Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Does a Code for Acute Myocardial Infarction Mean the Same in All Norwegian Hospitals? A Likelihood Approach to a Medical Record Review

Jon Helgeland, Doris Tove Kristoffersen, Katrine Damgaard Skyrud

Division for Health Services, Norwegian Institute of Public Health, Oslo, Norway

Correspondence: Jon Helgeland, Norwegian Institute of Public Health, PO Box 222 Skøyen, Oslo, 0213, Norway, Tel +47 464 00 443, Email Jon.Helgeland@fhi.no

Objective: Health registries are important data sources for epidemiology, quality monitoring, and improvement. Acute myocardial infarction (AMI) is a common, serious condition. Little is known about variation in the positive predictive value (PPV) of a coded AMI diagnosis and its association with hospital quality indicators. The present study aimed to investigate the relationship between PPV and registry-based 30-day mortality after AMI admission and between-hospital variation in PPV.

Study Design and Setting: An electronic record review was performed in a nationwide sample of Norwegian hospitals. Clinical signs and cardiac troponin measurements were abstracted and analyzed using a mixture model for likelihood ratios and parametric bootstrapping.

Results: The overall PPV was estimated to be 97%. We found no statistically significant association between hospital PPV and the classification of hospitals into low, intermediate, and high registry-based 30-day mortality. There was significant variation between hospitals, with a PPV range of 91–100%.

Conclusion: We found no evidence that variation in PPV of AMI diagnosis can explain variation between hospitals in registry-based 30-day mortality after admission. However, PPV varied significantly between hospitals. We were able to use a very efficient statistical approach to the analysis and handling of various sources of uncertainty.

Keywords: health registries, quality indicators, finite mixture models, case fatality, cardiac troponins

Introduction

Acute myocardial infarction (AMI) is a serious condition with high short-term mortality and a high rate of subsequent disability among survivors. Hospital administrative databases are important sources for the study of AMI epidemiology as well as for quality monitoring and improvement. The probability of death within 30 days after hospital admission (here denoted 30D), based on coded diagnosis in electronic registries, has been used as a quality indicator for hospitals. Routine publication of quality indicators can potentially be used for quality improvement. To be useful, the indicators must be valid in the sense of having negligible bias and unobserved confounding. We may distinguish between two sources of bias when comparing hospitals: first, variation in diagnostic or coding practice, resulting in different medical conditions in the data from different hospitals; second, variation in disease severity, which cannot be controlled for by case mix adjustment. Although 30D is in routine use as a quality indicator in Norway and elsewhere, criticisms have been raised.[1–8] Addressing these issues is important to further the use of 30D in health system governance and clinical quality improvement.

The diagnosis of AMI rests on patient history, clinical information such as electrocardiography (ECG) abnormalities and the presence of chest pain, as well as biochemical markers. Cardiac troponins (cTns) are proteins that form parts of the heart muscle tissue (myocardium) and are released when heart tissue is damaged. For an overview of the diagnosis of AMI and the use of cTn, see e.g.[9,10] A loss of oxygen supply to the heart (ischemia) is necessary for AMI diagnosis. Type 2 AMI is defined as ischemia caused by acute conditions other than acute coronary atherothrombosis. Nonischemic myocardial injury may

Received: 30 April 2022
Accepted: 24 September 2022
Published: 13 October 2022

result in elevated cTn and must be distinguished from Type 2 AMI.[11,12] The proportion of Type 2 among AMI cases varies widely between studies.[13]

The positive predictive value (PPV) of a coded AMI diagnosis from administrative health registries has been studied by several authors.[14–20] Two systematic reviews[21,22] reported PPVs greater than 93% in the majority of studies and a range of 70–100%. We know of only one study on the variation in PPV.[23] For definite or probable AMI, the authors found significant differences in PPV between four communities of residence, ranging from 71% to 78%.

Many validation and PPV studies rely on establishing a gold standard by expert judgment and consensus. This introduces an element of subjectivity and lack of transparency and reproducibility. Arguably, the interpretation of clinical laboratory data in isolation is primarily a statistical decision problem. Classification rules based on likelihood ratios (LRs) are known to have strong optimality properties.[24] Likelihood data for categorical clinical signs are easily derived from data on sensitivity and specificity.

Two types of cardiac troponin are used for the routine evaluation of suspected AMI: troponin T (cTnT) and troponin I (cTnI). Additionally, several assays are used for the analysis of troponins in clinical care. There is variation between types and assays, and there is a lack of standardization that makes direct comparison difficult.[25] In a statistical model for cTn, the null distribution, that is, the probability distribution of cTn measurements taken from a reference population known not to have AMI, must be known.

Study Objectives

The background for the present study is a validation project aimed at investigating the properties of 30-day mortality after AMI, based on coded diagnosis in electronic registries (30D), as a quality indicator for hospitals. A set of clinical observations as well as cTn measurements were obtained from record review in a sample of hospitals. No similarly detailed data from a known healthy or other non-AMI population were available. We were interested in the variation in diagnostic and coding practice, particularly the PPV of an AMI diagnosis, as recorded in hospitals' administrative systems.

The aims of the present paper are to

1. Study variation between hospitals in PPV.
2. Investigate whether differences in hospital registry-based 30D can be explained by differences in PPV.
3. Describe a statistical approach for analyzing medical records with discrete, continuous, and correlated serial data types, as well as uncertainties in external data.

Note that we do not intend to study any causal relationship or other association between the PPV in a hospital and the true mortality of a patient with AMI admitted to that hospital. These are, at least in principle, two independent quantities. Additionally, we do not a priori make assumptions about the accuracy of 30D as an estimate of the true mortality.

Materials and Methods

Patient administrative data from all Norwegian general hospitals for admissions in the period 2002–2009 were retrieved using an in-house system. Records were matched across hospitals using the unique Norwegian person identification number and date of death retrieved from the National Registry, where applicable. The underlying population consists of patients with an emergency hospital admission and ICD-10 categories I21.0–9 as the primary or secondary diagnosis. The risk-adjusted 30-day in-and-out-of-hospital mortality probability per hospital was computed as described elsewhere.[26] The hospitals were eventually classified as low, intermediate, or high 30D hospitals. These categories are denoted L30D, M30D, and H30D, respectively. In the first stage, hospitals were drawn randomly in these categories, constrained by requiring balance between the Norwegian hospital regions and between hospital types: local and intermediate size hospitals as well as regional referral hospitals. Only general hospitals with at least 250 cases in the period were included. Within each hospital, patients with an AMI admission in the period 2007–2009 were sampled within patient strata, defined by admission year and case severity in three levels: (i) survived 30 days after admission, with length of stay not exceeding the median; (ii) survived 30 days with hospital stay greater than the median; and (iii) died within 30 days. For patients with more than one AMI admission, the most recent was chosen. The objective of the stratified sampling of

patients was to maximize statistical power, as PPV could be expected to vary with case severity. The relative stratum sample sizes were chosen accordingly. To estimate the overall PPV for the total population of AMI cases, estimates were first computed for each patient stratum separately and aggregated using a weighted mean, with the proportions of the strata in the population as weights. Records were reviewed following a detailed manual and using structured entry forms by experienced medical record abstracters. They were given initial training and were followed up throughout the process. Each record required approximately one hour of abstraction time. The presence or absence of clinical signs and examination results, as well as laboratory results, were registered as "present/not present" if mention was found in the records and as "not recorded" otherwise. Records for transferred patients were reviewed at the initially admitting hospital. Further details of patient sampling and medical record abstraction have been described previously.[27]

A small double abstraction substudy was performed, where a reduced item set was recorded by an abstractor not otherwise engaged in the AMI study. Fifty records were drawn randomly from a single hospital.

Observation Model

Three categories were used for clinical signs: recorded as present, recorded as absent, and not recorded.

Let X, Y denote observed and true indicator variables, respectively, where Y takes the values 0 (denoting sign absent) and 1 (denoting present) and X takes the additional value ω (denoting sign not recorded). We assume that a sign actually present (or not) had a probability q_1 (q_0) of being recorded. Let I=1,2 denote the underlying reference non-AMI and AMI populations, respectively. We assume that

$$\Pr(Y = 1 \mid I = i) = p_i,$$
$$\Pr(X = x \mid Y = y, I = i) = \begin{cases} q_y & \text{if} x = y, \\ 1 - q_y & \text{if} x = \omega. \end{cases}$$

Consequently,

$$\Pr(X = x \mid Y = y, I = i) = \begin{cases} p_i q_1 & x = 1, \\ (1 - p_i) q_0 & x = 0, \\ p_i(1 - q_1) + (1 - p_i)(1 - q_0) & x = \omega. \end{cases}$$

For the likelihood ratio $L(x) = \Pr(X = x \mid I = 1)/\Pr(X = x \mid I = 2)$, we find

$$L(1) = \frac{p_1}{p_2},$$
$$L(0) = \frac{1-p_1}{1-p_2},$$
$$L(\omega) = \frac{p_1(1-q_1)+(1-p_1)(1-q_0)}{p_2(1-q_1)+(1-p_2)(1-q_0)}.$$

To estimate q_0, q_1 from the data, we assume that they are constant across questionnaire items (signs). With a slight change of notation, let Y_j, p_{1j}, p_{2j} denote indicator variable and probabilities for item j, N the number of questionnaires, and π_i the probability of population i. With Z_j denoting the number of ones and W_j denoting the number of zeroes for item j, we see that $Z_j \sim bin(n, q_1 a_j)$, $W_j \sim bin(n, q_0 b_j)$ where

$$a_j = \Pr(Y_j = 1) = \pi_1 p_{j1} + (1 - \pi_1) p_{j2},$$
$$b_j = \Pr(Y_j = 0) = \pi_1(1 - p_{j1}) + (1 - \pi_1)(1 - p_{j2}).$$

Assuming a value for π_1, reasonable estimators for q_0, q_1 are

$$\hat{q}_0 = \sum_j b_j W_j / N \sum_j b_j, \ \hat{q}_1 = \sum_j a_j Z_j / N \sum_j a_j. \tag{1}$$

Alternatively, if we assume $q_1 = q_0 = q$, we see that the overall number of zeroes or ones has the binomial distribution with probability parameter q, with the natural estimator

$$\hat{q} = \frac{1}{N} \sum_j (Z_j + W_j).$$

Mixture Model

We assume that the data contain an unknown proportion of AMI patients, denoted population 2, as well as non-AMI patients, denoted population 1. The statistical analysis framework is thus that of finite mixture modeling.[28]

The data from patient k are denoted X_k. We assume that our observations are independent and identically distributed with a two-component mixture distribution and that the density of X_k is

$$g_k = \pi_1 f_k^1 + (1-\pi_1) f_k^2, k = 1, \ldots, N,$$

where f_k^i is the density of X_k when patient k belongs to population i=1,2 and N is the total number of patients. The proportion of AMI patients $\pi_2 = 1 - \pi_1$ is the PPV of the AMI diagnosis. We will initially assume that the component densities are known and contain no unknown parameters. The log-likelihood for the complete dataset becomes

$$L(\pi_1) = \sum_k \log\left(\pi_1 f_k^1 + (1-\pi_1) f_k^2\right) = const + \sum_k \log(1 + \pi_1(\Lambda_k - 1)),$$

where $\Lambda_k = f_k^1 / f_k^2$ is the likelihood ratio for the k-th patient. LRs from different signs and variables were combined by multiplication.

For the case where the population is cross-classified by two factors, the model was extended. With a slight change of notation, we assume that the mixing probability is a constant π_{ij} for each combination of levels i, j of the first and second factors and that

$$\log\left(\pi_{ij} / \left(1 - \pi_{ij}\right)\right) = \mu + \alpha_i + \beta_j,\ i = 1, \ldots, I,\ j = 1, \ldots, J, \tag{2}$$

where for identifiability, $\alpha_1 = \beta_1 = 0$. To account for the stratified sampling plan, we used patient stratum as the first factor.

Likelihood ratios for clinical signs were taken from a systematic review,[29] except for chest pain and coronary artery stenoses. For those, we used the DerSimonian–Laird method[30] to derive LRs and confidence intervals based on data in the literature.[31–39] The LRs used in the analysis are shown in Table 1.

Table 1 Likelihood Ratios for Clinical Signs and Findings, with 95% Confidence Intervals

Sign	Sign Present	Sign Absent
Chest pain	1.7 (1.4, 2.1)	0.43 (0.33, 0.56)
Nausea	1.7 (1.3, 2.3)	0.80 (0.70, 0.90)
Diaphoresis	2.1 (1.8, 2.5)	0.70 (0.60, 0.80)
Systolic blood pressure <100 mm Hg	3.0 (2.0, 6.5)	0.96 (0.90, 1.00)
ST elevation	22.0 (16.0, 30.0)	0.60 (0.60, 0.60)
ST depression	4.5 (3.6, 5.6)	0.80 (0.70, 0.90)
T-wave inversion	2.2 (1.8, 2.6)	0.90 (0.80, 1.00)
New Q wave	22.0 (7.6, 62.0)	0.80 (0.80, 0.90)
Bundle branch block	2.4 (0.4, 15.0)	1.00 (0.80, 1.10)
Normal ECG	0.2 (0.1, 0.3)	1.50 (1.40, 1.60)
Angina	1.2 (0.9, 1.8)	0.90 (0.80, 1.10)
Prior myocardial infarction	1.3 (1.0, 1.8)	0.90 (0.80, 1.00)
Heart failure	0.7 (0.6, 0.9)	1.10 (1.00, 1.20)
Coronary artery stenoses	7.1 (3.6, 14.0)	0.07 (0.03, 0.16)

EM Fitting of Models for the Troponin Data

We did not collect data from patients known not to suffer from AMI. The guidelines rely on the general population of healthy subjects as the non-AMI reference population, specifically the upper 99th percentile of cTn concentrations. This is complicated by the fact that reference populations can be, and have been, selected after different criteria with widely different results.[40,41] In our selected patient population with suspected or confirmed AMI, we may reasonably expect only a negligible proportion of healthy subjects. We therefore chose the general hospitalized patient population as a non-AMI reference.

A distinguishing feature of troponin release due to AMI is that the troponin concentration displays a characteristic pattern of rise and gradual fall. Other causes of troponin release will typically result in cTn concentrations fluctuating around a constant level, although exceptions are known to occur. We fitted a model for the cTn data using fractional polynomials, which can reflect a rise-and-fall pattern in a flexible way.

For the log-transformed observations X_{kl} from patient k at measurement time s_{kl}, when belonging to population i, the final model was

$$\begin{gathered} X_{kl} = \alpha^{(i)} + \xi_k + \beta_1^{(i)} \log(s_{kl}) \sqrt{s_{kl}} + \beta_2^{(i)} \sqrt{s_{kl}} + \varepsilon_{kl}, \\ \beta_1^{(1)} = \beta_2^{(1)} = 0, \alpha^{(1)} = \alpha_0, \end{gathered} \tag{3}$$

where $\xi_k \sim N(0 \mid \tau)$ and $\varepsilon_{kl} \sim N(0|\sigma)$ are all independent stochastic variables.

Here, α_0 is a constant (common for both models) derived from published parameters for the non-AMI reference population. The EM algorithm was used to fit the model. We only included measurements within 72 hours of admission. Observation series were censored at the time of angiography, where applicable.

Our data included measured concentrations of both troponin types (cTnT and cTnI), as well as different assays. These measurements are not directly comparable. In the following analyses, the original data were standardized to a common standard, high sensitivity cTn (hs-cTnT), by calibration functions. Calibration parameters were found in published studies with samples analyzed by more than one method. The calibration functions were linear on the original or logarithmic scale, with intercepts and slopes shown in Table S2.

Published statistics, such as the median and interquartile range, from various non-AMI populations were used to specify the reference non-AMI population. This was done in two steps: (i) For each study included, a lognormal distribution was fit by least squares, or in the case of histograms, minimum Chi-square. The fitted parameters are shown in Table S1. Calibration to a common standard was first carried out when necessary. These distributions were regarded as components of a mixture. (ii) Eventually, a normal distribution was fitted, with prescribed variance and free mean, to have the best approximation of median and 75%-fractile to the mixture distribution. The fitted mean was used as the reference population value α_0 of Equation (3). The fixed variance value was taken, for convenience, from an initial exploratory model for the cTn dataset. Figure S2 shows the various fitted lognormal distributions, together with the resulting non-AMI reference population distribution.

Hypothesis Testing

Assuming Model (2) for stratified samples, we tested the hypothesis of homogeneity across the second factor while controlling for the first:

$$\begin{gathered} H_0 : \ \beta_2 = \cdots = \beta_J = 0 \text{ against} \\ H_1 : \beta_j \neq 0, \ \text{for at least one } j. \end{gathered}$$

We used a parametric bootstrap test for H_0 based on the likelihood ratio test statistic

$$R = 2\left(\sup_{\pi \in H_1} L - \sup_{\pi \in H_0} L \right).$$

To obtain p values, bootstrapping was used with 400 bootstrap replications of the test statistic R. The uncertainty in calibration was taken into account by randomly drawing the intercept and slope parameters of the calibration functions and subsequently restandardizing the data for each bootstrapping replication. The parameters of the LR for the cTn data

in each replication were estimated from parametric bootstrap samples. The LR for each clinical sign was drawn from a (one-dimensional) logit-normal distribution,[42] scaled to have the confidence intervals reported in Table 1 as limits and the point estimates as medians. The dispersion parameter σ of the logit-normal distribution was set to 1.

To address the research objectives, we tested the hypotheses of homogeneity between hospitals as well as between hospital 30D categories, controlling for patient stratum.

Sensitivity Analyses

Several analyses of robustness and sensitivity were performed. The bootstrap test was used to test for differences between abstracters, controlling for patient stratum. We also investigated two alternative models for cTn, incorporating the variable "time from symptom onset to admission", using either the recorded time interval as a categorical explanatory variable or imputed as interval midpoints. Note that this variable could not be used in the main analysis because of a high proportion of missing values.

We repeated the tests with altered datasets: (i) Due to the high frequency of transfers, we considered the possibility that angiography results were inconsistently recorded. We excluded the variable 'coronary artery stenosis'. (ii) As length of stay may be subject to variation due to hospital policies, introducing differences in allocation to the patient strata, we also repeated the tests after collapsing the two strata with lowest severity. (iii) The tests were also repeated with unequal response probabilities for present and not present signs, estimated by Equation (1). We considered the possibility that our choice of cTn parameters for the reference (non-AMI) population influenced the main analysis unduly. We therefore repeated the tests with three different choices of the fixed mean parameter α_0: (iv) corresponding to a 99%-fractile of 30 ng/L, which was the standard threshold for an AMI diagnosis in Norway before 2013; (v) with the non-AMI reference mean decreased by one; and (vi) with the non-AMI reference mean increased by one.

For each case of the double abstraction substudy, LRs were computed for the two datasets, and the samples were compared using the Wilcoxon signed rank test.

Results

The final dataset included 1146 patients from 12 hospitals, with 544, 518, and 84 patients in the low, intermediate, and high case severity strata, respectively. A flowchart is provided in Figure S1. Patient characteristics for the sample strata are shown in Table 2.

Assuming for simplicity that $q_1 = q_0$, we found an estimated common value of 0.799, used in the derivation of LRs.

After censoring, the number of cTn measurements ranged from 1 to 5, with a mean of 2.8. The fitted mean values for cTn are shown in Figure 1.

The overall PPV was 97.1%. Table 3 shows the estimated PPVs for each combination of hospital and case severity, as well as the overall value. The test for homogeneity between hospitals, while controlling for case severity, was significant with a p value of 0.015.

Table 2 Characteristics of the Study Sample After Hospital 30D Category

	L30D	M30D	H30D
Age (years)	71 (59–81)	74 (61–83)	73 (61–84)
Length of stay (days)	5.7 (3.8–10.4)	6.1 (4.1–11.3)	5.2 (3.4–8)
Females (%)	38	41	36
Transfers (%)	24	57	55
ST-elevation (%)	40	35	36
Previous coronary artery disease (%)	20	24	17
In-hospital death (%)	5.3	6.4	7.2

Note: Entries are percentages for categorical variables, median and interquartile range for continuous variables.

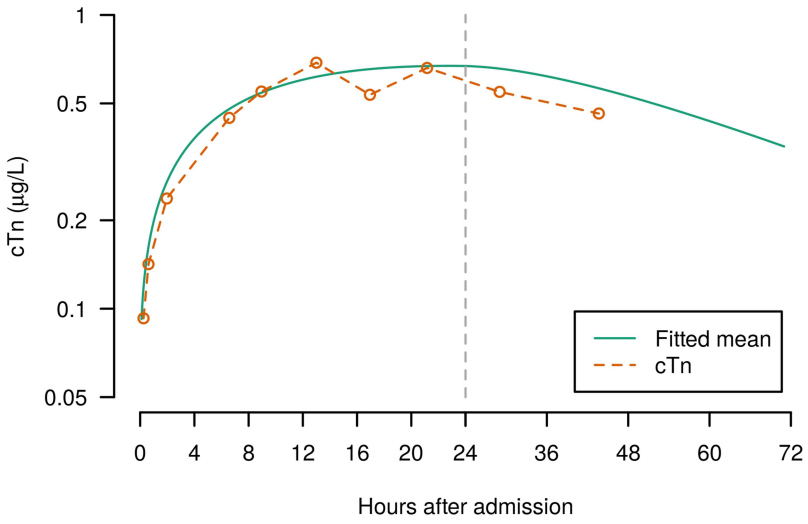

Figure 1 Fitted mean troponin vs time for the AMI population. Additionally, shown are medians of cTn over time intervals (notice break of time scale).

Table 4 shows the estimated PPVs for each combination of hospital 30D category and case severity, as well as the overall PPVs. The test for homogeneity between 30D categories, while controlling for case severity, was not significant, with a p value of 0.38. Note that the table does not give any indication of a possible monotone relationship between the 30D category and PPV, which would be the most plausible alternative to homogeneity.

Sensitivity Analyses

There was no significant difference between the abstracters using the bootstrap test (p value 0.76). As shown in Figure S3, the estimated mean cTn-values did not change materially when the time from symptom onset was included in the model. The estimates $\hat{q}_0, \hat{q}_1$ given by (1) were 0.57 and 0.91, respectively. The results of the other sensitivity analyses are shown in Table 5. The results for testing homogeneity between hospital 30D categories did not change,

Table 3 PPV After Hospital and Patient Case Severity

Hospital	Case Severity			Overall
	Low	Intermediate	High	
Fredrikstad	92.9	93.5	100	94.0
Gjøvik	89.6	90.4	100	91.2
Haraldsplass	100.0	100.0	100	100.0
Harstad	99.7	99.8	100	99.8
Haukeland	96.2	96.5	100	96.8
Kristiansand	100.0	100.0	100	100.0
Kristiansund	89.7	90.5	100	91.3
Levanger	100.0	100.0	100	100.0
Sandnessjøen	98.2	98.4	100	98.5
St. Olav	98.3	98.5	100	98.6
Tromsø	96.8	97.0	100	97.3
Vesterålen	98.4	98.6	100	98.7

Table 4 PPV After Hospital 30D Category and Patient Case Severity

Hospital 30D Category	Case Severity			Overall
	Low	Intermediate	High	
L30D	97.2	97.5	98.2	97.5
M30D	95.4	95.9	96.9	95.8
H30D	98.0	98.2	98.7	98.1

Table 5 Sensitivity Analysis: p values for Homogeneity Tests Under Alternate Specifications

	30D Categories	Hospitals
(i) Without the coronary artery stenosis variable	0.252	0.005
(ii) Low and intermediate case severity strata collapsed	0.382	0.010
(iii) Unequal response probabilities for present and absent signs	0.688	0.090
(iv) 99%-fractile of 30 ng/L used for non-AMI reference	0.698	0.428
(v) Non-AMI reference mean decreased by one	0.640	0.125
(vi) Non-AMI reference mean decreased by one	0.250	0.012

whereas the inhomogeneity between hospitals became nonsignificant if the reference value for cTn was set to a low value. In particular, this applied to the conventional decision limit for cTn. Additionally, the double abstraction substudy showed no significant difference (p value 0.062).

Discussion

We analyzed a set of medical record abstracts from a population of patients with a coded diagnosis of AMI taken from a nationally representative sample of hospitals. The data consisted of clinical signs and examination findings as well as serial cTn measurements. Parameters of the distribution of variables from non-AMI and AMI populations were determined from published data and, in the case of cTn measurements for AMI patients, estimated from the data using an EM algorithm. A finite mixture model was used to derive parametric bootstrap tests. We found significant heterogeneity between hospitals, with a PPV range of 91–100%. The test for homogeneity in the proportion of AMI patients between hospitals with low, intermediate, and high registry-based 30-day mortality was nonsignificant, while controlling for case severity stratum.

The variation in hospital PPV and its possible effect on the quality indicator 30-day mortality have not previously been studied, as far as the authors are aware. Our overall PPV estimate is in the upper part of the range reported in the literature. Recent Scandinavian studies show similar results, ranging from 95.1% to 97%.[18–20] The variation between hospitals may seem small compared with previous findings in the literature. We believe this is a result of the relative homogeneity of the Norwegian public hospital system. It is also a statistical necessity, given the high overall PPV.

Our approach to data abstraction and analysis obviated the need for a more or less subjective assessment of diagnosis as the gold standard. The use of LRs is a method that is optimal in the sense of statistical power. We were also able to account for the uncertainties due to standardizing measurements of different cTn types and in deriving likelihood ratios for clinical signs. The category "not recorded" for clinical signs was handled explicitly in the model. The hospital sample included local, intermediate and university hospitals from all Norwegian hospital regions. The analysis of variation between hospital 30D categories was robust against sensitivity tests.

For clinical practice, the present study will presumably only be indirectly useful by contributing to the use of quality indicators in quality and patient safety work. However, we think that our model for cTn measurements could profitably be investigated as a starting point for clinical decision rules.

The most critical assumption in our analysis is that the presence or absence of clinical signs and other relevant patient data are ascertained, recorded, and retrieved in a reasonably uniform manner across hospitals. For a hypothetical example, if one hospital only recorded the presence of signs, while another only recorded abscence, the PPV estimates would obviously be very different. The records were to a certain degree unstructured, and information was sometimes difficult to locate. We have seen that data from different abstracters varied. However, as they followed a structured record navigation sequence, we think that the consequence was increased variability, which the statistical analysis could accommodate, rather than bias. Double abstraction would reduce the variability in recording but increase the variation between patients due to the eventual halving of the sample size. Note that an alternative study design, with diagnosis adjucated by experts, would be subject to the same problems. The sensitivity and specificity data were based on data collected in presumably rather diverse settings, and it might be thought that they would not necessarily apply to our study and our choice of non-AMI reference population. However, the parametric bootstrapping, where LRs were drawn randomly for each replication, accommodated this source of uncertainty. The same applies to the standardization of the different types of cTn data. As shown by the sensitivity analysis, the analysis of variation between hospitals was to some extent dependent on the underlying assumptions.

Our data were obtained by a somewhat old medical record review and may not be completely representative of present practice. In particular, the high-sensitivity cTn assays that are used today were only used in a small fraction of cases. We think, however, that the results on differences between hospitals and types of hospitals are still valid. We were not able to investigate the relationship between Type 2 AMI and PPV, as the data were insufficient to distinguish between Type 1 and 2 AMI.

Uniformity in PPV across hospitals or some hospital categories ensures that the population of AMI cases is homogeneous in regard to the main medical condition. To conclude that the diagnosis and coding of AMI is the same across hospitals, the sensitivity would have to be uniform as well. The study design precluded the study of sensitivity. However, this would have been a formidable undertaking, at least with record review by humans as method.

Conclusions

The PPV of a coded diagnosis of AMI is high but somewhat variable between Norwegian hospitals. However, there was no significant variation in PPV between hospital categories with low, intermediate, or high registry-based 30-day mortality. We have presented a statistical approach that may be of wider interest. The concern motivating our study was the validity of registry-based 30-day mortality after admission for AMI as a quality indicator for hospitals. We conclude that the validity is not compromised by differences between hospitals in diagnosis and coding. Further research should aim to investigate potential heterogeneity in case severity.

Acknowledgments

The authors wish to thank Kristin Aakre and Kaare Bønaa for providing information about the troponin types and assays used in Norwegian hospitals and Tomislav Dimoski for developing the software for extraction from the hospitals' administrative databases. The conclusions of the paper are the authors' own.

Disclosure

The authors declare no competing interests in this work.

References

1. Haug C. Forskning uten filter [Research without filter]. *Tidsskr nor Laegeforen*. 2005;125(23):3243. Norwegian.
2. Lilford R, Pronovost P. Using hospital mortality rates to judge hospital performance: a bad idea that just won't go away. *Br Med J*. 2010;340.
3. McKee M. Hospital standardised mortality rates should not be used to make interhospital comparisons. *BMJ*. 2013;347:f6155. doi:10.1136/bmj.f6155
4. Gestel YR, Lemmens VE, Lingsma HF. The hospital standardized mortality ratio fallacy: a narrative review. *Med Care*. 2012;50(8):662–667. doi:10.1097/MLR.0b013e31824ebd9f
5. Coombes R. Experts disagree about usefulness of hospital mortality data. *BMJ*. 2014;349:g5658. doi:10.1136/bmj.g5658

6. Ranstam J, Wagner P, Robertsson O, Lidgren L. Health-care quality registers: outcome-orientated ranking of hospitals is unreliable. *J Bone Joint Surg Br*. 2008;90(12):1558–1561. doi:10.1302/0301-620X.90B12.21172
7. Bosch WF, Silberbusch J, Roozendaal KJ, Wagner C. Variatie in codering patientengegevens beinvloedt gestandaardiseerd ziekenhuissterftecijfer (HSMR). [Variations in patient data coding affect hospital standardized mortality ratio (HSMR)]. *Ned Tijdschr Geneeskd*. 2010;154:A1189. Dutch.
8. Ridgeway G, Norgaard M, Rasmussen TB, et al. Benchmarking Danish hospitals on mortality and readmission rates after cardiovascular admission. *Clin Epidemiol*. 2019;11:67–80. doi:10.2147/CLEP.S189263
9. Chapman AR, Adamson PD, Mills NL. Assessment and classification of patients with myocardial injury and infarction in clinical practice. *Heart*. 2017;103(1):10–18. doi:10.1136/heartjnl-2016-309530
10. Thygesen K, Alpert JS, Jaffe AS, et al. Fourth universal definition of myocardial infarction (2018). *Circulation*. 2018;138(20):e618–e651. doi:10.1161/CIR.0000000000000617
11. Sandoval Y, Smith SW, Sexter A, Schulz K, Apple FS. Use of objective evidence of myocardial ischemia to facilitate the diagnostic and prognostic distinction between type 2 myocardial infarction and myocardial injury. *Eur Heart J*. 2020;9(1):62–69. doi:10.1177/2048872618787796
12. Chapman AR, Shah AS, Lee KK, et al. Long-term outcomes in patients with type 2 myocardial infarction and myocardial injury. *Circulation*. 2018;137(12):1236–1245. doi:10.1161/CIRCULATIONAHA.117.031806
13. DeFilippis AP, Chapman AR, Mills NL, et al. Assessment and treatment of patients with type 2 myocardial infarction and acute nonischemic myocardial injury. *Circulation*. 2019;140(20):1661–1678. doi:10.1161/CIRCULATIONAHA.119.040631
14. Madsen M, Davidsen M, Rasmussen S, Abildstrom SZ, Osler M. The validity of the diagnosis of acute myocardial infarction in routine statistics: a comparison of mortality and hospital discharge data with the Danish MONICA registry. *J Clin Epidemiol*. 2003;56(2):124–130. doi:10.1016/S0895-4356(02)00591-7
15. Pajunen P, Paakkonen R, Juolevi A, et al. Trends in fatal and non-fatal coronary heart disease events in Finland during 1991-2001. *Scand Cardiovasc J*. 2004;38(6):340–344. doi:10.1080/14017430410010996
16. Joensen AM, Jensen MK, Overvad K, et al. Predictive values of acute coronary syndrome discharge diagnoses differed in the Danish national patient registry. *J Clin Epidemiol*. 2009;62(2):188–194. doi:10.1016/j.jclinepi.2008.03.005
17. Coloma PM, Valkhoff VE, Mazzaglia G, et al. Identification of acute myocardial infarction from electronic healthcare records using different disease coding systems: a validation study in three European countries. *BMJ Open*. 2013;3:6. doi:10.1136/bmjopen-2013-002862
18. Sundbøll J, Adelborg K, Munch T, et al. Positive predictive value of cardiovascular diagnoses in the Danish national patient registry: a validation study. *BMJ Open*. 2016;6(11):e012832. doi:10.1136/bmjopen-2016-012832
19. Govatsmark RES, Janszky I, Slordahl SA, et al. Completeness and correctness of acute myocardial infarction diagnoses in a medical quality register and an administrative health register. *Scand J Public Health*. 2020;48(1):5–13. doi:10.1177/1403494818803256
20. Varmdal T, Mathiesen EB, Wilsgaard T, et al. Validating acute myocardial infarction diagnoses in national health registers for use as endpoint in research: the Tromsø study. *Clin Epidemiol*. 2021;13:675–682. doi:10.2147/CLEP.S321293
21. McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of myocardial infarction diagnoses in administrative databases: a systematic review. *PLoS One*. 2014;9(3):e92286. doi:10.1371/journal.pone.0092286
22. Rubbo B, Fitzpatrick NK, Denaxas S, et al. Use of electronic health records to ascertain, validate and phenotype acute myocardial infarction: a systematic review and recommendations. *Int J Cardiol*. 2015;187:705–711. doi:10.1016/j.ijcard.2015.03.075
23. Rosamond WD, Chambless LE, Sorlie PD, et al. Trends in the sensitivity, positive predictive value, false-positive rate, and comparability ratio of hospital discharge diagnosis codes for acute myocardial infarction in four US communities, 1987-2000. *Am J Epidemiol*. 2004;160(12):1137–1146. doi:10.1093/aje/kwh341
24. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. *Lancet*. 2005;365(9469):1500–1505. doi:10.1016/S0140-6736(05)66422-7
25. Apple FS. Counterpoint: standardization of cardiac troponin I assays will not occur in my lifetime. *Clin Chem*. 2012;58(1):169–171. doi:10.1373/clinchem.2011.166165
26. Hassani S, Lindman AS, Kristoffersen DT, Tomic O, Helgeland J. 30-day survival probabilities as a quality indicator for Norwegian hospitals: data management and analysis. *PLoS One*. 2015;10(9):e0136547. doi:10.1371/journal.pone.0136547
27. Helgeland J, Kristoffersen DT, Skyrud KD, Lindman AS. Variation between hospitals with regard to diagnostic practice, coding accuracy, and case-mix. A retrospective validation study of administrative data versus medical records for estimating 30-day mortality after Hip fracture. *PLoS One*. 2016;11(5):e0156075. doi:10.1371/journal.pone.0156075
28. McLachlan G, Peel D. *Finite Mixture Models*. Wiley; 2000.
29. Chun AA, McGee SR. Bedside diagnosis of coronary artery disease: a systematic review. *Am J Med*. 2004;117(5):334–343. doi:10.1016/j.amjmed.2004.03.021
30. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7(3):177–188. doi:10.1016/0197-2456(86)90046-2
31. Björck L, Nielsen S, Jernberg T, Zverkova-Sandström T, Giang KW, Rosengren A. Absence of chest pain and long-term mortality in patients with acute myocardial infarction. *Open Heart*. 2018;5(2):e000909. doi:10.1136/openhrt-2018-000909
32. Canto JG, Rogers WJ, Goldberg RJ, et al. Association of age and sex with myocardial infarction symptom presentation and in-hospital mortality. *JAMA*. 2012;307(8):813–822. doi:10.1001/jama.2012.199
33. Herlitz J, Starke M, Hansson E, Ringvall E, Karlson BW, Waagstein L. Early identification of patients with an acute coronary syndrome as assessed by dispatchers and the ambulance crew. *Am J Emerg Med*. 2002;20(3):196–201. doi:10.1053/ajem.2002.33003
34. Pope JH, Ruthazer R, Beshansky JR, Griffith JL, Selker HP. Clinical features of emergency department patients presenting with symptoms suggestive of acute cardiac ischemia: a multicenter study. *J Thromb Thrombolysis*. 1998;6(1):63–74. doi:10.1023/A:1008876322599
35. Pedersen CK, Stengaard C, Friesgaard K, et al. Chest pain in the ambulance; prevalence, causes and outcome-A retrospective cohort study. *Scand J Trauma Resusc Emerg Med*. 2019;27(1):1–10. doi:10.1186/s13049-019-0659-6
36. Widimský P, Anděl M. Prevalence of coronary atherosclerosis in asymptomatic population. *Eur Heart J*. 2000;21(1):13–14. doi:10.1053/euhj.1999.1889
37. Kang WY, Jeong MH, Ahn YK, et al. Are patients with angiographically near-normal coronary arteries who present as acute myocardial infarction actually safe? *Int J Cardiol*. 2011;146(2):207–212. doi:10.1016/j.ijcard.2009.07.001
38. Larsen AI, Galbraith PD, Ghali WA, et al. Characteristics and outcomes of patients with acute myocardial infarction and angiographically normal coronary arteries. *Am J Cardiol*. 2005;95(2):261–263. doi:10.1016/j.amjcard.2004.09.014

39. Gehrie ER, Reynolds HR, Chen AY, et al. Characterization and outcomes of women and men with non- ST-segment elevation myocardial infarction and nonobstructive coronary artery disease: results from the can rapid risk stratification of unstable angina patients suppress adverse outcomes with early implementation of the ACC/AHA guidelines (CRUSADE) quality improvement initiative. *Am Heart J*. 2009;158(4):688–694.
40. Cardinaels EP, Mingels AM, Jacobs LH, Meex SJ, Bekers O. A comprehensive review of upper reference limits reported for (high-)sensitivity cardiac troponin assays: the challenges that lie ahead. *Clin Chem Lab Med*. 2012;50(5):791–806. doi:10.1515/cclm-2011-0895
41. Hickman PE, Lindahl B, Cullen L, Koerbin G, Tate J, Potter JM. Decision limits and the reporting of cardiac troponin: meeting the needs of both the cardiologist and the ED physician. *Crit Rev Clin Lab Sci*. 2015;52(1):28–44. doi:10.3109/10408363.2014.972497
42. Aitchison J, Shen SM. Logistic-normal distributions: some properties and uses. *Biometrika*. 1980;67(2):261–272. doi:10.2307/2335470