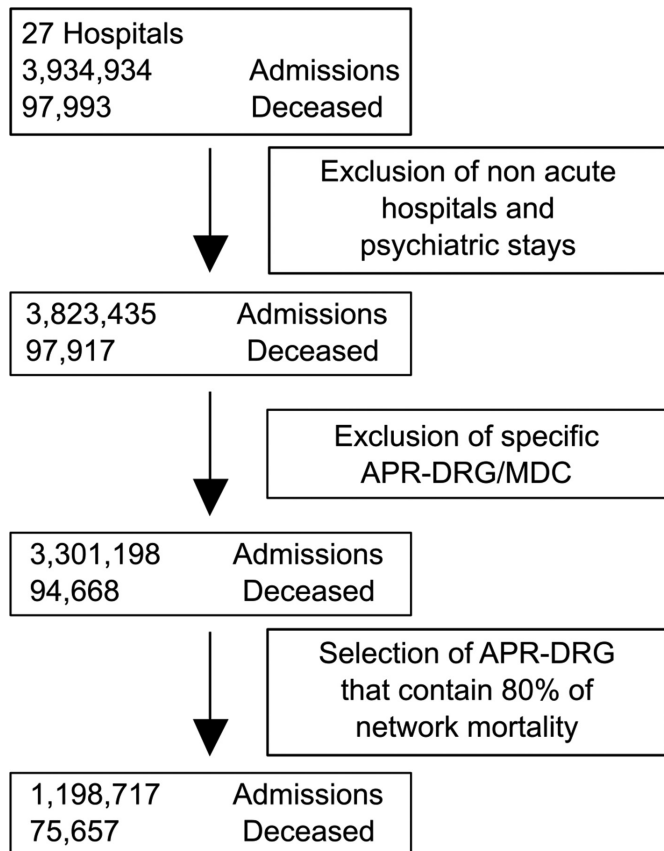


Figure 1 Inclusion criteria and selection of All Patient Refined—Diagnosis-Related Groups (APR-DRG). The hospital discharge sets contained information on 3.9 million hospital stays for 7 years for 27 hospitals. Psychiatric stays and hospital stays in non-acute hospitals were excluded. Specific APR-DRGs and APR-Major Diagnostic Categories (MDCs) (*excluded pathology groups*, online supplementary table 1) were excluded. From this dataset, APR-DRGs were selected that contributed most to 80% of the mortality. These selected APR-DRGs represent 36% of the hospital admissions (after the exclusion of *excluded pathology groups* and psychiatric stays).

Statistical analyses

The methodology developed by Jarman *et al*¹ was used to calculate a mortality benchmark per APR-DRG, SMR, for the participating hospitals. SMR is the expression of the observed mortality over the expected mortality. The latter is calculated by a logistic regression model with mortality as dependent variable and gender, age, comorbidity, admission source, admission type and discharge year as independent variables. Age was categorised as 10-year age bands which were, for each APR-DRG, grouped to contain at least 10 deaths to ensure a minimum number of events in each category. The Elixhauser Comorbidity Score was calculated according to international standards²⁰ and included as a continuous variable. In brief, the comorbidity score is the result of the sum of weighted coefficients of a separate logistic regression for mortality with 30 binary comorbidities as covariates. The delineation of each of the 30 comorbidities was accomplished using the ICD-9-CM and ICD-10-CM mappings of the

Agency for Healthcare Research and Quality.^{21 22} Admission source was categorised as follows: ‘Nursing home’, ‘Other hospital’, ‘Home’ or ‘On the road’. Admission type was defined as ‘Emergency’ or ‘Elective’. Discharge year was coded as a categorical variable (2009–2016).

The logistic regression procedure was employed with automated backward variable selection to estimate SMRs. For each of the selected APR-DRGs, the deletion criterion was set at $\alpha=0.10$ to prevent the unwanted deletion of relevant variables. As an internal validation, a leave-one-out cross-validation procedure was used. The C-statistics (a global assessment of model fit) depicted in online supplementary table 2, were calculated using these cross-validated values.²³ CIs (95%) were calculated using Byar’s approximation. The data analysis was generated using SAS software, V.9.4 of the SAS System for Windows.

To identify APR-DRG with large interhospital differences in SMR, we calculated for each SMR the percentage of hospitals for whom their ratio of expected versus observed mortality signalled deviation from the benchmark. This was defined as a hospital’s 95% CI excluding 1, which is the value for the number of observed deaths equalling the number of expected deaths.

The evolution in HSMR (SMR aggregated per hospital) over the seven discharge years was analysed for homogeneity of variance and differences between discharge years were statistically evaluated. This analysis was generated using the SAS Generalized Linear Mixed Models (GLIMMIX) procedure.

Development of an online platform

An online reporting platform was developed by the Management Information & Reporting Department of University Hospitals Leuven using a performant database system and R/Shiny technology.¹⁷ The output from the logistic regression models is formatted and stored on a password-protected SQL Server database (Microsoft SQL Server 2014).

The R/Shiny package¹⁷ is an open-source technology which allows to integrate the graphical and analytical capabilities of the R language to be used in a web application. This fits with the requirements for the web application set by the Flemish hospital network and has been used in other healthcare applications.²⁴ The workflow prior to publishing the web application entails querying the data and precalculating and aggregating the data in data slices. Each data slice contains the data needed to visualise the various graphs and are relatively small in size, increasing the speed of browsing the web application. The R datasets containing the data slices and the Shiny application files are stored on a Shiny server.²⁵ All figures were prepared in R V.3.4.1²⁶ using the Tidyverse package.

An illustrative video regarding the use of the online platform is available as online supplementary video (online supplementary video 1). To enable CMOs, hospital managers and clinicians to benchmark SMRs in the Flemish hospital network, different visualisations were designed (online supplementary figure 1). The

MDC-DRG grid (online supplementary figure 1 panel A) provides a quick overview on the performance of the hospital for each of the 61 included APR-DRGs. The user can navigate using this grid to each of the APR-DRGs where the hospital benchmark graphs are depicted in the form of funnel plots, bar charts and trend graphs (online supplementary figure 1 panels B, C and D). In all figures, a red colour indicates a significantly elevated SMR, a blue colour indicates an SMR that falls within the 95% confidence limits and a green colour indicates a significantly lowered SMR. In addition, this information is also available in a tabular format. For each pathology group, the model coefficients, Receiver operating characteristic curve and model fit statistics are provided, enabling full transparency regarding the modelling approach, the predictive value of the model and the variance explained by the model (online supplementary figure 1 panels E and F). This allows the user to understand the expected mortality for individual hospital stays (online supplementary figure 1 panel G).

Patient and public involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for implementation of the study, or were they asked to advise on interpretation or writing up of results. This research reflects a key step in developing system-level (hospital, medical unit, care programme) interventions whereby patients benefit through improved quality of care. Study results and the use of the online platform will be presented at local symposia and conferences that are attended by patients, patient advocates, healthcare providers and the broader community.

RESULTS

Calculation of standardised mortality ratios

In total, 1 198 717 hospital stays were included for fitting the models for 61 APR-DRGs (figure 1). This selection contains 30% of the total Flemish hospital network population, 77% of the Flemish hospital network mortality and 80% of the Flemish hospital network mortality after the exclusion of psychiatric stays and the *excluded pathology groups* (online supplementary table 1).

Crude mortality increased with increasing age group. Likewise, crude mortality rates were higher for patients with higher comorbidity scores. The majority of the included hospital deaths were emergency admissions (7.8%), and elective admissions had lower mortality rates (3.2%) (table 1). Higher crude mortality rates were observed for patients originating from nursing homes (16%) and other hospitals (12%), compared with patients coming from home (6%) (table 1). Average crude hospital mortality for the selected patient groups (figure 1) ranged from 4.9% to 8.6% across hospitals and from 0.7% (249—*non-bacterial gastroenteritis, nausea and vomiting*) to 81% (196—*cardiac arrest*) in the included APR-DRGs (online supplementary table 2).

The overall C-statistic of the SMR-models was 88%. The C-statistic ranged from 64% to 93% across models, with 52 of 61 models having a C-statistic higher than 70%.

Per APR-DRG one logistic regression model was fitted and predictors are listed in online supplementary table 2. All models included age, 31 models included gender, 48 models included admission type, 58 models included the comorbidity score and 57 models included admission source.

The direction of the effects of different covariates was not always consistent. For example, in general, men had a higher odds of dying in a hospital; however, this was not true for APR-DRG 045—*Cerebral Vascular Accident and Precerebral Occlusion with infarct*, 174—*Percutaneous Cardiovascular Procedures With Ami* and 279—*Hepatic Coma and Other Major Acute Liver Disorders*. Similarly, patients originating from another care setting usually have increased odds of dying. The odds of dying increased when a patient originated from a nursing home or was transferred from another hospital versus a patient coming from home. However, in two APR-DRGs, 281 *Malignancy Of Hepatobiliary System and Pancreas* and 308 *Hip and Femur Procedures For Trauma, Except Joint Replacement*, origin of the nursing home was protective. For APR-DRG 130 *Respiratory System Diagnosis W Ventilator Support 96+Hours*, transfer from another hospital showed a statistically significant lower odd of dying.

Finally, the percentage of hospitals with SMRs per APR-DRG signalling deviation from the benchmark was analysed (figure 2). All APR-DRGs with both a high crude mortality and high percentage of hospitals with SMRs signalling deviation from the benchmark involved malignancies. In addition, large variation in palliative coding between the hospitals was observed for these APR-DRGs (online supplementary table 3). The IQR for the percentage of palliative patients exceeded 20% for the following APR-DRGs (figure 2): 382—*Malignant Breast Disorders* (IQR=25%, Max – Min=52%), 136—*Respiratory Malignancy* (IQR=23%, Max – Min=58%), 530—*Female Reproductive System Malignancy* (IQR=23%, Max – Min=49%), 500—*Malignancy, Male Reproductive System* (IQR=21%, Max – Min=45%) and 041—*Nervous System Malignancy* (IQR=20%, Max – Min=50%). The APR-DRGs with the highest percentage of hospitals with SMRs signalling deviation from the benchmark, but not involving malignancies, were 021—*Craniotomy Except for Trauma*, 190—*Acute Myocardial Infarction* and 194—*Heart Failure*, 042—*Degenerative Nervous system disorders Exclusive Multiple Sclerosis*, 139—*Other pneumonia*, 380—*Skin Ulcers and 308—Hip and Femur Procedures for Trauma, Except Joint Replacement* (figure 2). All SMR models for these APR-DRGs had C-statistics of >0.70, except for 194—*Heart failure* (C-statistic=0.67).

Use of an online platform

For SMR reporting, the developed R/Shiny application was embedded in the Flemish hospital network website from December 2014 onwards. Dissemination of the

Table 1 Description of explanatory covariates and mortality

	2009	2010	2011	2012	2013	2014	2016	7 years
n (% of admissions)								
Admissions	155 086	167 483	171 343	174 455	176 242	177 376	176 732	1 198 717
Deaths	10 350 (6.67%)	10 948 (6.54%)	10 839 (6.33%)	11 321 (6.49%)	11 145 (6.32%)	10 638 (6.00%)	10 416 (5.89%)	75 657 (6.31%)
Crude mortality rate								
Gender								
M	7.2%	7.0%	6.9%	6.9%	6.8%	6.5%	6.3%	6.8%
F	6.2%	6.1%	5.8%	6.1%	5.9%	5.4%	5.5%	5.8%
Age								
(0,10)	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
(10,20)	0.3%	0.2%	0.2%	0.2%	0.2%	0.3%	0.2%	0.2%
(20,30)	0.6%	0.7%	0.6%	0.6%	0.8%	0.5%	0.5%	0.6%
(30,40)	1.2%	1.2%	0.9%	0.9%	1.0%	0.9%	0.9%	1.0%
(40,50)	2.0%	1.9%	1.7%	2.1%	1.9%	1.8%	1.8%	1.9%
(50,60)	4.3%	4.0%	3.9%	3.9%	3.8%	3.3%	3.3%	3.8%
(60,70)	5.8%	5.6%	5.5%	5.5%	5.5%	5.1%	4.9%	5.4%
(70,80)	7.9%	7.6%	7.5%	7.5%	7.1%	6.8%	6.4%	7.3%
(80,90)	12.2%	11.7%	10.9%	10.8%	10.4%	10.0%	9.6%	10.8%
(90,.)	17.4%	16.8%	16.1%	16.6%	16.1%	14.7%	14.7%	15.9%
Admission source								
Home	5.9%	5.8%	5.7%	5.8%	5.6%	5.3%	5.1%	5.6%
Nursing home	16.7%	15.9%	15.4%	16.1%	15.9%	15.2%	15.8%	15.8%
Other	7.4%	7.9%	5.6%	5.7%	7.6%	6.6%	6.0%	6.6%
Other hospital	12.4%	12.0%	10.7%	11.1%	11.1%	11.5%	12.2%	11.6%
Public space	5.2%	5.1%	4.7%	4.3%	4.5%	4.1%	3.7%	4.5%
Comorbidity score								
(0,1)	4.3%	4.2%	3.9%	4.0%	3.8%	3.4%	3.3%	3.9%
(1,2)	18.7%	18.2%	17.7%	17.3%	16.9%	16.0%	15.6%	17.1%
(2,3)	30.3%	28.4%	29.3%	29.3%	27.4%	27.9%	26.9%	28.3%
(3,.)	40.0%	38.8%	40.2%	41.1%	43.8%	39.6%	38.7%	40.1%
Admission type								
Emergency	8.4%	8.2%	7.8%	8.1%	7.7%	7.3%	7.2%	7.8%
Elective	3.3%	3.2%	3.3%	3.2%	3.3%	3.2%	3.1%	3.2%

Crude mortality in each demographic subgroup of the included patient population is depicted as percentage of the total number of included admissions for a specific discharge year or overall. M=male, F=female.

benchmark to the member hospitals was phased. In the first instance, access was given only to the CMOs of the participating hospitals and hospital identification numbers were anonymous to build a trusted environment. In order to validate the HSMR model, each of the participating CMOs was asked to examine SMRs for their respective hospital guided by an analytical scheme. The analytical scheme provides a stepped approach to analyse SMR: (1) select a patient group with a high percentage of hospitals signalling deviation from the benchmark

within the Flemish hospital network (figure 2); (2) check whether the set of ICD9/10 CM codes for all patients reflects the profile of the patients; (3) examine whether structure or process parameters could have contributed to a deficit in quality of care; and (4) perform root cause analysis at individual patient level for deaths during hospitalisation with a low risk of mortality. CMOs subsequently discuss findings in group meetings organised by the Flemish hospital network (figure 3). In the next phase, improvement strategies and the organisation of mortality

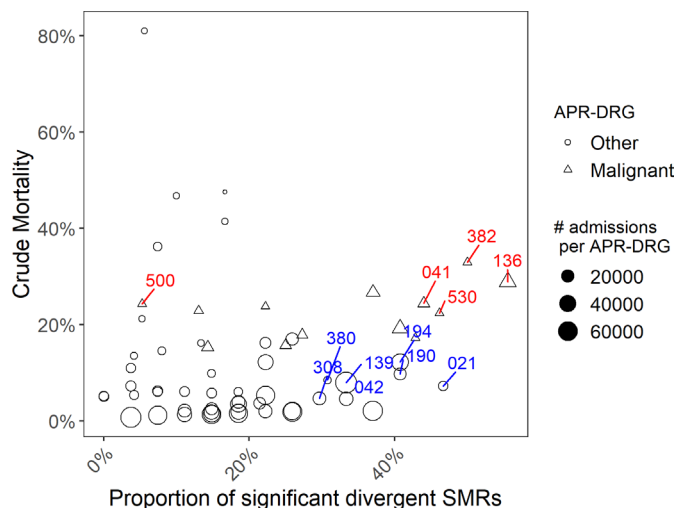


Figure 2 Crude mortality. The proportion of hospitals standardised mortality rate (SMRs) per All Patient Refined—Diagnosis-Related Groups (APR-DRG) signalling deviation from the benchmark is depicted on the X-axis. The Y-axis depicts the crude mortality. The size of the shapes indicates the number of included stays for that APR-DRG. APR-DRGs with malignancies are depicted with triangles. APR-DRGs highlighted in red have the highest proportion of palliative patients. APR-DRGs highlighted in blue have the highest proportion of significant signalling SMRs, excluding APR-DRGs entailing malignancies.

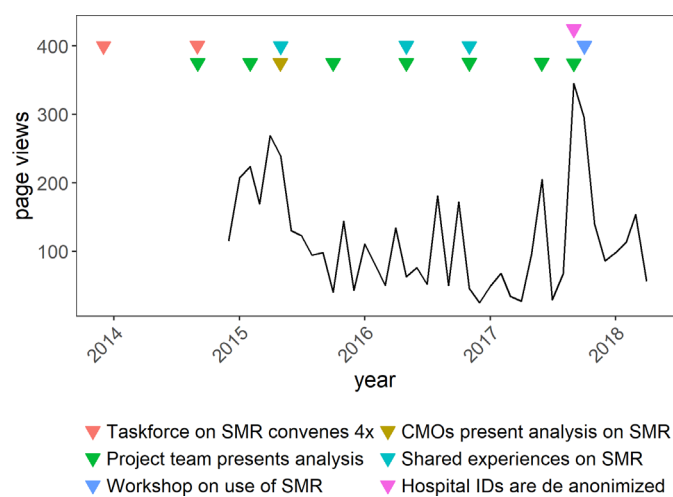


Figure 3 Development track and use of the web application. The development and use of the standardised mortality rate (SMR) benchmark are depicted. The trend in page views of the web application is depicted with a line diagram. Specific events in the hospital network are indicated with triangles. (Red) A taskforce consisting of the medical coding experts and chief medical officers (CMOs) of six member hospitals decided on inclusion and exclusion criteria. (Orange) Hospitals analysed their SMRs and presented the conclusions. (Light blue) Experiences and learnings are shared during quarterly medical coding experts meetings and quarterly CMO meetings. (Green) The project group provided regular updates on the models and web application in the CMO meetings. (Blue) A hands on workshop on the web application was organised. (Purple) The web application was deanonymised with regard to hospital ID.

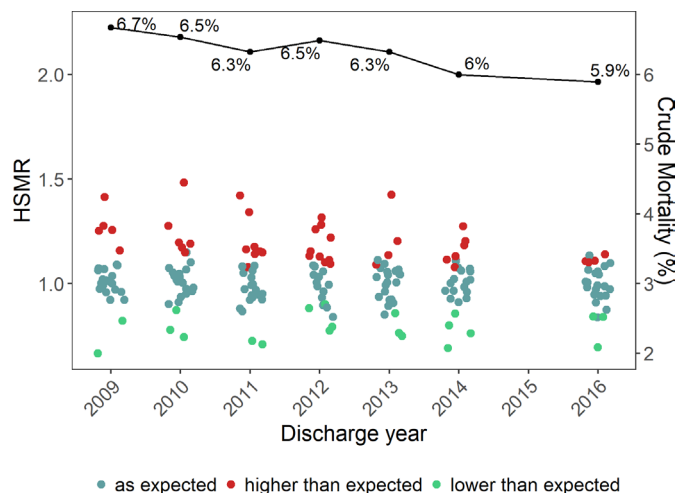


Figure 4 Hospital standardised mortality ratio (HSMR) variation over time. The HSMR for each hospital aggregated by discharge year is depicted as a dot. Blue dots indicate HSMR within 95% CI, red dots indicate HSMR higher than upper limit of 95% CI and green dots indicate HSMR lower than lower limit of 95% CI. In addition, the crude mortality within selected All Patient Refined—Diagnosis-Related Groups in the Flemish hospital network is shown as a line diagram.

and morbidity rounds^{27–29} were shared within the CMO meetings. This involved both seminars on quality topics as well as testimonials from CMOs on the actions taken within their hospital. In addition, analytical and medical coding experts quarterly discuss the usability and technical aspects of the web application. In the last phase, when hospitals were experienced with the use of the information, the application was deanonymised with regard to hospital name.

The use of the web application has increased moderately over time, but it is overly clear that the increased use of the web application coincided with specific events like the presentation by the CMOs of the analysed SMR, the workshops on the use of the application and the deanonymisation (with regard to the hospital name) of the web application (figure 3). In addition, within working groups, peers reported that the web application was frequently used for preparing reports for CMOs and boards of directors.

Finally, we observed that crude mortality in the Flemish hospital network decreased from 6.7% to 5.9% for the selected APR-DRGs between 2009 and 2016 (table 1). Although variation in HSMR appeared to decrease over time, no heterogeneity in variance could be identified ($p=0.60$) (figure 4) nor could a significant effect of discharge year on HSMR be demonstrated. Discharge year 2016 did however show an OR significantly lower than 1 (online supplementary table 2) in 25 SMRs.

DISCUSSION

We report on the development of an SMR model and the use of a framework to benchmark results, share analyses

and initiate quality improvement initiatives. We experienced that our approach incentivises critical appraisal of hospital mortality by the CMOs and enables various quality improvement initiatives in the member hospitals. We note that in using the tool as a smoke signal and further analysing processes for those patient groups which got an alert, healthcare practitioners gather around the 'patient process' and establish a path towards better care.

We developed an SMR model for APR-DRGs which comprises 80% of the Flemish hospital network mortality. The design of the presented modelling strategy is similar to other public reported SMR models and generally the same set of independent variables is used.^{30–32} The use of an arbitrary cut-off at 80% of all pooled mortality limits the number of included APR-DRGs. This may cause selection bias in aggregated analyses (over all APR-DRG) at the level of the hospital (HSMR), but in analyses at the level of APR-DRG this is not relevant. In some countries, information on socioeconomic deprivation is also included.^{30–32} Such information is not collected by the participating hospitals but we acknowledge that the addition of information about socioeconomic context can be valuable.³³ Some authors add an independent variable for admission season,^{30 32} which could be linked to the comorbidity level (eg, seasonal influenza). This may however also mask variability in quality of care due to seasonal variation in occupancy rates in Belgian hospitals.³⁴ As expected, odds of dying in hospitals increased with increasing age.^{1 35} Patients admitted via the emergency room and patients with higher comorbidity showing higher odds of dying for most APR-DRGs was as was shown in previous studies.^{1 35} For factors of source of admission, results were less clear; origin of the nursing home being protective for two APR-DRGs may indicate that other factors are at play, than just patient frailty. Such factors may include the organisation of the healthcare system, geographical and demographical factors.

The choice of grouping software could have an effect on modelling as the degree of homogeneity of the patient groups can have an impact on case mix adjustment. Many authors use Clinical Classification Software to group patients, which is solely based on the principal diagnosis.^{30 36} In this study, the APR-DRG grouping system was used which also takes into account information on procedures to assign patients to groups. This grouping system is a familiar format used for benchmarking based on HDDs in Belgium. It is however not clear which grouping system is better for comparing hospital mortality.³⁷ Further analysis should be done to demonstrate how grouping systems can affect SMR results. In addition, registration bias in the diagnosis codes and covariates may lead to the misclassification of patients to pathology groups and the misclassification of predictors.³⁸ The current study did not evaluate the existence of registration bias in the discharge datasets. However, in the stepped approach to analyse SMR within the *Flemish hospital network*, we included checks for coding bias. Although hospital mortality is a clear and tractable measure, the use of HSMR methodology for hospital

comparison has been subject to debate for a number of reasons.^{11–13 39} First, in-hospital mortality is only relevant to a limited number of pathologies. For instance, in our dataset, only 30% of the hospital stays were included in the analysis. Consequently, in-hospital mortality must not be used as a single measure to compare overall hospital quality of care. In addition, incomparability of HSMR arising from issues such as insufficient adjustment of disease severity, referral bias, disparity in end-of-life care, unmeasured case mix variation, variation in coding and other specific artefacts of data collection or analysis has been described.^{12 40–42} Manktelow *et al* showed that even with optimal risk adjustment the comparison of two SMRs can be perturbed by the Yule-Simpson effect that may result from differences in case-mix and volume of patients.¹³ Several authors reported the lack of a correlation between avoidable deaths and SMR.^{11 43} Taken together, these reports underline the necessity of alerting the user for misinterpretation or misuse of SMR. It is clear that the SMR models will need continuous refinement to improve patient-mix correction and limit additional biases. Studies in the UK and the Netherlands pointed to the importance of including postdischarge deaths to avoid bias related to differences in discharge policies.^{9 44} It has also been shown that an adjustment for the frequency of readmissions should be considered, since such a model showed more favourable quality metric characteristics compared with a model without such adjustment.⁴⁵ Moreover, in collaboration with clinicians of the hospital, additional clinical risk-adjustment variables can be identified which are very likely to increase the chance of identifying failing processes that explain variation in mortality.^{43 46} It is clear that prespecifying a model based on clinical input is a better approach than the automated backward variable deletion methodology used in this study which has been shown to be unstable with regard to predictor selection.⁴⁷ In addition, the use of clinical insight could help to identify unaccounted predictors and to avoid residual confounding.

As an alternative to the models described based on administrative HDDs, recent advances in deep learning methods with electronic health records hold promise to predict mortality with high accuracy.⁴⁸ The downside is that these techniques may be challenging to interpret clinically.

Even with more refined models, the appropriate use of HSMR methodology raises another debate. Public reporting of SMR data makes it intuitively attractive to compare hospitals in a competitive ranking. The goal of the mortality benchmark within the *Flemish hospital network* is not to create a competitive benchmark but aims at enabling quality improvement initiatives for specific patient groups. As illustrated in this study, we propose a methodology to select patient groups with high variation of SMR. APR-DRGs with malignancies suffer from large variation in palliative coding, making them less appropriate for SMR analysis. The presented reporting platform of the Flemish hospital network and the collaborative

approach where hospital management and CMOs have a forum to discuss findings and experiences give the Flemish hospital network the opportunity to largely overcome the weaknesses of SMR methodology and exploit its strengths. In addition, the *Flemish hospital network* is continuously working on gathering the data needed to include postdischarge deaths in order to investigate the effect of discharge policies on SMR in the Flemish hospital network hospitals. Further quantitative and qualitative investigation will look into the causes of mortality through the investigation of structure and process variables. It will prove interesting to further investigate this at the hospital level, at the pathology level and the individual patient level. We already established benchmarks in the web application on length of stay and healthcare personnel staffing levels. The relation between these general process measures and outcome measures like mortality and readmission has been previously reported.¹⁹ Within care-program groups, the *Flemish hospital network* is striving to design interventions and implement improvement strategies to reduce mortality in selected APR-DRGs identified by focussing on patient groups with high variation of SMR and high crude mortality. One such approach are mortality and morbidity rounds^{27–29} which were initiated in multiple member hospitals as a consequence of SMR reporting.

The current study reported on the development of a user-friendly web application accessible to clinicians of the Flemish hospital network. The use of the web application increased over time; however, the measurement thereof was limited to the number of page views. Future studies could develop surveys or other metrics to monitor the use of the web application more thoroughly.

CONCLUSION

The *Flemish hospital network* successfully used a phased approach for introducing SMR which enabled mutual trust among the members and incentivises quality improvement. The presented framework for mortality measure reporting and the phased collaborative approach seems to serve as a stepping stone for quality improvement initiatives. Key components of this approach were technical support in creating hospital benchmarks and providing support for interpretation, establishment of a platform to anonymously disseminate the results of hospital benchmarks and sharing analyses on SMR and steps taken to improve. Further initiatives should be pathology specific and entail the refinement of the SMR model (if relevant) with the inclusion of disease-specific risk factors, development of other relevant quality measures and concrete interventions.^{45 48} The combination thereof could facilitate an increase in the quality of care in a hospital network which could lead to a decrease in hospital mortality.

Author affiliations

¹University Hospitals Leuven, Leuven, Belgium

²Management Information and Reporting, University Hospitals Leuven, Leuven, Belgium

³Department of Radiation Oncology, University Hospitals Leuven, Leuven, Belgium

⁴Department of Urology, University Hospitals Leuven, Leuven, Belgium

⁵Leuven Institute for Healthcare Policy, KU Leuven, Leuven, Belgium

⁶Department of Quality Management, University Hospitals Leuven, Leuven, Belgium

Acknowledgements The authors would like to thank the Flemish hospital network members for the long lasting collaboration, Jonas Tundo (former junior member of MIR UZ Leuven), Paul Aylin and Alex Bottle (Imperial College London), Robert Douce and Steven Middleton (Dr Foster, London) for advice regarding standardised mortality modelling.

Contributors WT, GVB, DDR, FR and CW designed the study. PS, LB and WT wrote the manuscript. PS, GVB and PM programmed the algorithms for fitting the models. PM and PS programmed the web application. DDR, KV, FR and CW reviewed the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Approval by the University Hospitals Leuven ethics committee was obtained (S61450).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- Jarman B, Gault S, Alves B, *et al*. Explaining differences in English Hospital death rates using routinely collected data. *BMJ* 1999;318:1515–20.
- Heijink R, Koolman X, Pieter D, *et al*. Measuring and explaining mortality in Dutch hospitals; the hospital standardized mortality rate between 2003 and 2005. *BMC Health Serv Res* 2008;8:73.
- Brien SE, Ghali WA. Public reporting of the hospital standardized mortality ratio (HSMR): implications for the Canadian approach to safety and quality in health care. *Open Med* 2008;2.
- Fekri O, Amuah JE, Herasimovich V, *et al*. Palliative care coding practices in Canada since the introduction of guidelines and the HSMR indicator. *BMJ Open* 2015;5:e008753.
- Shinjo D, Fushimi K. The degree of severity and trends in hospital standardized mortality ratios in Japan between 2008 and 2012: a retrospective observational study. *Int J Qual Heal Care* 2017;29:705–12.
- NHS Scotland. *Hospital standardised mortality ratio quarterly release*. 39, 2017.
- Jacques RM, Fotheringham J, Campbell MJ, *et al*. Did hospital mortality in England change from 2005 to 2010? A retrospective cohort analysis. *BMC Health Serv Res* 2013;13:216.
- Chong CA, Nguyen GC, Wilcox ME. Trends in Canadian Hospital standardised mortality ratios and palliative care coding 2004–2010: a retrospective database analysis. *BMJ Open* 2012;2:e001729.
- Bottle A, Jarman B, Aylin P. Hospital standardized mortality ratios: sensitivity analyses on the impact of coding. *Health Serv Res* 2011;46:1741–61.
- Julious SA, George S. Are Hospital League tables calculated correctly? *Public Health* 2007;121:902–4.
- Girling AJ, Hofer TP, Wu J, *et al*. Case-Mix adjusted hospital mortality is a poor proxy for preventable mortality: a modelling study. *BMJ Qual Saf* 2012;21:1052–6.
- Mohammed MA, Deeks JJ, Girling A, *et al*. Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. *BMJ* 2009;338:780.
- Manktelow BN, Evans TA, Draper ES. Differences in case-mix can influence the comparison of standardised mortality ratios even with

- optimal risk adjustment: an analysis of data from paediatric intensive care. *BMJ Qual Saf* 2014;23:782–8.
14. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc* 1984;79:516–24.
 15. Lee WC. Properties of the geometrically averaged ratio, an alternative standardized measure. *Epidemiology* 1999;10:456–9.
 16. Shahian DM, Normand S-LT. Comparison of "risk-adjusted" hospital outcomes. *Circulation* 2008;117:1955–63.
 17. Ghaferi AA, Birkmeyer JD, Dimick JB. Variation in hospital mortality associated with inpatient surgery. *N Engl J Med* 2009;361:1368–75.
 18. Pearse RM, Moreno RP, Bauer P, et al. Mortality after surgery in Europe: a 7 day cohort study. *Lancet* 2012;380:1059–65.
 19. Aiken LH, Sloane DM, Bruyneel L, et al. Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study. *Lancet* 2014;383:1824–30.
 20. van Walraven C, Austin PC, Jennings A, et al. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Med Care* 2009;47:626–33.
 21. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8–27.
 22. Moore BJ, White S, Washington R, et al. Identifying increased risk of readmission and in-hospital mortality using Hospital administrative data: the AHRQ Elixhauser comorbidity index. *Med Care* 2017;55:698–705.
 23. Steyerberg EW, Harrell FE, Borsboom GJ, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 2001;54:774–81.
 24. Badgeley MA, Shameer K, Glicksberg BS, et al. EHDViz: clinical dashboard development using open-source technologies. *BMJ Open* 2016;6:e010579–12.
 25. Chang W, Joe Cheng JJ, Allaire YX. *Shiny: web application framework for R*, 2014.
 26. R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2015.
 27. Kobewka DM, van Walraven C, Turnbull J, et al. Quality gaps identified through mortality review. *BMJ Qual Saf* 2017;26:141–9.
 28. Higginson J, Walters R, Fulop N. Mortality and morbidity meetings: an untapped resource for improving the governance of patient safety? *BMJ Qual Saf* 2012;21:576–85.
 29. Benassi P, MacGillivray L, Silver I, et al. The role of morbidity and mortality rounds in medical education: a scoping review. *Med Educ* 2017;51:469–79.
 31. Canadian Institute for Health Information. *Hospital standardized mortality ratio (HSMR) about HSMR*, 2014: 1–20.
 32. ISD Scotland. *Scottish Hospital Standardised Mortality Ratio (HSMR): Methodology & Specification Document*, 2016: 1–14.
 33. Foster HME, Celis-Morales CA, Nicholl BI, et al. The effect of socioeconomic deprivation on the association between an extended measurement of unhealthy lifestyle factors and health outcomes: a prospective analysis of the UK Biobank cohort. *Lancet Public Health* 2018;3:e576–85.
 34. Schilling PL, Campbell DA, Englesbe MJ, et al. A comparison of in-hospital mortality risk conferred by high Hospital occupancy, differences in nurse staffing levels, weekend admission, and seasonal influenza. *Med Care* 2010;48:224–32.
 35. Jarman B, Pieter D, van der Veen AA, et al. The hospital standardised mortality ratio: a powerful tool for Dutch hospitals to assess their quality of care? *Qual Saf Health Care* 2010;19:9–13.
 36. Agency for Healthcare Research and Quality. *Beta clinical classifications software (CCS) for ICD-10-CM/PCS*. Healthcare Cost and Utilization Project (HCUP), 2018.
 37. De Bleser L, Vlayen J, Vanhaecht K, et al. Classifying clinical pathways. *Stud Health Technol Inform* 2004;110:9–14.
 38. Manuel DG, Rosella LC, Stukel TA. Importance of accurately identifying disease in studies using electronic health records. *BMJ* 2010;341:4226.
 39. van Erven JA, van Galen LS, Hettinga-Roest AA, et al. Hospital standardised mortality ratio: a reliable indicator of quality of care? *Neth J Med* 2018;76:72–7.
 40. van Gestel YRBM, Lemmens VEPP, Lingsma HF, et al. The hospital standardized mortality ratio fallacy: a narrative review. *Med Care* 2012;50:662–7.
 41. Goodacre S, Campbell M, Carter A. What do hospital mortality rates tell us about quality of care? *Emerg Med J* 2015;32:244–7.
 42. Pitches DW, Mohammed MA, Lilford RJ. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. *BMC Health Serv Res* 2007;7:91.
 43. Hogan H, Zipfel R, Neuburger J, et al. Avoidability of hospital deaths and association with hospital-wide mortality ratios: retrospective case record review and regression analysis. *BMJ* 2015;351.
 44. Pouw ME, Peelen LM, Moons KGM, et al. Including post-discharge mortality in calculation of hospital standardised mortality ratios: retrospective analysis of hospital episode statistics. *BMJ* 2013;347:5913.
 45. van den Bosch WF, Spreeuwenberg P, Wagner C. Variations in hospital standardised mortality ratios (HSMR) as a result of frequent readmissions. *BMC Health Serv Res* 2012;12:91.
 46. Siregar S, Pouw ME, Moons KGM, et al. The Dutch Hospital standardised mortality ratio (HSMR) method and cardiac surgery: benchmarking in a national cohort using Hospital administration data versus a clinical database. *Heart* 2014;100:702–10.
 47. Austin PC, Tu JV. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. *J Clin Epidemiol* 2004;57:1138–46.
 48. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med* 2018;1.
 49. Pouw M. *Measuring standardised mortality ratios of hospitals: challenges and recommendations*, 2016.