Table of contents

MANUSCRIPT

Introduction

Rates of first-time hospitalization for acute myocardial infarction (MI) and subsequent 30-day mortality have declined by more than 50% in Denmark over the last three decades.[1] Still, 15% of Danish patients with first-time MI die within the first 30 days.[1] Moreover, short-term mortality is 27% among patients with a very severe comorbidity burden, making comorbidity burden a strong predictor of mortality.[1] The prevalence of a high comorbidity burden in patients with MI has increased over the last three decades[1] and with ageing of the population this prevalence is likely to become even higher.[2-4] There is thus a compelling need to better understand the effect of a comorbidity burden on MI prognosis.

Comorbidity indices (prediction models) are used widely for this purpose. Indices have been developed specifically for cardiac patients[5-8] and others in mixed populations with successive testing in cardiac patients.[9-12] The Charlson Comorbidity Index (CCI) is one of the most commonly used comorbidity indexes in research.[9] The CCI was initially developed using a small patient group consisting of 559 patients admitted to a medical center during a 1-month period in 1984.[9] Since then, the impact of comorbidities on survival has changed, with improvements in prophylaxis and treatment leading to longer survival.[1] Also, the CCI does not include psychiatric diagnoses.

Another common comorbidity index is the van Walraven-weighted version of the Elixhauser index.[10] This index was also developed using a mixed patient group to predict in-hospital mortality. This may not be ideal for assessing the comorbidity burden in Danish MI patients either.

In the current study, we aimed to construct comorbidity indices to adjust accurately for confounding variables in prognosis studies of MI patients. We conducted a population-based cohort study to develop and validate comorbidity-based prognostic indices that predict 1-year mortality after first-time MI. We developed indices both with and without cardiovascular comorbidities.

Methods

Setting and data sources

The Danish National Health Service provides universal tax-supported health care, guaranteeing free access to general practitioners and hospitals in Denmark.[13] All Danish residents are assigned a unique central personal registry (CPR) number at birth or upon immigration.[14] This number is used to record health data in Danish registries, facilitating registry linkage at the individual level.[14] The registries used in the current study are described below.

The *Danish Civil Registration System* contains data on birth, vital status and migration for the entire Danish population since 1968, with daily updates.[14]

The *Danish National Patient Registry* (DNPR) includes data on all non-psychiatric hospital admissions in Denmark since 1977, and on outpatient clinic and emergency room contacts since 1995.[15] Each contact is registered using one primary diagnosis and potentially several secondary diagnoses classified according to the International Classification of Diseases, Eighth Revision (ICD-8) until the end of 1993 and the Tenth Revision (ICD-10) thereafter.[15]

The *Aarhus University Prescription Database* contains data on prescriptions. The type of drug prescribed (using the Anatomical Therapeutic Chemical classification system), and the date the drug was dispensed are transferred electronically from pharmacies to the prescription database.[16]

The *Clinical Laboratory Information System Research Database* (LABKA) contains test results from all laboratory analyses performed on blood samples drawn at hospitals or by general practitioners and submitted to departments of clinical biochemistry in the North and Central Denmark Regions.[17] Complete geographical coverage was achieved in 1997 in the North Denmark Region and in 2000 in the Central Denmark region.[17]

Study population

We used the DNPR to identify a population-based cohort consisting of all patients aged 15 years or older admitted with a first-time hospitalization for MI (ICD-10: I21) in the North and Central Denmark regions between 1 January 2000 and 31 December 2013. To ensure that we identified incident MI diagnoses, we excluded patients with previous inpatient or outpatient MI diagnoses from any Danish hospital recorded in the DNPR. Of note, the DNPR included MI patients who died in the ambulance on the way to the hospital or during admission, but not if they died at home.

Outcome

The outcome of interest was time to all-cause mortality within 1 year of hospital admission, which we ascertained through linkage to the Danish Civil Registration System.[14] Follow-up continued until 31 December 2014, ensuring that all patients could potentially be followed for one year without missing data.

Potential predictors

A list of comorbidities was assembled from previously constructed indices and clinical knowledge. To identify all medical conditions that could have an impact on the prognosis after MI, medical conditions included in the ICD-10 were reviewed thoroughly. Conditions that could be considered a symptom or a complication of MI, such as cardiac shock or arrest, were not included.[18]

We obtained information on comorbid conditions from hospital inpatient and outpatient clinic diagnoses recorded in the DNPR within the 5 years before hospitalization for MI. We also included diagnoses recorded during the index admission for MI, except for diagnoses for possible complications of MI, antithrombotic treatment, or associated immobilization. Complications included stable angina pectoris, heart failure, deep venous thrombosis in a lower limb, pulmonary

embolism, atrial fibrillation, heart block, ventricular tachycardia, cardiac valve disease, and stroke. These conditions were therefore only included if they were recorded before MI hospitalization.

Based on Danish 5-year mortality estimates,[19] we categorized cancer as high-risk cancer (survival <30%) or lower-risk cancer (survival ≥30%). High-risk cancers included cancers of the hypopharynx, esophagus, stomach, liver, gallbladder, pancreas, trachea and lung, as well as mesothelioma, acute myeloid leukemia, unspecified leukemia, and secondary cancer. All remaining types of cancer were considered lower-risk cancers.

Certain conditions, such as diabetes, chronic pulmonary disease, and hypertension, may be treated solely in general practice and thus not be registered in the DNPR.[15] The same applies for conditions treated at psychiatric hospitals such as affective disorders and schizophrenia. We therefore supplemented records from DNPR with information from the LABKA database[17] or the Aarhus University Prescription Database (Table 1).[16]

The final list of potential predictors included 41 individual comorbidities (Table 1). In addition to constructing the DANish Comorbidity index for Acute Myocardial Infarction (DANCAMI) accounting for both cardiovascular and non-cardiovascular comorbidities, we wished to help researchers planning to study the effect of individual cardiovascular conditions separately while adjusting for non-cardiovascular comorbidities. From the 41 comorbidities, we therefore identified 24 non-cardiovascular conditions for inclusion in a second comorbidity index restricted to non-cardiovascular conditions (restricted (r)DANCAMI).

For comparison with our own comorbidity scores, we calculated patients' CCI[9] and Elixhauser[10] scores. As in the original CCI weighting paper,[9] we based the CCI score on 18 comorbidities (MI excluded). For the Elixhauser index, we based the score on 30 comorbidities. For descriptive statistics, we categorized both CCI scores (0, 1, 2, 3+) and Elixhauser scores (≤0, 1-5, 6-13, 14+).

We treated scores from both indices as continuous variables for predictive validation models and estimating nonparametric correlation with the DANCAMI and the rDANCAMI.

External validation cohort

We validated the performance of the DANCAMI and the rDANCAMI using patients diagnosed with first-time MI in New Zealand between 1 January 2007 and 31 December 2016. We used the unique New Zealand National Health Index (NHI) number, assigned to patients at entry into the public health system (>98% of the population),[20] to link the New Zealand National Minimum Dataset (hospital inpatient data),[20] the Mortality Collection (vital status),[21] and the Pharmaceutical Collection (dispensed prescriptions).[22] The National Minimum Dataset includes nationwide information on all patients discharged from publicly funded hospitals, including admission dates, primary diagnoses, and secondary diagnoses (*i.e.,* comorbidities).[20] Except for HbA1c data (which were unavailable), we designed the validation cohort using an approach identical to that used for the Danish MI cohort, including eligibility criteria, outcome, and definitions of baseline characteristics.

Statistical analysis

Model development

We defined the outcome as time to all-cause mortality within one year from the date of MI hospital admission. We calculated frequencies of categorical covariates and the median and interquartile range of age. We used Cox regression models adjusted for sex and age to compute hazard ratios (HRs) with 95% confidence intervals (CIs) for the minimally-adjusted association between each comorbidity and 1-year mortality.

For selection of variables to include in the complete DANCAMI, we included all 41 available comorbidities (both cardiovascular and non-cardiovascular conditions), and sex and age in

multivariable Cox models, regardless of the results from the minimally-adjusted analyses.[23] We used fractional polynomials[24] to examine the assumption that age had a linear association with 1-year mortality in multivariable models and found it valid. We eliminated comorbidities with a HR <1.10 or a 95% CI that overlapped 1, and fitted revised models with the remaining comorbidities, and sex and age. We repeated this procedure until the models included only comorbidities with a HR ≥1.10. We tested the proportionality assumption using the global test based on scaled Schoenfeld residuals[25] and with log-log plots for variables that appeared non-proportional. We assigned weights to each comorbidity in the final indices by multiplying the beta coefficient from the multivariable models by ten and rounding to the nearest integer to yield the score components. The score components were added to form the final score.[26,27] We repeated the above steps for the rDANCAMI model, including only the 24 non-cardiovascular comorbidities.

Model performance

We evaluated the performance of the DANCAMI/rDANCAMI in the Danish MI cohort for internal validation and in the New Zealand MI cohort for external validation.

We assessed performance in the Danish MI cohort using the following statistics: (1) a modified version of Nagelkerke's R^2 to measure overall performance with explained variation;[28] (2) Harrell's C-statistic to measure discriminative ability, for binary outcomes, Harrell's C-statistic is equivalent to the area under the Receiver Operating Characteristic curve and indicates the proportion of all pairs of patients in which the patient who died first had lower predicted mortality;[23] (3) five different predictive Cox regression models to contrast the performance measures in the baseline model (age and sex) *vs.* the baseline model plus the CCI, Elixhauser, DANCAMI, or rDANCAMI score; (4) the integrated discrimination improvement (IDI) and the continuous Net Reclassification Index (NRI) to compare the baseline model with each of the four comorbidity models. IDI and NRI

indicate how much a predictor adds to a model's discriminatory power and are joint measures of a model's comparative improvement in sensitivity and specificity.[29,30] The NRI represents the net proportion of patients with a change in predicted risk in the correct direction when the comorbidity score is added to a baseline model containing age and sex.[29] The IDI integrates the NRI over all possible cut-offs for the probability of an outcome and is the difference between predicted probabilities in those who do and those who do not experience the outcome. It is identical to the difference in discrimination slopes of two models.[30] A positive NRI or IDI indicates better prediction in the newer model than in the comparison model.

Sensitivity analyses

We examined how the DANCAMI/rDANCAMI performed when restricted to selected populations. We measured the performance in men and women separately; among patients aged 75 or above; and among patients surviving the initial MI hospitalization.

External validation

We estimated the predicted probability of 1-year mortality in the New Zealand MI cohort by using the score components determined during the DANCAMI/rDANCAMI development. We then compared the predicted probabilities with actual outcomes in the New Zealand cohort, using the same methods, performance validations, and sensitivity analyses described above. Within the New Zealand MI cohort, we also measured performance in self-reported ethnicity groups (European, Maori, Pacific, Indian, and Chinese/other Asian).

All statistical analyses were conducted using Stata Version 14.2 (Stata Corp, College Station, Texas, USA). The study was approved by the Danish Data Protection Agency (record number 2013-41-1924).

Results

Development

The Danish MI cohort included 36,685 MI patients (61% men) with a median age of 72 years (interquartile range: 61-81 years). Table 1 presents prevalence and associations with 1-year mortality (adjusted for age and sex) for each comorbidity. The most prevalent comorbidity in the population was hypertension (53%), followed by chronic pulmonary disease (22%) and stable angina pectoris (19%).

The final number of comorbidities included in the DANCAMI was 24 (Table 2), and the final number included in the rDANCAMI was 17 (Table 3). In both models, high-risk cancer received the highest score, with a severity weight of 10. Other comorbidities with a high severity weight were schizophrenia, hemiplegia, moderate to severe liver disease, and chronic pancreatitis, all receiving a severity weight of 5 or above in both indices.

One-year mortality in the Danish MI cohort was 24% (Table 4). Among members of the cohort, 71% had at least one cardiovascular or non-cardiovascular comorbidity, while 43% had at least one non-cardiovascular comorbidity (Table 4). Survival decreased with increasing DANCAMI scores, although for rDANCAMI it was similar for patients with scores of 1-2 and 3-4 (Figure 1).

Performance

Compared with a baseline model containing only age and sex, a prediction model with age, sex, and DANCAMI score was 1.04 times better at discriminating between patients with a high and a low 1-year mortality risk (C-statistic: 0.753) (Table 5). The explained variance (R^2) was 1.20 times that in the baseline model (R^2: 0.332). When the DANCAMI score was added to the baseline model, the IDI was 0.054 and the total NRI was 0.519, with 77% of non-events and 49% of events receiving a better predicted probability of 1-year mortality (*i.e.,* improved discrimination) compared with

predictions from the baseline model. Compared with the CCI and the Elixhauser index, DANCAMI was superior in all four performance measures (Table 5).

A prediction model containing age, sex, and the rDANCAMI score was also superior in all four performance measures compared with the CCI and the Elixhauser index, but not when compared with the DANCAMI (Table 5). The rDANCAMI discriminated 1.03 times better than the baseline model (C-statistic: 0.746) and R^2 was 1.15 times higher than for the baseline model (R^2: 0.318). When the rDANCAMI score was added to the baseline model, the IDI was 0.038 and the total NRI was 0.428, with 68% of non-events and 54% of events receiving a more correct predicted probability of 1-year mortality. Overall, the Elixhauser index had the poorest performance in all measures compared with the baseline model (Table 5).

External validation

The New Zealand MI cohort included 75,069 MI patients. One-year mortality was lower than in Denmark (Table 4). The proportion of men was 59%, and the median age was 71 years. The proportion of MI patients with at least one comorbidity (DANCAMI >0) was lower (67%), but higher in MI patients with at least one non-cardiovascular comorbidity (rDANCAMI >0) (47%). As in the Danish MI cohort, the two most prevalent comorbidities were hypertension (38%) and chronic pulmonary disease (17%). The third most prevalent comorbidity was diabetes with end-organ failure (16%) (Table 4).

In the New Zealand MI cohort, DANCAMI scores also added to the predictive performance compared with the baseline model. Discrimination was 1.07 times that of the baseline model (C-statistic 0.773) and R^2 was 1.32 times that of the baseline model (R^2: 0.373). IDI was 0.079 and NRI was 0.682, with 78% of non-events and 56% of events receiving a more correct predicted probability of 1-year mortality compared with the predictions of the baseline model. Performance of

the CCI and the Elixhauser index was nearly identical to that of DANCAMI, except for NRI where the CCI performance was lower than both the DANCAMI and the Elixhauser index (Table 5). rDANCAMI performance was lower compared with the other three indices (Table 5).

Sensitivity analyses

In the subpopulation analyses, the models performed better for males than for females. However, this finding was mostly attributable to the baseline model and not the added comorbidity score (Table 6a and 6b). The models performed best among patients surviving the initial MI hospitalization and worst among patients aged 75 or older. We observed this pattern in both the Danish and the New Zealand MI cohorts (Table 6a, 6b, 7a and 7b). In the New Zealand MI cohort, DANCAMI had the best performance in the European ethnicity group, while the Elixhauser index outperformed the other comorbidity indices in the other ethnicity groups (Table 8a and 8b).

Discussion

We developed two comorbidity indices predicting 1-year for mortality after MI based on (1) any type of comorbidity (DANCAMI), or (2) non-cardiovascular comorbidities alone (rDANCAMI). In a Danish MI cohort, our indices outperformed other common comorbidity indices. In the context of external validity, DANCAMI also showed satisfactory performance for MI patients in New Zealand. Thus, DANCAMI provides a valuable approach to adjusting the impact of a comorbidity burden on 1-year mortality in future MI prognostic studies.

To our knowledge, rDANCAMI is the first comorbidity index for MI patients to include only non-cardiovascular comorbidities. However, other comorbidity indices have been developed specifically for MI patients. A 1994 US study used Medicare data to develop a comorbidity index predicting 2-year mortality in MI patients.[5] Patients were diagnosed in 1987 and all were 30-day

survivors; thus, the index may not be generalizable to all MI patients. A Chinese comorbidity index was developed in 2016 to predict in-hospital mortality in MI patients admitted to a Beijing hospital during 2006-2010.[6] The investigators aimed to develop a method to adjust for heterogeneity between hospitals in China. In contrast to the DANCAMI, the Chinese index includes conditions such as cardiac arrest and shock. We excluded these conditions, as they could be complications of MI. A Spanish comorbidity index was developed in 2011 based on patients with non-ST-segment elevation acute coronary syndrome who were hospitalized between 2002 and 2008.[7] The Spanish index does not generalize to all MI patients, since its focus is patients with non-ST-segment elevation acute coronary syndrome. Moreover, the development cohort included only 1017 patients. Finally, a 2001 Canadian study developed two separate comorbidity indices predicting 30-day and 1-year mortality among MI patients with age group and sex included in the indices.[8] Unfortunately, the authors reported only regression coefficients and odds ratios with 95% CIs, without generating a simpler scoring system.

Unlike other comorbidity indices, both the DANCAMI and the rDANCAMI include multiple mental and behavioural disorders, including alcohol/drug abuse, schizophrenia and affective disorders, which are assigned relatively high weights of 3 to 5. In some indices, these disorders were not included or not considered for inclusion.[6-9] The Elixhauser index[10] and the 1994 US study[5] both include psychiatric diagnoses. In the Elixhauser index, drug abuse and depression score less than zero, while alcohol abuse and psychoses have a score of zero. In the US study, the prevalence of these disorders was very low compared with the Danish MI cohort. This could be due to use of different definitions of these diagnoses.

Both the DANCAMI and the rDANCAMI showed higher discrimination in the New Zealand validation cohort than in the development cohort, which was unexpected. However, the CCI and the Elixhauser index also showed higher discrimination in the New Zealand MI cohort compared to the

Danish MI cohort. These findings may reflect different case mixes in the two national cohorts, *e.g.,* a more ethnically diverse population in New Zealand than in Denmark. DANCAMI was slightly superior in the European cohort that is likely to be more comparable with the Danish population compared to the other ethnicity groups in the New Zealand population.

Major strengths of our study are its population-based design, the large sample size, and importantly, an external validation, which indicated that both the DANCAMI and the rDANCAMI indices generalize well outside the Danish cohort where it was developed. Furthermore, the rDANCAMI allows researchers to study the effect of individual cardiovascular diseases separately while adjusting for non-cardiovascular comorbidities. As well, we used recommended methods to generate summary scores in our final indices and considered a variety of variables for both indices, including psychiatric diagnoses, which are rarely included in comorbidity indices.

Additionally, the comorbidities included in the DANCAMI/rDANCAMI were identified both prior to and during hospital admission, and excluded conditions that could be a direct consequence of the MI admission. This approach ensured that all comorbidities were present before the MI admission.

Although we used a five-year look-back period to identify comorbidities and defined variables using a variety of approaches (*i.e.*, diagnoses from the index hospitalization, prescription redemptions, and laboratory data), misclassification of some conditions must be expected since we used routine secondary care data sources.[31] Like previous studies,[5,10] we found several comorbidities that were associated with a decreased 1-year mortality (*e.g.* stable angina pectoris and anxiety) in our multivariable model. These seemingly protective comorbidities could result from a coding bias in which severity of overall patient illness may inversely affect the coding of chronic and nonfatal comorbidities.[10] We therefore excluded these comorbidities from our final indices.

Future validation studies are needed to examine generalizability to other look-back periods for identifying comorbidities.

Another concern is that we lacked clinical information, *e.g.,* electrocardiogram results and cardiac biochemical markers, which may be important predictors particularly of short-term mortality. This is evident in clinical risk prediction models such as the Global Registry of Acute Coronary Events (GRACE) risk score[32] and may explain the superior performance of our indices among patients surviving hospital admission. However, clinical information is often not available in routine secondary care data which makes it less useful as predictors in this setting.

In conclusion, we found that the comorbidity burden is a strong predictor of mortality in MI patients and must be controlled for accurately when studying outcomes in these patients. We have developed two separate comorbidity indices that can be used to control for the comorbidity burden in MI patients; an overall comorbidity index (DANCAMI) showing superior performance compared with other commonly used comorbidity indices, and a novel index restricted to non-cardiovascular comorbidities (rDANCAMI). The indices have potential to be used for adjusting for comorbidity burden in observational studies of MI patients in Western countries similar to Denmark and New Zealand.

SUPPLEMENTARY

Background

Prediction studies

Prediction studies predict outcomes from multiple variables rather than investigating whether a single variable may be prognostic,[33] and they aim at predicting the probability that an outcome will occur in an individual.[33] Prediction models are developed to estimate the probability of a future event (prognostic prediction model) or the probability that a certain disease or condition is present (diagnostic prediction model).[34] In medicine, the prognostic prediction model can help assess an individual's risk of developing a specific state of health based on his or her individual risk profile.[34]

Prediction models can be developed with different purposes. In a clinical setting, they are developed to: inform patients about the future course of their illness; guide doctors and patients in decisions involving choice of treatment; or help identify relevant patients for therapeutic research.[34] Examples of clinical prediction models in cardiology are the CHA2DS-VASc-score and HAS-BLED-score which estimate the risk of stroke or bleeding, respectively, in patients with atrial fibrillation. These risk scores are applied when deciding whether or not a patient should receive anticoagulation treatment. Another example is the Framingham risk score which estimates individuals' 10-year cardiovascular risk.

In a research setting, a prediction model can be useful in adjusting for confounding variables in observational research where correct adjustment is essential.[23] Prediction models are also applicable in comparing differences in performances between hospitals.[33] When analyzing factors like budgets or patient outcome it is important to adjust for case-mix, for example the proportion of severely ill patients.

Comorbidity indices are developed in prediction studies, and they use comorbidities to predict a chosen outcome in a specific population. Comorbidity indices are often used in research to adjust for the comorbidity burden between individual patients or patient groups.

Methodological considerations

Design

Prognostic prediction studies are inherently longitudinal studies in nature as opposed to diagnostic studies, which are often designed as cross-sectional studies.[23] To develop our comorbidity indices, we conducted a historical cohort study using prospectively collected data. A cohort is defined as a group of individuals who are followed over a period of time.[35] In a cohort study, you can measure the occurrence of one or multiple outcomes and compare these outcomes across baseline characteristics of the individual study participants.[35] A historical cohort study is conducted from previously recorded information and the time of exposure has taken place before the beginning of the study.[35] In general, a historical cohort study has several advantages: it usually takes shorter time to conduct compared with a non-historical cohort study and it is often more cost-effective since the information has been gathered beforehand.[35] The large amount of different information gathered in the Danish registries makes this study design ideal.

A different study design applicable for prognosis studies are prospective cohort studies including randomized clinical trials (RCTs). In a prospective cohort study, the predictors and the outcomes are collected concurrently with the conduct of the study.[35] Ideally, prognosis studies require a high number of outcome events to reduce the risk of overestimating the predictive performance of the model;[33] it has been proposed that at least 10 events are required for each candidate predictor studied.[33] This makes studies with longer follow-up time using a prospective study design more

expensive and time consuming. Moreover, prognostic models developed from RCTs often have a limited generalizability due to factors like strict eligibility criteria for the trial.[33]

Covariates and data source

Definition of comorbidity

The definition of comorbidity is not clearly distinguished, and the concepts of comorbidity, complications and multimorbidity are used interchangeably.[18] A clear differentiation between the concepts is important when they are applied in predictive models as well as causal studies. The definitions we applied in our study have previously been proposed: Comorbidities are medical conditions that exist at the time of diagnosis of the index disease (the main condition under study) and they are not a consequence of the index disease; complications are adverse events occurring after the diagnosis of the index disease; multimorbidity is the existence of two or more chronic diseases.[18]

Still, comorbidity is classified differently in various studies. Some define comorbidities as currently active conditions that should have a cogent impact on the prognosis.[9] Thus, resolved conditions (*e.g.,* previous pneumonia) and a history of operation for current inactive conditions (*e.g.,* cholecystectomy) were not included in this definition.[9] Other studies emphasize that the comorbidity should be present at admission and not be related directly to the main reason for hospitalization.[36] Moreover, comorbidities are defined as conditions that increase the intensity of resources used or increase the likelihood of a poor outcome.[36]

For our study, the collective list of comorbidities was assembled from previously applied comorbidity lists as well as clinical knowledge. The ICD-10 was thoroughly examined to identify all medical conditions that could be considered having an impact on the prognosis after an MI.

Extracting data on comorbidity

For the identification of a patient's comorbidities, we applied a look-back period of five years before hospital admission.

The sensitivity and specificity of comorbidities cannot be considered perfect when using databases to extract the diagnoses since there are diagnoses with relatively low validity, *e.g.,* heart failure with a PPV of 76%.[37] This means that among those with heart failure in the registry, only 76% actually have the condition when compared to a reference standard. Another measure of data quality is completeness. Completeness is the proportion of true cases with a disease that is correctly captured in the registry. Completeness can be measured in relation to all individuals in the general population with a specific disease or all patients admitted or treated for the specific disease, and it is largely dependent on the sensitivity of the registry.[15] When inpatient data alone are used to identify comorbidity, the total comorbidity burden may be underestimated.[38] Diseases like uncomplicated diabetes and chronic pulmonary disease are often treated by the general practitioner and registration in the DNPR may be incomplete.[15] Therefore, in addition to the ICD-10 diagnosis codes, we used LABKA data and prescription data in the DNPR to identify more patients with these diseases. This has previously been proposed elsewhere.[15] 1) In addition to the prognosis codes, we identified diabetes from HbA1c >48 mmol/L[17] or an antidiabetic prescription.[16] 2) We supplemented chronic pulmonary disease diagnosis codes with any prescription record for a drug used against obstructive airway disease.[16] 3) We defined hypertension as a hospital diagnosis, redemption of combination antihypertensive tablets or redemption of at least two prescriptions for antihypertensive drug classes within 90 days before admission.[39] 4) We supplemented diagnosis codes for schizophrenia and affective disorder with prescriptions suggesting pharmacotherapy for these disorders in the last 90 days.[16]

Development of the comorbidity indices

Management and selection of predictors

First, we assessed the univariate association between baseline characteristics and 1-year mortality adjusted for sex and age. Independently of the univariate analyses, we tested all comorbidities in the multivariable analysis. Another approach could have been to only include comorbidities with a p-value below a desired value to enter the multivariable analysis. However, this approach could lead to a wrong inclusion or exclusion of important comorbidities from the final model.[40] This is due to the fact that the univariate analysis cannot account for possible confounding that may exist between the comorbidities and the outcome,[40] and a non-significant association in the univariate analysis does not always mean that a variable is unimportant.[40]

We adjusted for age and sex in both univariate and multivariable analyses because this information is always available, and age is a surrogate measure of comorbidity.[41] Age was included as a continuous variable since more predictive information is retained compared with using categorization of continuous variables.[42] Age was tested for a linear association with the outcome using fractional polynomial approach. When a continuous variable is included as a linear term, it is assumed that the effect of an increase by one unit is the same at each part of the range of the variable.[23] If this assumption is incorrect, it can lead to misinterpretation of the influence of the variable. This can affect a model's overall predictive ability in new patients.[42] In our study, we did observe a linear association between age and 1-year mortality, and age was included as a linear term in all analyses.

It is expected that mortality is higher in the acute phase after an MI relative to the subsequent phase.[1] By applying a Cox proportional hazards regression in our analyses, we could account for time of death for patients dying within the first year after MI admission. We tested the

proportionality assumptions by the global test based on scaled Shoenfeld residuals,[25] and plotting log(-log(survival)) versus log(time) for variables that appeared non-proportional.

There exists no consensus on the best method when it comes to variable selection in prediction models.[42] Our approach was to eliminate covariates with a HR<1.10 or the appertaining 95% CI overlapping 1.00. We evaluated implication of this approach by conducting a sensitivity analysis using a HR cut-off of 1.20 in our variable selection process.

Weight assignment for summary risk indices

Weights were assigned to all comorbidities in the final models using the beta coefficients from the multivariable models after variable selection. Each beta coefficient was multiplied by ten and rounded to the nearest digit which formed the score components for each comorbidity in the final indices. This is in line with recommended best practice for risk index-development with studies demonstrating better performance of prediction models with this approach compared with hazard ratio-based risk indices.[26,27]

Validation

Before a prediction model (or a comorbidity index) should be implemented in the clinic or in research, it is essential to know if the model is valid and makes accurate predictions. A model with high validity makes satisfactory predictions that can be interpreted by the people using it. Assessing model performance is essential, which is why a model should be tested for its accuracy and its generalizability. Accuracy is the degree to which the model predictions match the observed outcome in the same cohort based on which the model was developed. Generalizability is the ability of the model to provide accurate prediction in different samples of patients.[43]

Performance measures

Accuracy can be tested with an overall performance measure like R^2. R^2 is used to quantify how close predictions are to the actual outcome.[30] The distance between observed and predicted outcomes is related to the concept of "goodness-of-fit" of a model. In a valid model there is a small distance between the predicted and the observed outcomes.[23] When performance is evaluated in the development cohort, it is usually defined as "goodness-of-fit", while performance measured in a cohort that differs from the development cohort is usually defined as predictive performance.[30]

Accuracy is also tested in two separate measurements, discrimination and calibration.[43] Discrimination describes the model's ability to rank patients according to their risk of the outcome. If there is an error in discrimination, the relative ranking of the individuals' risks is out of order.[43] Discrimination is important when the model is used to stratify patients by stage of severity, for example when comparing a treatment within different disease stages.[43] Calibration describes the model's reliability. If there is an error in calibration, the predicted probability become either too high or too low.[43] Calibration is especially important when a model is to be used for counseling patients.[43] Here it is essential that if the model predicts a 90% risk of an outcome, then the outcome should in fact occur in nine out of ten patients. The purpose of our comorbidity indices is not to counsel patients. Instead, our indices should be applied to adjust for comorbidities in observational studies. Therefore, we focused on discrimination and not calibration for the performance measures.

We assessed discrimination using different measures. First, C-statistic is equivalent to the probability that a prediction model assigns a higher predicted mortality to the patient dying first than the patient dying last, out of all pairs of subjects.[23] For binary outcomes, C-statistic is equivalent to the area under the Receiver Operating Characteristic (ROC) curve, a plot of sensitivity (true positive rate) against 1-(false-positive rates) for all possible outcomes.[30] A straight line with a

slope equal to one signifies complete lack of ability to discriminate. The farther the curve lies up and to the left of this line, the greater is the model's discriminating power.[44]

Second, continuous net reclassification improvement (NRI) and integrated discrimination improvement (IDI) indicate how much a new predictor adds to a model's ability to discriminate.[30] NRI is based on the concept of reclassification tables. If adding a new predictor to a prediction model will give cases (people who will experience the event early) a higher predicted risk and conversely give controls (people who will not experience the event) a lower predicted risk, the reclassification is successful.[29] The continuous NRI is estimated as the difference between the proportion of cases moving up and the proportion of cases moving down and the corresponding difference in proportions controls moving up and the proportion of controls moving down, and taking a difference of these two differences.[29]

$$\widehat{NRI} = \left(\hat{p}_{up,controls} - \hat{p}_{down,controls}\right) - \left(\hat{p}_{up,cases} - \hat{p}_{down,cases}\right)$$

The IDI integrates the NRI over all possible cut-offs for the probability of the outcome, and it is identical to the difference in discrimination slopes of two models.[30] Compared to NRI that assigns any correct reclassification with 1 point and the corresponding incorrect reclassification with -1 point, IDI assign to each individual the actual difference in predicted probabilities. IDI can be estimated as the difference between the mean predicted probability in cases and control in the model with the added predictor (*new model*) compared with the difference between the mean predicted probability in cases and control in the original model (*old model*).[29]

$$\widehat{IDI} = \left(\bar{\hat{p}}_{new,cases} - \bar{\hat{p}}_{new,controls}\right) - \left(\bar{\hat{p}}_{old,cases} - \bar{\hat{p}}_{old,controls}\right)$$

Continuous NRI and IDI were used in addition to C-statistics because improvement in C-statistics has been argued to be dependent on the baseline model's ability to discriminate.[45] Compared with a model with low baseline discrimination, a model with good discrimination will have a smaller improvement in C-statistics when adding a new predictor, even when the predictor

has a strong effect size. IDI is less dependent on the comparator model's discrimination while the continuous NRI has been proven to only be dependent on the effect size of the added predictor and not on the strengths of the comparator model.[45]

External validation

External validation, also called generalizability, is the ability of the model to provide accurate predictions among patients drawn from a different population from where it was developed, this can be a different time setting, geographical location, or in a population with a different disease composition.[43] If a model has low generalizability it may be caused by either underfitting or overfitting.[43] Underfitting occurs when important independent predictors of outcome are missing from the model. Overfitting occurs when the model is influenced by random noise from the development dataset.[43] We assessed generalizability of our comorbidity indices by examining their performances in a cohort of New Zealand MI patients. These analyses demonstrated satisfying performance by our indices in the New Zealand MI cohort with better discriminatory ability compared with a baseline model containing age and sex. DANCAMI performed approximately equally as good as other commonly used comorbidity indices.

Additional strengths and limitations

In addition to the validity of the prediction model as described above, it is important to discuss the validity of the study in general when performing epidemiologic research. In the literature, two broad types of error are described: random error and systematic error.[35]

Random error

Random error is the variability in data that lead to normal variation of an estimate.[35] To indicate the precision of an estimate, we use CIs which is a range of values around the estimate. Estimates with high precision have narrow confidence intervals and estimates with low precision have broad CIs. For example: the estimates of the univariate association between comorbidities and mortality have various precisions. This is illustrated in the different ranges of CIs. Rare comorbidities like hemiplegia have broad CIs and common comorbidities like hypertension have narrow CIs. The impact of random error is susceptible to the size of the study population.[35] With an increasing study population, the precision of the estimates will increase and random error will be reduced.[35]

Systematic error

In contrast to random error, systematic error is not susceptible to study population size.[35] Systematic error is divided into selection bias, information bias and confounding.[35]

Selection bias

Selection bias arises when the study population is not representative to the population you wish to examine in your study, for example if the association between exposure and outcome differs between study participants and non-participants.[35] If selection of study participants or a loss to follow-up depends on both exposure and outcome, it can lead to selection bias.

As mentioned, we used population-based registries to select study participants for our study. We included all patients with a first-time hospitalization for MI in the North and Central Denmark regions. The MI diagnosis has been proved to be accurately recorded in the DNPR with a positive predictive value of 97%.[37] Moreover, we had complete 1-year follow-up, leading to a minimal risk of selection bias in our study.

Information bias

Another way to introduce systematic error in research studies is by including erroneous information from or about the study participants. If this information causes patients to be placed in an incorrect category, the information is referred to as misclassified.[35] If the information depends on another study variable, the misclassification is said to be differential. If not, the misclassification is non-differential.[35] Non-differential misclassification of a dichotomous exposure will mostly lead to dilution of the final estimate with a value closer to "no effect". Differential misclassification, on the other hand, can lead to both overestimation and underestimation of the true effect.[35]

In our study, predictors were sex, age and comorbidities. The outcome was 1-year all-cause mortality. As regards to the outcome, information on all-cause mortality is very unlikely to be misclassified in the Danish registries. With respect to the predictors, information about sex and age is also unlikely to be misclassified. However, as described above, information on comorbidities from registries may still be incomplete.[15] Since the information in the registries is recorded prospectively and independently of the study outcome, we have no reason to suspect these errors to be directly related to 1-year mortality or be caused by recall bias. However, earlier studies have argued that doctors may record fewer comorbidities in severely ill patient admitted to a hospital which may result in coding bias.[46] Therefore, some comorbidities may be affected by this coding bias with coding of chronic and nonfatal comorbidities being inversely related to the severity of overall patient illness.[10] Thus, patients who are not severely ill will have more of this type of comorbidity recorded during a hospitalization. If they in general are in better health, they would be expected to live longer, even after an MI. These types of comorbidities would therefore be associated with better survival, and the effect on mortality would be underestimated. This may explain why some comorbidities in our analyses also seem to be associated with a decreased risk of 1-year mortality.

Confounding

Confounding is confusion of effects.[35] This implies that the effect of one exposure is mixed with the effect of another variable. A confounder is characterized by three things: 1) it is independently associated with the outcome; 2) it is associated with the exposure of interest; and 3) it is not an intermediate step between the exposure and the outcome.[35] Confounding can be minimized by applying different methods, either in the design of the study (*e.g.*, randomization, restriction and matching) or in data analyses (*e.g.*, stratification, standardization and adjustment).[35]

Confounding is a major concern in most observational research. Since randomization is often impossible, the studied groups may not be directly comparable. If these differences are not adjusted for, it can result in systematic error. Adjustment is essential in this type of research.[23]

Contrary to observational research, the goal in prediction models is not to explain whether an outcome can be attributed to a particular risk factor. Instead the purpose is to develop a model that accurately predicts a certain risk of a future outcome.[33] Therefore, predictors in prediction studies are not necessarily causally associated with the outcome, and controlling for confounding is not a consideration when building a prediction model. In fact, a non-causal factor can be useful as a predictor if it can replace a well-known causal factor that is more difficult to measure.[33]

Additional sensitivity analyses

We conducted three sensitivity analyses to evaluate how decisions made during model development affected the final indices: (1) We changed the HR cut-off from 1.10 to 1.20; (2) We used the exact rather than the rounded beta coefficients for score components; and (3) We used the HRs instead of the beta coefficients for score components.

Changing the inclusion threshold to an HR of 1.20 removed all comorbidities with a score of one from the DANCAMI models. The modified threshold had very limited effect on the remaining

comorbidities or the performance of the models (Supplementary table A). Changing the severity weights to precise beta coefficients did not alter model performance either (Supplementary table A). However, the models performed more poorly on almost all performance measures when we used HRs as severity weights (Supplementary table A). This is consistent with previous research.[26,27]

Split sample internal validation

In an additional sensitivity analysis, we performed split sample internal validation.[23] Following recommendations from the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) Initiative,[34] we split the Danish MI cohort into two subcohorts by time (temporal validation) rather than randomly. The two Danish subcohorts included a development cohort with all Danish MI patients diagnosed in 2000-2009, and a validation cohort with all Danish MI patients diagnosed in 2010-2013. We developed a new complete model and a restricted model by fitting the comorbidities from the two DANCAMI models in the development cohort. We assessed severity weights from the new beta coefficients, as described previously. We then assessed performance in the validation cohort using the same performance measures discussed above (R^2, C-statistics, IDI and NRI).

For temporal validation, 1-year mortality was higher in the development cohort (26%) than in the validation cohort (21%), but the baseline characteristics were similar (Supplementary table B). We refitted both the DANCAMI and the rDANCAMI in the development subcohort. Compared with the original indices, six comorbidities received different severity weights in the refitted DANCAMI and five in the refitted rDANCAMI, with 7 to 4 and 6 to 3 being the largest changes (not shown). When the refitted DANCAMI/rDANCAMI were tested in the validation subcohort, they performed better than the CCI and the Elixhauser index (Supplementary table C).

Additional discussion and perspectives

In our performance analyses, we used summary scores to compare the DANCAMI, rDANCAMI, the CCI and the Elixhauser index. The CCI and the Elixhauser index have previously been validated in MI patient populations to examine their predictive ability in this patient group. In studies performed in the US,[47] Taiwan,[48] and five different European countries,[49] the Elixhauser index outperformed the CCI in predicting in-hospital[47-49] and 1-year mortality.[48] These studies differ from ours as they included comorbidities as separate variables in their performance analyses instead of using a summary score. A Japanese study compared the performance of the CCI and the Elixhauser index using individual comorbidities *vs.* a summary score,[50] and found that the Elixhauser index performed better with individual comorbidities. However, the CCI and the Elixhauser index performed similarly when the summary score was applied. The performances with summary scores were generally lower than the performances with individual comorbidities.[50]

In observational studies, in which multiple variables often are included in regression analyses, it may be best to adjust for a summary comorbidity score due to limited data. In our performance analyses of summary scores, the CCI showed better performance than the Elixhauser index in the Danish MI cohort. In contrast, the Elixhauser index performed marginally better than the CCI in the New Zealand MI cohort. This was also true for the other performance measures.

This demonstrates that performance of the individual comorbidity indices can vary depending on their application. Using comorbidity indices to adjust for comorbidity burden is a useful tool in observation research. However, it is also a simplification of the more complex association between comorbidities. Researchers should be aware of these limitations when applying comorbidity indices and scores in future research.

REFERENCES

1. Schmidt, M., Jacobsen, J.B., Lash, T.L., Botker, H.E. & Sorensen, H.T. 25 year trends in first time hospitalisation for acute myocardial infarction, subsequent short and long term mortality, and the prognostic impact of sex and comorbidity: a Danish nationwide cohort study. *BMJ (Clinical research ed.)* **344**, e356 (2012).
2. Wolff, J.L., Starfield, B. & Anderson, G. Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Archives of internal medicine* **162**, 2269-2276 (2002).
3. Eyre, H., Kahn, R. & Robertson, R.M. Preventing cancer, cardiovascular disease, and diabetes: a common agenda for the American Cancer Society, the American Diabetes Association, and the American Heart Association. *Diabetes care* **27**, 1812-1824 (2004).
4. Fraccaro, P.*, et al.* Predicting mortality from change-over-time in the Charlson Comorbidity Index: A retrospective cohort study in a data-intensive UK health system. *Medicine* **95**, e4973 (2016).
5. Normand, S.L., Morris, C.N., Fung, K.S., McNeil, B.J. & Epstein, A.M. Development and validation of a claims based index for adjusting for risk of mortality: the case of acute myocardial infarction. *Journal of clinical epidemiology* **48**, 229-243 (1995).
6. Qu, Z., Zhao, L.P., Ma, X. & Zhan, S. Building a Patient-Specific Risk Score with a Large Database of Discharge Summary Reports. *Medical science monitor : international medical journal of experimental and clinical research* **22**, 2097-2104 (2016).
7. Sanchis, J.*, et al.* Influence of comorbid conditions on one-year outcomes in non-ST-segment elevation acute coronary syndrome. *Mayo Clinic proceedings* **86**, 291-296 (2011).
8. Tu, J.V.*, et al.* Development and validation of the Ontario acute myocardial infarction mortality prediction rules. *Journal of the American College of Cardiology* **37**, 992-997 (2001).
9. Charlson, M.E., Pompei, P., Ales, K.L. & MacKenzie, C.R. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of chronic diseases* **40**, 373-383 (1987).
10. van Walraven, C., Austin, P.C., Jennings, A., Quan, H. & Forster, A.J. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Medical care* **47**, 626-633 (2009).
11. Li, P., Kim, M.M. & Doshi, J.A. Comparison of the performance of the CMS Hierarchical Condition Category (CMS-HCC) risk adjuster with the Charlson and Elixhauser comorbidity measures in predicting mortality. *BMC health services research* **10**, 245 (2010).
12. Holman, C.D., Preen, D.B., Baynham, N.J., Finn, J.C. & Semmens, J.B. A multipurpose comorbidity scoring system performed better than the Charlson index. *Journal of clinical epidemiology* **58**, 1006-1014 (2005).
13. Ministry of Interior and Health. Health Care in Denmark 2008. [Accessed May 31, 2018]. Available from: https://www.sum.dk/English/~/media/Filer - Publikationer_i_pdf/2016/Healthcare-in-dk-16-dec/Healthcare-english-V16-dec.ashx.
14. Schmidt, M., Pedersen, L. & Sorensen, H.T. The Danish Civil Registration System as a tool in epidemiology. *European journal of epidemiology* **29**, 541-549 (2014).
15. Schmidt, M.*, et al.* The Danish National Patient Registry: a review of content, data quality, and research potential. *Clinical epidemiology* **7**, 449-490 (2015).

16. Ehrenstein, V., Antonsen, S. & Pedersen, L. Existing data sources for clinical epidemiology: Aarhus University Prescription Database. *Clinical epidemiology* **2**, 273-279 (2010).
17. Grann, A.F., Erichsen, R., Nielsen, A.G., Froslev, T. & Thomsen, R.W. Existing data sources for clinical epidemiology: The clinical laboratory information system (LABKA) research database at Aarhus University, Denmark. *Clinical epidemiology* **3**, 133-138 (2011).
18. Ording, A.G. & Sorensen, H.T. Concepts of comorbidities, multiple morbidities, complications, and their clinical epidemiologic analogs. *Clinical epidemiology* **5**, 199-203 (2013).
19. NORDCAN database. [Accessed June 07, 2018]. Available from: http://www-dep.iarc.fr/NORDCAN/DK/frame.asp.
20. National Health Board. National minimum dataset (hospital events) data dictionary. (Ministry of Health Wellington, 2014).
21. Ministry of Health. Mortality Collection Data Dictionary. (Ministry of Health Wellington, 2009).
22. Ministry of Health. Pharmaceutical Claims Data Mart data dictionary. (Ministry of Health Wellington).
23. Steyerberg, E.W. *Clinical Prediction Models - A Practical Approach to Development, Validation, and Updating*, (Springer, 2009).
24. Royston, P.S., W. *Multivariable Model-Building, A pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables*, (2008).
25. Schoenfeld, D. Partial Residuals for The Proportional Hazards Regression Model. *Biometrika*, 239-241 (1982).
26. Mehta, H.B., Mehta, V., Girman, C.J., Adhikari, D. & Johnson, M.L. Regression coefficient-based scoring system should be used to assign weights to the risk index. *Journal of clinical epidemiology* **79**, 22-28 (2016).
27. Knottnerus, J.A., Tugwell, P. & Wells, G. Editorial comment: Ratios should be multiplied, not added. *Journal of clinical epidemiology* **79**, 30 (2016).
28. Royston, P. Explained variation for survival models. *The Stata Journal* **6**, 83-96 (2006).
29. Pencina, M.J., D'Agostino, R.B., Sr., D'Agostino, R.B., Jr. & Vasan, R.S. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Statistics in medicine* **27**, 157-172; discussion 207-112 (2008).
30. Steyerberg, E.W*., et al.* Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology (Cambridge, Mass.)* **21**, 128-138 (2010).
31. Lash, T.L*., et al.* Methodology, design, and analytic techniques to address measurement of comorbid disease. *The journals of gerontology. Series A, Biological sciences and medical sciences* **62**, 281-285 (2007).
32. Granger, C.B*., et al.* Predictors of hospital mortality in the global registry of acute coronary events. *Archives of internal medicine* **163**, 2345-2353 (2003).
33. Moons, K.G., Royston, P., Vergouwe, Y., Grobbee, D.E. & Altman, D.G. Prognosis and prognostic research: what, why, and how? *BMJ (Clinical research ed.)* **338**, b375 (2009).
34. Moons, K.G*., et al.* Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Annals of internal medicine* **162**, W1-73 (2015).
35. Rothman, K.J. *Epidemiology: an introduction. 2nd ed. New York, NY: Oxford University Press 2012*.

36. Elixhauser, A., Steiner, C., Harris, D.R. & Coffey, R.M. Comorbidity measures for use with administrative data. *Medical care* **36**, 8-27 (1998).
37. Sundboll, J.*, et al.* Positive predictive value of cardiovascular diagnoses in the Danish National Patient Registry: a validation study. *BMJ open* **6**, e012832 (2016).
38. Klabunde, C.N., Potosky, A.L., Legler, J.M. & Warren, J.L. Development of a comorbidity index using physician claims data. *Journal of clinical epidemiology* **53**, 1258-1267 (2000).
39. Olesen, J.B.*, et al.* Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study. *BMJ (Clinical research ed.)* **342**, d124 (2011).
40. Sun, G.W., Shook, T.L. & Kay, G.L. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *Journal of clinical epidemiology* **49**, 907-916 (1996).
41. Schneeweiss, S. & Maclure, M. Use of comorbidity scores for control of confounding in studies using administrative databases. *International journal of epidemiology* **29**, 891-898 (2000).
42. Royston, P., Moons, K.G., Altman, D.G. & Vergouwe, Y. Prognosis and prognostic research: Developing a prognostic model. *BMJ (Clinical research ed.)* **338**, b604 (2009).
43. Justice, A.C., Covinsky, K.E. & Berlin, J.A. Assessing the generalizability of prognostic information. *Annals of internal medicine* **130**, 515-524 (1999).
44. Poses, R.M., Cebul, R.D. & Centor, R.M. Evaluating physicians' probabilistic judgments. *Medical decision making : an international journal of the Society for Medical Decision Making* **8**, 233-240 (1988).
45. Pencina, M.J., D'Agostino, R.B., Pencina, K.M., Janssens, A.C. & Greenland, P. Interpreting incremental value of markers added to risk prediction models. *American journal of epidemiology* **176**, 473-481 (2012).
46. Iezzoni, L.I.*, et al.* Comorbidities, complications, and coding bias. Does the number of diagnosis codes matter in predicting in-hospital mortality? *Jama* **267**, 2197-2203 (1992).
47. Stukenborg, G.J., Wagner, D.P. & Connors, A.F., Jr. Comparison of the performance of two comorbidity measures, with and without information from prior hospitalizations. *Medical care* **39**, 727-739 (2001).
48. Chu, Y., Ng, Y. & Wu, S. Comparison of different comorbidity measures for use with administrative data in predicting short- and long-term mortality. *BMC health services research* (2010).
49. Gutacker, N., Bloor, K. & Cookson, R. Comparing the performance of the Charlson/Deyo and Elixhauser comorbidity measures across five European countries and three conditions. *European journal of public health* **25 Suppl 1**, 15-20 (2015).
50. Yamana, H., Matsui, H., Sasabuchi, Y., Fushimi, K. & Yasunaga, H. Categorized diagnoses and procedure records in an administrative database improved mortality prediction. *Journal of clinical epidemiology* **68**, 1028-1035 (2015).

TABLES AND FIGURES

Table 1. Definitions and prevalence of DANCAMI candidate diseases and their effect on mortality in the Danish MI cohort (N=36,685)

Variables	ICD-10 codes	Prevalence (%)	Univariate HR[1] (95% CI)
Cardiovascular diseases			
Stable angina pectoris	I20.1, I20.8, I20.9, I25.1, I25.9	6,852 (19)	0.9 (0.9;0.9)
Heart failure	I50, I11.0, I13.0, I13.2	1,603 (4.4)	1.8 (1.7;1.9)
Cardiomyopathy	I25.5, I42, I43	405 (1.1)	1.4 (1.2;1.6)
Intermittent arterial claudication	I73.9	841 (2.3)	1.6 (1.4;1.8)
Aorta disease	I71	458 (1.3)	1.4 (1.2;1.7)
Deep venous thrombosis in the lower limb	I80.1-I80.3	289 (0.8)	1.1 (0.9;1.4)
Pulmonary embolism	I26.0, I26.9	186 (0.5)	1.3 (1.0;1.6)
Atrial fibrillation	I48	2,784 (7.6)	1.3 (1.2;1.4)
Heart block (atrioventricular block, left bundle-branch block, fascicular block)	144, I45	422 (1.2)	1.1 (1.0;1.3)
Ventricular tachycardia	I47.2	79 (0.2)	1.9 (1.4;2.7)
Cardiac valve disease	I05-I09, I34-I39	1,336 (3.6)	1.5 (1.4;1.7)
Stroke	I60, I61, I63, I64	1,048 (3.8)	1.6 (1.5;1.7)
Hypertension	I10-I13, I15, I67.4[2]	19,389 (53)	1.3 (1.3;1.4)
Neoplasms			
High-risk cancer	C13, C15, C16, C22-C26, C33, C34, C45, C77-C79, C92.0, C92.3-C92.9, C95	631 (1.7)	3.4 (3.1;3.8)
Lower-risk cancer	C00-C12, C14, C17-C21, C30-C32, C37-C44, C46-C76, C80-C91, C92.1, C93, C94, C96, C97	2,270 (6.2)	1.5 (1.4;1.6)
Diseases of the blood, blood-forming organs and immune system disorders			
Nutritional anemia	D50-D53	607 (1.7)	1.2 (1.1;1.4)
Coagulopathy and other blood disorders	D55-D61, D63, D64, D66-D72, D74-D77	1,783 (4.9)	1.5 (1.4;1.6)
Immune system disorder	D80-D84, D89	15 (0.0)	1.4 (0.6;3.2)
HIV	B21-B24	27 (0.1)	0.5 (0.1;3.6)
Endocrine, nutritional and metabolic diseases			
Diabetes non-complicated	E10.0, E10.1, E10.9, E11.0, E11.1, E11.9 E12.0, E12.1, E12.9 E13.0, E13.1, E13.9 E14.0, E14.1, E14.9[3]	3,141 (8.6)	1.2 (1.1;1.3)
Diabetes with end-organ damage	E10.2-E10.8, E11.2-E11.8, E12.2-E12.8, E13.2-E13.8, E14.2-E14.8, H36.0	2,108 (5.8)	1.7 (1.6;1.8)
Endocrine disorder (not diabetes)	E01-E03, E05, E06.2, E06.3, E06.5, E07, E20-27, E31, E32, E34.8, E34.9	1,096 (3.0)	1.1 (1.0;1.2)
Obesity	E65-E68	621 (1.7)	1.4 (1.2;1.6)

Mental and behavioral disorders			
Dementia	F00-F03, F05.1, G30	459 (1.3)	1.7 (1.5;1.9)
Alcohol and drug abuse	F10-F19, Z50.2, Z50.3, Z71.4, Z71.5	672 (1.8)	2.0 (1.7;2.3)
Schizophrenia	F20-F22, F25, F28, F29[4]	1,160 (3.2)	1.9 (1.8;2.1)
Affective	F30-F34, F38, F39[5]	4,766 (13)	1.6 (1.5;1.7)
Anxiety and behavioral disorder	F40-F45, F48, F50, F55, F59-F66, F68, F69	156 (0.4)	1.3 (0.9;1.8)
Diseases of the nervous system			
Transient ischemic attack	G45	859 (2.3)	1.1 (1.0; 1.3)
Epilepsy	G40, G41	334 (0.9)	1.8 (1.5; 2.1)
Atrophy, degenerative disease or demyelination of CNS	G10-G13, G20-23, G25.5, G31.2, G31.8, G31.9, G35-G37, G90, G93.4	348 (1.0)	1.5 (1.3; 1.8)
Hemiplegia	G81, G82	75 (0.2)	2.1 (1.5; 3.0)
Diseases of the genitourinary system			
Chronic kidney disease	E10.2, E11.2, E14.2, I12, I13, N03, N05, N11.0, N14, N16, N18, N19, N26.9, Q61.1-Q61.4, Z99.2	1,576 (4.3)	2.0 (1.9; 2.2)
Diseases of the respiratory system			
Chronic pulmonary disease	J40-J47, J60-J67, J68.4, J70.1, J70.3, J84.1, J92.0, J96.1, J98.2, J98.3[6]	8,015 (22)	1.4 (1.3; 1.5)
Diseases of the digestive system			
Ulcer disease	K22.1, K25-K28	1,087 (3.0)	1.5 (1.3; 1.6)
Mild liver disease	B18, K70.1-K70.3, K70.9, K71, K73, K74, K76.0	209 (0.6)	2.1 (1.7; 2.7)
Moderate to severe liver disease	B15.0, B16.0, B16.2, B19.0, K70.4, K72, K76.6, I85	66 (0.2)	3.0 (2.1; 4.2)
Inflammatory bowel disease	K50, K51	275 (0.8)	1.1 (0.9; 1.4)
Chronic pancreatitis	K86.0, K86.1	71 (0.2)	2.5 (1.7; 3.8)
Diseases of the musculoskeletal system and connective tissue			
Connective tissue disease	M05, M06, M08, M09, M30-M36, D86	1,153 (3.1)	1.2 (1.1; 1.3)
Bone disorder	M80-M83, M85, M86.3-M86.6, M88	1,428 (3.9)	1.2 (1.1; 1.3)

[1] *Estimated with a Cox regression model adjusted for sex and age*
[2] *Or prescriptions of antihypertensive drugs (A10A, A10B)*
[3] *Or HbA1c >48 mmol/L (only in Danish cohort) or prescription for antidiabetic drugs*
[4] *Or prescription for antipsychotic drug (N05A)*
[5] *Or prescription for antidepressiva drug (N06A, except N06AX12)*
[6] *Or prescription for chronic pulmonary disease drugs (R03)*

Table 2. The DANCAMI model of cardiovascular and non-cardiovascular comorbidities

Covariate	Beta coefficient	SE	Hazard ratio	95% CI	Severity weight
Heart failure	0.320	0.037	1.38	1.28; 1.48	3
Intermittent arterial claudication	0.229	0. 055	1.26	1.13; 1.40	2
Aorta disease	0.209	0. 082	1.23	1.05; 1.45	2
Cardiac valve disease	0.233	0.042	1.26	1.16; 1.37	2
Stroke	0.254	0.042	1.29	1.19; 1.40	3
Hypertension	0.121	0.025	1.13	1.08; 1.18	1
High risk cancer	1.043	0.053	2.84	2.56; 3.15	10
Lower-risk cancer	0.190	0.036	1.21	1.13; 1.30	2
Blood disorder	0.127	0.037	1.14	1.06; 1.22	1
Diabetes non-complicated	0.183	0.034	1.20	1.12; 1.28	2
Diabetes with end-organ damage	0.315	0.040	1.37	1.27; 1.48	3
Dementia	0.327	0.063	1.39	1.23; 1.57	3
Alcohol and drug abuse	0.302	0.080	1.35	1.16; 1.58	3
Schizophrenia	0.464	0.048	1.59	1.45; 1.75	5
Affective disorder	0.255	0.027	1.29	1.22; 1.36	3
Epilepsy	0.287	0.090	1.33	1.12; 1.59	3
Atrophy, degenerative disease or demyelination of CNS	0.286	0.085	1.33	1.13; 1.57	3
Hemiplegia	0.577	0.183	1.78	1.24; 2.55	6
Chronic kidney disease	0.373	0.047	1.45	1.32; 1.59	4
Chronic pulmonary disease	0.226	0.024	1.25	1.20; 1.31	2
Ulcer	0.176	0.048	1.19	1.08; 1.31	2
Mild liver disease	0.286	0.129	1.33	1.03; 1.71	3
Moderate to severe liver disease	0.664	0.190	1.94	1.34; 2.82	7
Chronic pancreatitis	0.500	0.207	1.65	1.10; 2.47	5

Table 3. The DANCAMI model restricted to non-cardiovascular comorbidities (rDANCAMI)

Covariate	Beta coefficient	SE	Hazard ratio	95% CI	Severity weight
High-risk cancer	1.041	0.053	2.83	2.55; 3.14	10
Lower-risk cancer	0.193	0.036	1.21	1.13; 1.30	2
Coagulopathy and other blood disorders	0.260	0.037	1.30	1.21; 1.39	3
Obesity	0.248	0.085	1.28	1.09; 1.51	2
Dementia	0.362	0.063	1.44	1.27; 1.62	4
Alcohol and drug abuse	0.336	0.080	1.40	1.20; 1.64	3
Schizophrenia	0.470	0.048	1.60	1.46; 1.76	5
Affective disorder	0.299	0.027	1.35	1.28; 1.42	3
Epilepsy	0.392	0.090	1.48	1.24; 1.76	4
Atrophy, degenerative disease or demyelination of CNS	0.295	0.085	1.34	1.14; 1.59	3
Hemiplegia	0.637	0.183	1.89	1.32; 2.71	6
Chronic pulmonary disease	0.265	0.024	1.30	1.24; 1.36	3
Ulcer	0.247	0.048	1.28	1.16; 1.41	2
Mild liver disease	0.359	0.130	1.43	1.11; 1.85	4
Moderate to severe liver disease	0.554	0.191	1.74	1.20; 2.53	6
Chronic pancreatitis	0.643	0.207	1.90	1.27; 2.85	6
Connective tissue disease	0.105	0.533	1.11	1.00; 1.23	1

Abbreviation: SE: Standard error, CI: Confidence interval

Table 4. Characteristics of the Danish and New Zealand Myocardial Infarction cohorts

	Denmark	New Zealand
Number of patients, n (%)	36,685 (100)	75,069 (100)
Follow-uptime, person years	29,293	63,263
1-year mortality, n (%)	8,974 (24)	14,951 (20)
In-hospital mortality, n (%)	5,014 (14)	7,095 (9.5)
Sex, n (%)		
Female	14,255 (39)	30,514 (41)
Male	22,430 (61)	44,555 (59)
Age, years, Median (IQR)	72 (61-81)	71 (59-81)
>75 years, n (%)	14,978 (41)	31,027 (41)
Prevalent comorbidities		
Most prevalent, % Second most prevalent, % Third most prevalent, %	Hypertension, 53 Chronic pulmonary disease, 22 Stable angina pectoris, 19	Hypertension, 38 Chronic pulmonary disease, 17 Diabetes with end-organ failure, 16
DANCAMI score, n (%)		
0	10,725 (29)	25,047 (33)
1-2	10,016 (27)	14,260 (19)
3-4	7,393 (20)	12,002 (16)
5+	8,551 (23)	23,760 (32)
rDANCAMI score, n (%)		
0	20,775 (57)	39,558 (53)
1-2	2,134 (6.0)	4,637 (6.2)
3-4	8,201 (22)	14,000 (19)
5+	5,575 (15)	16,874 (22)
Charlson score, n (%)		
0	21,893 (60)	37,008 (49)
1	6,515 (18)	8,633 (12)
2	4,232 (12)	11,841 (16)
3+	4,045 (11)	17,587 (23)
Elixhauser score, n (%)		
≤0	22,705 (62)	39,427 (53)
1-5	9,285 (25)	14,559 (19)
6-13	3,923 (11)	12,363 (16)
14+	772 (2.1)	8,720 (12)
Ethnicity, n (%)	NA	European, 58,315 (78) Maori, 7,544 (10) Pacific, 3,915 (5.2) Indian, 2,412 (3.2) Chinese/other Asian, 1,693 (2.3) Other, 1,190 (1.6)

Abbreviation: NA: Not available

Table 5. Overall performance and discrimination of the DANCAMI indices in the Danish MI cohort (development) and New Zealand MI cohort (validation)

	Danish patient registry cohort (95% CI)		New Zealand cohort (95% CI)	
R^2				
Baseline[1]	0.276 (0.265; 0.287)	ref.	0.282 (0.275; 0.291)	ref.
DANCAMI[2]	0.332 (0.322; 0.342)	1.20[3]	0.373 (0.366; 0.381)	1.32[3]
rDANCAMI[2]	0.318 (0.307; 0.327)	1.15[3]	0.362 (0.354; 0.369)	1.28[3]
Charlson[2]	0.315 (0.305; 0.325)	1.14[3]	0.372 (0.365; 0.379)	1.32[3]
Elixhauser[2]	0.306 (0.296; 0.315)	1.13[3]	0.375 (0.368; 0.384)	1.33[3]
Harrell's C				
Baseline[1]	0.726 (0.721; 0.731)	ref.	0.726 (0.722; 0.730)	ref.
DANCAMI[2]	0.753 (0.748; 0.758)	1.04[4]	0.773 (0.770; 0.777)	1.07[4]
rDANCAMI[2]	0.746 (0.741; 0.751)	1.03[4]	0.765 (0.762; 0.769)	1.05[4]
Charlson[2]	0.744 (0.740; 0.749)	1.03[4]	0.773 (0.770; 0.777)	1.07[4]
Elixhauser[2]	0.740 (0.735; 0.745)	1.02[4]	0.774 (0.771; 0.778)	1.07[4]
IDI				
Baseline[1] vs. DANCAMI[2]	0.054	-	0.079	-
Baseline[1] vs. rDANCAMI[2]	0.038	-	0.068	-
Baseline[1] vs. Charlson[2]	0.038	-	0.077	-
Baseline[1] vs. Elixhauser[2]	0.029	-	0.081	-
NRI				
Baseline[1] vs. DANCAMI[2]	**0.519**	-	**0.682**	-
Cases with increased probabilities	49%		56%	
Cases with decreased probabilities	51%		44%	
Controls with increased probabilities	23%		22%	
Controls with decreased probabilities	77%		78%	
Baseline[1] vs. rDANCAMI[2]	**0.428**	-	**0.573**	-
Cases with increased probabilities	54%		49%	
Cases with decreased probabilities	46%		51%	
Controls with increased probabilities	32%		23%	
Controls with decreased probabilities	68%		79%	
Baseline[1] vs. Charlson[2]	**0.409**	-	**0.582**	-
Cases with increased probabilities	39%		54%	
Cases with decreased probabilities	61%		46%	
Controls with increased probabilities	18%		25%	
Controls with decreased probabilities	82%		75%	
Baseline[1] vs. Elixhauser[2]				
Cases with increased probabilities	**0.402**	-	**0.676**	-
Cases with decreased probabilities	49%		56%	
Controls with increased probabilities	51%		44%	
Controls with decreased probabilities	29%		22%	
	71%		78%	

[1] Baseline model defined as a Cox model including sex and age
[2] All model performances were examined in a Cox model including sex, age & individual model score
[3] Difference in R^2 relative to baseline model
[4] Difference in Harrell's C relative to baseline model
95% CI for R^2 were calculated using 1000 bootstrap replications
95% CI for C-statistics were calculated using Jackknife
Abbreviation: Ref.: Reference

Table 6a. Sensitivity analysis, results for the Danish MI cohort. Performance in selected groups

	Men (95% CI)		Women (95% CI)		MI patients age>75 (95% CI)		Patients surviving hospital admission (95% CI)		Danish cohort overall (95% CI)	
R^2										
Baseline[1]	0.316 (0.301; 0.330)	Ref.	0.204 (0.189; 0.218)	Ref.	0.065 (0.056; 0.073)	Ref.	0.360 (0.345; 0.375)	Ref.	0.276 (0.265; 0.287)	Ref.
DANCAMI[2]	0.383 (0.370; 0.396)	1.21[3]	0.248 (0.235; 0.262)	1.22[3]	0.119 (0.109; 0.130)	1.83[3]	0.450 (0.436; 0.464)	1.25[3]	0.332 (0.322; 0.342)	1.20[3]
rDANCAMI[2]	0.367 (0.352; 0.380)	1.16[3]	0.236 (0.222; 0.251)	1.16[3]	0.101 (0.090; 0.111)	1.54[3]	0.431 (0.415; 0.445)	1.20[3]	0.318 (0.307; 0.327)	1.15[3]
Charlson[2]	0.361 (0.348; 0.375)	1.15[3]	0.235 (0.220; 0.249)	1.15[3]	0.101 (0.093; 0.112)	1.55[3]	0.428 (0.414; 0.442)	1.19[3]	0.315 (0.305; 0.325)	1.14[3]
Elixhauser[2]	0.348 (0.334; 0.362)	1.10[3]	0.229 (0.214; 0.244)	1.12[3]	0.093 (0.084; 0.103)	1.43[3]	0.406 (0.393; 0.420)	1.13[3]	0.306 (0.296; 0.315)	1.11[3]
Harrell's C										
Baseline[1]	0.741 (0.734; 0.748)	Ref.	0.688 (0.681; 0.696)	Ref.	0.597 (0.590; 0.604)	Ref.	0.764 (0.757; 0.771)	Ref.	0.726 (0.721; 0.731)	Ref.
DANCAMI[2]	0.772 (0.766; 0.779)	1.04[4]	0.712 (0.705; 0.719)	1.03[4]	0.630 (0.623; 0.637)	1.06[4]	0.808 (0.802; 0.814)	1.06[4]	0.7523 (0.748; 0.758)	1.04[4]
rDANCAMI[2]	0.764 (0.758; 0.771)	1.03[4]	0.706 (0.698; 0.713)	1.02[4]	0.619 (0.613; 0.626)	1.04[4]	0.797 (0.790; 0.803)	1.04[4]	0.746 (0.741; 0.751)	1.03[4]
Charlson[2]	0.764 (0.757; 0.770)	1.03[4]	0.703 (0.695; 710)	1.02[4]	0.619 (0.612; 0.626)	1.04[4]	0.801 (0.794; 0.807)	1.05[4]	0.744 (0.740; 0.749)	1.03[4]
Elixhauser[2]	0.757 (0.750; 0.764)	1.02[4]	0.701 (0.694; 708)	1.02[4]	0.615 (0.608; 0.622)	1.03[4]	0.790 (0.784; 0.797)	1.03[4]	0.740 (0.735; 0.745)	1.02[4]

[1] Baseline model defined as a Cox model including sex and age
[2] All model performances were examined in a Cox model including sex, age & individual model score
[3] Difference in R^2 relative to baseline model
[4] Difference in Harrell's C relative to baseline model
95% CI for R^2 were calculated using 1000 bootstrap replications
95% CI for C-statistics were calculated using Jackknife
Abbreviation: Ref.: Reference

Table 6b. Sensitivity analysis, results for the Danish MI cohort. Performance in selected groups

	Men	Women	MI patients age >75	Surviving hospital admission	Danish cohort overall
IDI					
Baseline[1] vs. DANCAMI[2]	0.064	0.043	0.049	0.064	0.054
Baseline[1] vs. rDANCAMI[2]	0.047	0.030	0.031	0.046	0.038
Baseline[1] vs. Charlson[2]	0.045	0.030	0.034	0.048	0.038
Baseline[1] vs. Elixhauser[2]	0.032	0.025	0.026	0.035	0.029
NRI					
Baseline[1] vs. DANCAMI[2]	**0.552**	**0.478**	**0.392**	**0.636**	**0.519**
Cases with increased probabilities	50%	52%	50%	52%	49%
Cases with decreased probabilities	50%	48%	50%	48%	51%
Controls with increased probabilities	23%	28%	30%	20%	23%
Controls with decreased probabilities	77%	72%	70%	80%	77%
Baseline[1] vs. rDANCAMI[2]	**0.484**	**0.391**	**0.265**	**0.546**	**0.428**
Cases with increased probabilities	52%	53%	53%	57%	54%
Cases with decreased probabilities	48%	47%	47%	43%	46%
Controls with increased probabilities	28%	34%	40%	30%	32%
Controls with decreased probabilities	72%	66%	60%	70%	68%
Baseline[1] vs. Charlson[2]	**0.443**	**0.434**	**0.341**	**0.539**	**0.409**
Cases with increased probabilities	44%	45%	46%	44%	39%
Cases with decreased probabilities	56%	55%	54%	56%	61%
Controls with increased probabilities	22%	24%	29%	17%	18%
Controls with decreased probabilities	78%	76%	71%	83%	82%
Baseline[1] vs. Elixhauser[2]	**0.427**	**0.284**	**0.266**	**0.475**	**0.402**
Cases with increased probabilities	47%	49%	46%	52%	49%
Cases with decreased probabilities	53%	51%	54%	48%	51%
Controls with increased probabilities	25%	35%	33%	28%	29%
Controls with decreased probabilities	75%	65%	67%	72%	71%

[1] Baseline model defined as a Cox model including sex and age
[2] All model performances were examined in a Cox model including sex, age & individual model score

Table 7a. Sensitivity analysis, results for the New Zealand MI cohort. Performance in selected groups

	Men		Women		MI patients age>75		Patients surviving hospital admission		New Zealand cohort overall	
R^2										
Baseline[1]	0.320 (0.309; 0.331)	ref.	0.220 (0.209; 0.232)	ref.	0.081 (0.074; 0.088)	ref.	0.338 (0.327; 0.348)	ref.	0.282 (0.275; 0.291)	ref.
DANCAMI[2]	0.415 (0.405; 0.425)	1.30[3]	0.307 (0.295; 0.317)	1.40[3]	0.175 (0.165; 0.184)	2.16[3]	0.446 (0.437; 0.456)	1.32[3]	0.373 (0.366; 0.381)	1.32[3]
rDANCAMI[2]	0.406 (0.395; 0.416)	1.27[3]	0.290 (0.280; 0.301)	1.32[3]	0.160 (0.151; 0.170)	1.98[3]	0.438 (0.428; 0.447)	1.30[3]	0.362 (0.354; 0.369)	1.28[3]
Charlson[2]	0.410 (0.399; 0.421)	1.28[3]	0.310 (0.299; 0.321)	1.41[3]	0.172 (0.163; 0.182)	2.12[3]	0.438 (0.429; 0.447)	1.30[3]	0.372 (0.365; 0.379)	1.32[3]
Elixhauser[2]	0.417 (0.406; 0.427)	1.30[3]	0.308 (0.297; 0.320)	1.40[3]	0.173 (0.163; 0.182)	2.14[3]	0.436 (0.427; 0.446)	1.29[3]	0.375 (0.368; 0.384)	1.33[3]
Harrell's C										
Baseline[1]	0.741 (0.735; 0.746)	Ref.	0.693 (0.687; 0.699)	Ref.	0.606 (0.601; 0.612)	Ref.	0.754 (0.749; 0.759)	Ref.	0.726 (0.722; 0.730)	Ref.
DANCAMI[2]	0.791 (0.786; 0.796)	1.07[4]	0.739 (0.734; 0.744)	1.07[4]	0.663 (0.658; 0.668)	1.09[4]	0.812 (0.808; 0.816)	1.08[4]	0.773 (0.770; 0.777)	1.07[4]
rDANCAMI[2]	0.783 (0.778; 0.788)	1.06[4]	0.731 (0.725; 0.736)	1.05[4]	0.654 (0.649; 0.659)	1.08[4]	0.805 (0.800; 0.809)	1.07[4]	0.765 (0.762; 0.769)	1.05[4]
Charlson[2]	0.789 (0.784; 0.794)	1.07[4]	0.742 (0.737; 0.747)	1.07[4]	0.662 (0.657; 0.667)	1.09[4]	0.808 (0.804; 0.813)	1.07[4]	0.773 (0.770; 0.777)	1.07[4]
Elixhauser[2]	0.792 (0.787; 0.797)	1.07[4]	0.740 (0.735; 0.745)	1.07[4]	0.662 (0.657; 0.667)	1.09[4]	0.807 (0.803; 0.812)	1.07[4]	0.774 (0.771; 0.778)	1.07[4]

[1] Baseline model defined as a Cox model including sex and age
[2] All model performances were examined in a Cox model including sex, age & individual model score
[3] Difference in R^2 relative to baseline model
[4] Difference in Harrell's C relative to baseline model
95% CI for R^2 were calculated using 1000 bootstrap replications
95% CI for C-statistics were calculated using Jackknife
Abbreviation: Ref.: Reference

Table 7b. Sensitivity analysis, results for the New Zealand MI cohort. Performance in selected groups

	Men	Women	MI patients age >75	Surviving hospital admission	New Zealand cohort overall
IDI					
Baseline[1] vs. DANCAMI[2]	0.088	0.071	0.070	0.073	0.079
Baseline[1] vs. rDANCAMI[2]	0.078	0.058	0.058	0.065	0.068
Baseline[1] vs. Charlson[2]	0.082	0.074	0.067	0.065	0.077
Baseline[1] vs. Elixhauser[2]	0.091	0.072	0.067	0.067	0.081
NRI					
Baseline[1] vs. DANCAMI[2]	**0.721**	**0.590**	**0.503**	**0.734**	**0.682**
Cases with increased probabilities	57%	55%	54%	58%	56%
Cases with decreased probabilities	43%	45%	46%	42%	44%
Controls with increased probabilities	21%	25%	29%	22%	22%
Controls with decreased probabilities	79%	75%	71%	78%	78%
Baseline[1] vs. rDANCAMI[2]	**0.579**	**0.512**	**0.427**	**0.640**	**0.573**
Cases with increased probabilities	51%	50%	48%	51%	49%
Cases with decreased probabilities	49%	50%	52%	49%	51%
Controls with increased probabilities	22%	24%	27%	19%	22%
Controls with decreased probabilities	78%	76%	73%	81%	78%
Baseline[1] vs. Charlson[2]	**0.662**	**0.598**	**0.503**	**0.604**	**0.582**
Cases with increased probabilities	57%	50%	54%	54%	54%
Cases with decreased probabilities	43%	50%	46%	46%	46%
Controls with increased probabilities	24%	20%	28%	24%	25%
Controls with decreased probabilities	76%	80%	72%	76%	75%
Baseline[1] vs. Elixhauser[2]	**0.732**	**0.610**	**0.486**	**0.681**	**0.676**
Cases with increased probabilities	57%	54%	52%	56%	56%
Cases with decreased probabilities	43%	46%	48%	44%	44%
Controls with increased probabilities	20%	24%	28%	22%	22%
Controls with decreased probabilities	80%	76%	72%	78%	78%

[1] Baseline model defined as a Cox model including sex and age
[2] All model performances were examined in a Cox model including sex, age & individual model score

Table 8a. Sensitivity analysis, results for the New Zealand MI cohort. Performance in ethnicity groups

	European/ Other		Maori		Pacific		Indian		Chinese/ Other Asian	
R^2										
Baseline[1]	0.299 (0.290; 0.308)	ref.	0.227 (0.198; 0.255)	ref.	0.229 (0.193; 0.272)	ref.	0.350 (0.292; 0.413)	ref.	0.337 (0.245; 0.435)	ref.
DANCAMI[2]	0.385 (0.377; 0.393)	1.29[3]	0.347 (0.323; 0.372)	1.53[3]	0.341 (0.306; 0.375)	1.49[3]	0.399 (0.342; 0.453)	1.14[3]	0.391 (0.309; 0.487)	1.16
rDANCAMI[2]	0.377 (0.369; 0.386)	1.26[3]	0.330 (0.304; 0.356)	1.45[3]	0.308 (0.270; 0.345)	1.34[3]	0.387 (0.328; 0.446)	1.11[3]	0.381 (0.293; 0.475)	1.13
Charlson[2]	0.380 (0.372; 0.388)	1.27[3]	0.361 (0.337; 0.385)	1.59[3]	0.342 (0.307; 0.383)	1.49[3]	0.425 (0.371; 0.485)	1.21[3]	0.384 (0.303; 0.482)	1.14
Elixhauser[2]	0.383 (0.374; 0.391)	1.28[3]	0.361 (0.338; 0.387)	1.59[3]	0.368 (0.336; 0.403)	1.61[3]	0.445 (0.394; 0.501)	1.27[3]	0.416 (0.336; 0.518)	1.23
Harrell's C										
Baseline[1]	0.737 (0.733; 0.741)	Ref.	0.689 (0.675; 0.703)	Ref.	0.697 (0.677; 0.716)	Ref.	0.752 (0.723; 0.781)	Ref.	0.744 (0.694; 0.794)	Ref.
DANCAMI[2]	0.780 (0.776; 0.784)	1.06[4]	0.758 (0.745; 0.770)	1.10[4]	0.755 (0.738; 0.773)	1.08[4]	0.776 (0.748; 0.805)	1.03[4]	0.772 (0.724; 0.821)	1.04
rDANCAMI[2]	0.775 (0.772; 0.779)	1.05[4]	0.744 (0.731; 0.757)	1.08[4]	0.740 (0.722; 0.759)	1.06[4]	0.769 (0.740; 0.798)	1.02[4]	0.770 (0.721; 0.818)	1.03
Charlson[2]	0.779 (0.775; 0.783)	1.06[4]	0.763 (0.751; 0.775)	1.11[4]	0.753 (0.735; 0.770)	1.08[4]	0.786 (0.758; 0.813)	1.05[4]	0.771 (0.723; 0.819)	1.04
Elixhauser[2]	0.779 (0.776; 0.783)	1.06[4]	0.764 (0.752; 0.776)	1.11[4]	0.766 (0.749; 0.783)	1.10[4]	0.797 (0.770; 0.823)	1.06[4]	0.783 (0.738; 0.828)	1.05

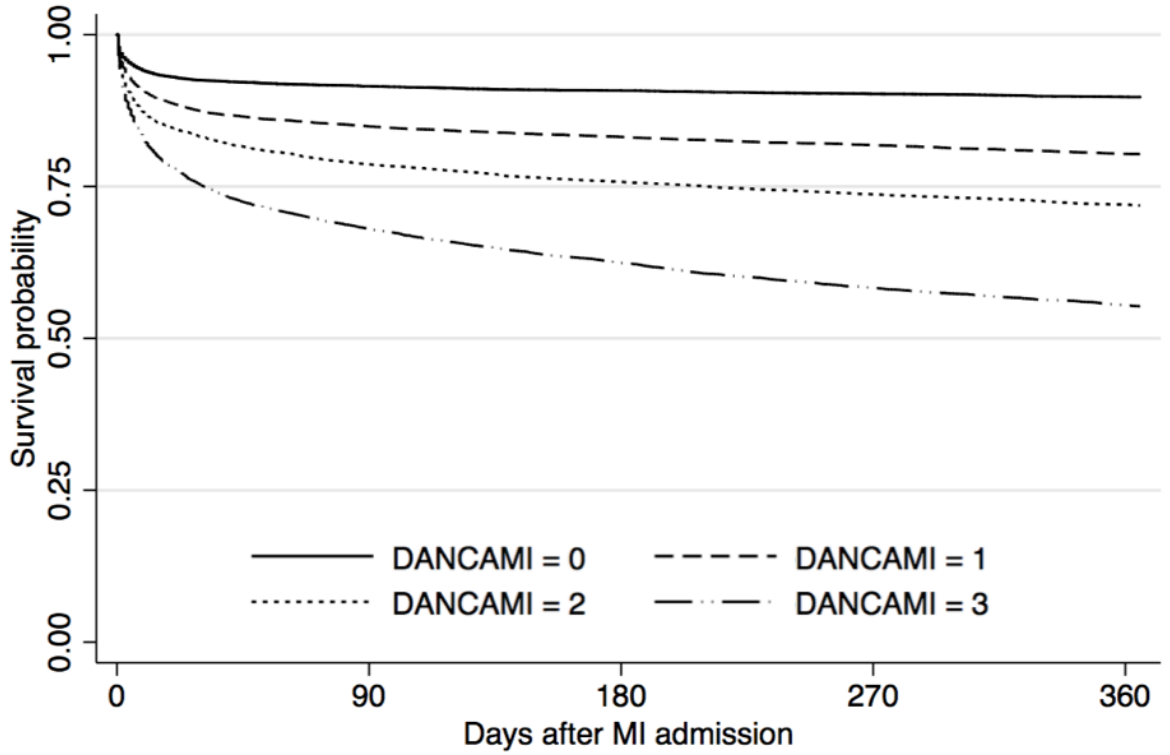
[1] Baseline model defined as a Cox model including sex and age
[2] All model performances were examined in a Cox model including sex, age & individual model score
[3] Difference in R^2 relative to baseline model
[4] Difference in Harrell's C relative to baseline model
95% CI for R^2 were calculated using 1000 bootstrap replications
95% CI for C-statistics were calculated using Jackknife
Abbreviation: Ref.: Reference

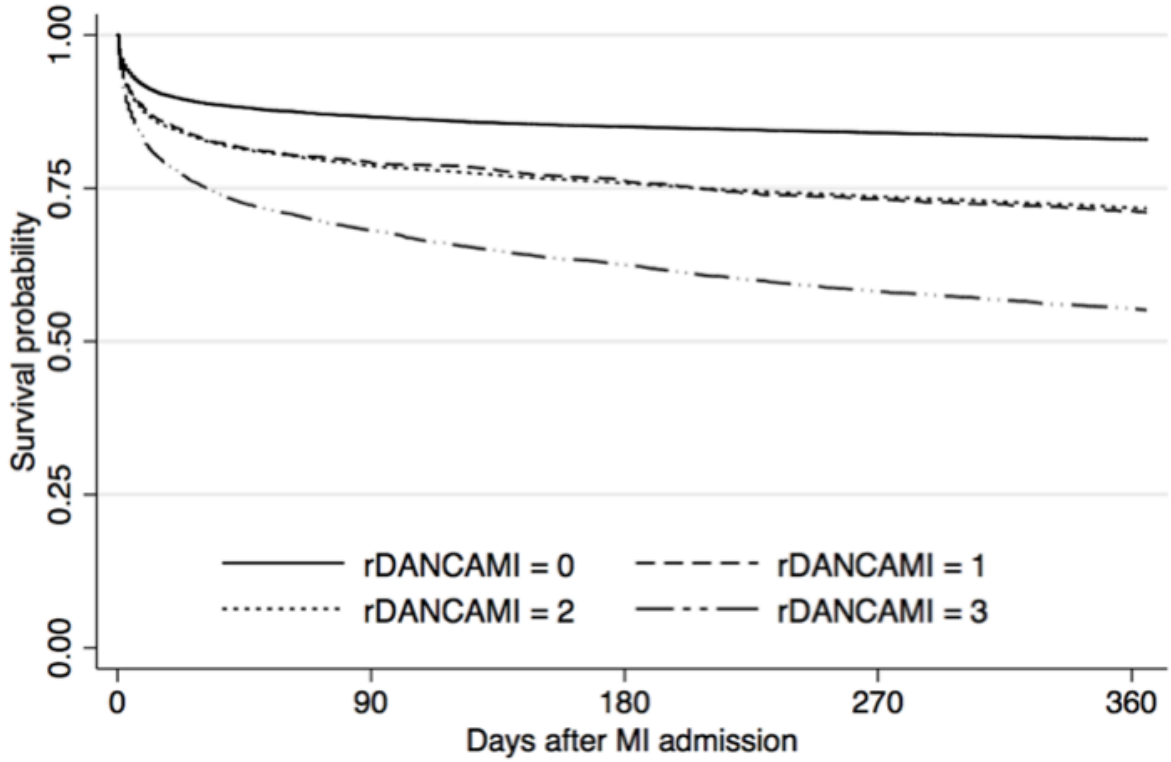
Table 8b. Sensitivity analysis, results for the New Zealand MI cohort. Performance in ethnicity groups

	European/ Other	Maori	Pacific	Indian	Chinese/ Other Asian
IDI					
Baseline[1] vs. DANCAMI[2]	0.078	0.088	0.078	0.037	0.036
Baseline[1] vs. rDANCAMI[2]	0.069	0.071	0.057	0.031	0.031
Baseline[1] vs. Charlson[2]	0.073	0.099	0.074	0.051	0.029
Baseline[1] vs. Elixhauser[2]	0.076	0.102	0.095	0.077	0.050
NRI					
Baseline[1] vs. DANCAMI[2]	**0.688**	**0.732**	**0.639**	**0.570**	**0.549**
Cases with increased probabilities	56%	61%	61%	55%	51%
Cases with decreased probabilities	44%	39%	39%	45%	49%
Controls with increased probabilities	21%	25%	29%	26%	23%
Controls with decreased probabilities	79%	75%	71%	74%	77%
Baseline[1] vs. rDANCAMI[2]	**0.608**	**0.590**	**0.477**	**0.405**	**0.451**
Cases with increased probabilities	50%	52%	48%	47%	44%
Cases with decreased probabilities	50%	48%	52%	53%	56%
Controls with increased probabilities	20%	23%	24%	27%	22%
Controls with decreased probabilities	80%	77%	76%	73%	78%
Baseline[1] vs. Charlson[2]	**0.610**	**0.728**	**0.617**	**0.599**	**0.415**
Cases with increased probabilities	55%	61%	59%	59%	46%
Cases with decreased probabilities	45%	39%	41%	41%	54%
Controls with increased probabilities	24%	24%	28%	29%	25%
Controls with decreased probabilities	76%	76%	72%	71%	75%
Baseline[1] vs. Elixhauser[2]	**0.665**	**0.748**	**0.652**	**0.639**	**0.683**
Cases with increased probabilities	55%	60%	56%	53%	55%
Cases with decreased probabilities	45%	40%	44%	47%	45%
Controls with increased probabilities	22%	23%	24%	21%	21%
Controls with decreased probabilities	78%	77%	76%	79%	79%

[1] *Baseline model defined as a Cox model including sex and age*
[2] *All model performances were examined in a Cox model including sex, age & individual model score*

Figure 1. Survival according to DANCAMI score[1] (top) and rDANCAMI score[1] (bottom)





[1] *Score: 0=0 comorbidities, 1=1-2 comorbidities, 2=3-4 comorbidities, 3=5+ comorbidities)*

Supplementary table A. Sensitivity analysis, results for the Danish MI cohort. Alternative Hazard Ratio threshold for inclusion in final models: 1.20, Precise Beta coefficients as score components, and Hazard Ratio instead of Beta-coefficients as score components

	HR threshold for inclusion: 1.20 (95% CI)	Precise Beta coefficients (95% CI)	HR instead of beta coefficients (95% CI)
R^2			
DANCAMI[2]	0.332 (0.322; 0.342)	0.333 (0.324; 0.343)	0.327 (0.317; 0.337)
rDANCAMI[2]	0.318 (0.308; 0.327)	0.318 (0.309; 0.328)	0.314 (0.303; 0.323)
Harrell's C			
DANCAMI[2]	0.752 (0.748; 0.757)	0.753 (0.748; 0.758)	0.751 (0.746; 0.755)
rDANCAMI[2]	0.746 (0.741; 0.751)	0.746 (0.741; 0.751)	0.745 (0.740; 0.749)
IDI			
Baseline[1] vs. DANCAMI[2]	0.053	0.054	0.049
Baseline[1] vs. rDANCAMI[2]	0.038	0.039	0.035
NRI			
Baseline[1] vs. DANCAMI[2]	**0.512**	**0.507**	**0.497**
Cases with increased probabilities	48%	49%	55%
Cases with decreased probabilities	52%	51%	45%
Control with increased probabilities	23%	24%	30%
Controls with decreased probabilities	77%	76%	70%
Baseline[1] vs. rDANCAMI[2]	**0.437**	**0.440**	**0.451**
Cases with increased probabilities	55%	54%	60%
Cases with decreased probabilities	45%	46%	40%
Controls with increased probabilities	33%	32%	37%
Controls with decreased probabilities	67%	68%	63%

[1] Baseline model defined as a Cox model including sex and age
[2] All model performances were examined in a Cox model including sex, age & individual model score
95% CI for R^2 were calculated using 1000 bootstrap replications
95% CI for C-statistics were calculated using Jackknife
Abbreviation: HR: Hazard Ratio

Supplementary table B. Split-sample validation, characteristics of the development subcohort (2000-2009) and validation subcohort (2010-2013)

	Total Danish MI cohort (2000-2013)	**Development subcohort (2000-2009)**	**Validation subcohort (2010-2013)**
Number of patients	36,685	27,716	8,969
Follow-uptime, person years	29,293	21,837	7456
1-year mortality, n (%)	8,974 (24)	7,118 (26)	1,856 (21)
In-hospital mortality, n (%)	5,104 (14)	4,136 (15)	968 (11)
Sex, n (%)			
Female	14,255 (39)	10,873 (39)	3,382 (38)
Male	22,430 (61)	16,843 (61)	5,587 (62)
Age, y, Median (IQR)	72 (61-81)	72 (61-81)	71 (60-80)
Prevalent comorbidities			
Most prevalent, %	Hypertension, 53	Hypertension, 52	Hypertension, 54
Second most prevalent, %	Chronic pulmonary disease, 22	Chronic pulmonary disease, 22	Chronic pulmonary disease, 23
Third most prevalent, %	Stable angina pectoris, 19	Stable angina pectoris, 19	Stable angina pectoris, 17
DANCAMI score, n (%)			
0	10,725 (29)	8,213 (30)	2,512 (28)
1-2	10,016 (27)	7,684 (28)	2,332 (26)
3-4	7,393 (20)	5,537 (20)	1,856 (21)
5+	8,551 (23)	6,282 (23)	2,269 (25)
rDANCAMI score, n (%)			
0	20,775 (57)	15,855 (57)	4,920 (55)
1-2	2,134 (5.98)	1,600 (5.8)	534 (6.0)
3-4	8,201 (22)	6,143 (22)	2,058 (23)
5+	5,575 (15)	4,118 (15)	1,457 (16)
Charlson score, n (%)			
0	21,893 (60)	16,543 (60)	5,350 (60)
1	6,515 (18)	5,019 (18)	1,496 (17)
2	4,232 (12)	3,176 (11)	1,056 (12)
3+	4,045 (11)	2,978 (11)	1,067 (12)
Elixhauser score, n (%)			
≤0	22,705 (62)	17,291 (62)	5,414 (60)
1-5	9,285 (25)	6,927 (25)	2,358 (26)
6-13	3,923 (11)	2,936 (11)	987 (11)
14+	772 (2.1)	562 (2.0)	210 (2.3)

Supplementary table C. Split-sample validation in Danish MI cohort, performance of refitted DANCAMI in a Danish MI validation subcohort

	Split sample models (2000-2009) tested in validation cohort (2010-2013)	
R^2		
Baseline[1]	0.312 (0.293; 0.335)	ref.
DANCAMI[2]	0.367 (0.348; 0.388)	1.18[3]
rDANCAMI[2]	0.349 (0.329; 0.371)	1.12[3]
Charlson[2]	0.359 (0.338; 0.379)	1.15[3]
Elixhauser[2]	0.352 (0.331; 0.374)	1.13[3]
Harrell's C		
Baseline[1]	0.742 (0.731; 0.753)	ref.
DANCAMI[2]	0.774 (0.764; 0.784)	1.04[4]
rDANCAMI[2]	0.764 (0.753; 0.774)	1.03[4]
Charlson	0.766 (0.756; 0.776)	1.03[4]
Elixhauser[2]	0.759 (0.748; 0.769)	1.02[4]
IDI		
Baseline[1] vs. DANCAMI[2]	0.059	-
Baseline[1] vs. rDANCAMI[2]	0.036	-
Baseline[1] vs. Charlson[2]	0.044	-
Baseline[1] vs. Elixhauser[2]	0.034	-
NRI		
Baseline[1] vs. DANCAMI[2]	**0.588**	
Cases with increased probabilities	51%	
Cases with decreased probabilities	49%	
Controls with increased probabilities	22%	
Controls with decreased probabilities	78%	
Baseline[1] vs. rDANCAMI[2]	**0.406**	
Cases with increased probabilities	47%	
Cases with decreased probabilities	53%	
Controls with increased probabilities	27%	
Controls with decreased probabilities	73%	
Baseline[1] vs. Charlson[2]	**0.475**	
Cases with increased probabilities	43%	
Cases with decreased probabilities	57%	
Controls with increased probabilities	19%	
Controls with decreased probabilities	81%	
Baseline[1] vs. Elixhauser[2]	**0.357**	
Cases with increased probabilities	47%	
Cases with decreased probabilities	53%	
Controls with increased probabilities	29%	
Controls with decreased probabilities	71%	

[1] Baseline model defined as a Cox model including sex and age
[2] All model performances were examined in a Cox model including sex, age & individual model score
[3] Difference in R^2 relative to baseline model
[4] Difference in Harrell's C relative to baseline model
95% CI for R^2 were calculated using 1000 bootstrap replications
95% CI for C-statistics were calculated using Jackknife
Abbreviation: Ref.: Reference