

Validity of coronary heart diseases and heart failure based on hospital discharge and mortality data in the Netherlands using the cardiovascular registry Maastricht cohort study

Audrey H. H. Merry · Jolanda M. A. Boer · Leo J. Schouten ·
Edith J. M. Feskens · W. M. Monique Verschuren ·
Anton P. M. Gorgels · Piet A. van den Brandt

Received: 11 November 2008 / Accepted: 11 March 2009 / Published online: 1 April 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract Incidence rates of cardiovascular diseases are often estimated by linkage to hospital discharge and mortality registries. The validity depends on the quality of the registries and the linkage. Therefore, we validated incidence rates of coronary heart disease (CHD), acute myocardial infarction, unstable angina pectoris, and heart failure, estimated by this method, against the disease registry of the cardiovascular registry Maastricht cohort study. The cohort consists of 21,148 persons, born between 1927 and 1977, who were randomly sampled from Maastricht and surrounding communities in 1987–1997. Incident cases were identified by linkage to the Netherlands causes of death registry and either the hospital discharge registry (HDR) or the cardiology information system (CIS) of the University Hospital Maastricht. Sensitivities and positive

predictive values were calculated using the CIS-based registry as gold standard. Relatively high sensitivities and positive predictive values were found for CHD (72 and 91%, respectively) and acute myocardial infarction (84 and 97%, respectively). These values were considerably lower for unstable angina pectoris (53 and 78%, respectively) and heart failure (43 and 80%, respectively). A substantial number of cases (14–47%) were found only in the CIS-based registry, because they were missed or miscoded in the HDR-based registry. As a consequence, the incidence rates in the HDR-based registry were considerably lower than in the CIS-based registry, especially for unstable angina pectoris and heart failure. Incidence rates based on hospital discharge and mortality data may underestimate the true incidence rates, especially for unstable angina pectoris and heart failure.

A. H. H. Merry (✉) · P. A. van den Brandt
Department of Epidemiology, School for Public Health and Primary Care (Caphri), Maastricht University, PO Box 616,
6200 MD Maastricht, The Netherlands
e-mail: audrey.merry@epid.unimaas.nl

J. M. A. Boer · W. M. M. Verschuren
National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

L. J. Schouten · P. A. van den Brandt
Department of Epidemiology, School for Oncology and Developmental Biology (GROW), Maastricht University,
Maastricht, The Netherlands

E. J. M. Feskens
Division of Human Nutrition, Wageningen University,
Wageningen, The Netherlands

A. P. M. Gorgels
Department of Cardiology, University Hospital Maastricht,
Maastricht, The Netherlands

Keywords Coronary Heart Disease · Heart Failure ·
Hospital discharge diagnoses · Mortality ·
Sudden Cardiac Arrest · Validity

Abbreviations

CABG	Coronary artery bypass grafting
CAREMA	Cardiovascular registry Maastricht
CHD	Coronary heart disease
CI	Confidence interval
CIS	Cardiology information system
HDR	Hospital discharge registry
HIS	Hospital information system
ICD	International classification of diseases
MCAR	Maastricht circulatory arrest registry
MORGEN	Monitoring project on chronic disease risk factors
N	Number

PPHVZ	Monitoring project on cardiovascular risk factors
PPV	Positive predictive value
PTCA	Percutaneous transluminal coronary angioplasty
Sd	Standard deviation
Se	Sensitivity
UHM	University hospital Maastricht
Yrs	Years

Introduction

Coronary heart disease (CHD) is the major cause of death in developed countries. In the last decades, the mortality rate of CHD has declined in the United States and Western Europe including the Netherlands [1–4]. Two possible explanations for this decreasing mortality rate are a decline in the population risk of CHD leading to a lower incidence, or a better survival of cases with CHD resulting in a lower case-fatality rate. Information about the trend in the incidence rate could be used to distinguish between these two explanations [5].

Because cardiovascular disease registries are lacking in most countries, record linkage with national hospital discharge and mortality data is often used to estimate the incidence of CHD [5, 6]. Recently, Koek et al. [7] estimated the incidence rate of a first acute myocardial infarction in 2000 by record linkage to the Dutch hospital discharge register (HDR) and causes of death registry from Statistics Netherlands. They found a crude incidence rate (per 100,000 persons per year) of 293 in men and 174 in women.

The validity of these estimates, however, depends on the completeness and accuracy of the data in the national registers and the accuracy of the linkage. Therefore, several studies have investigated the validity of data about CHD or acute myocardial infarction in national registers by comparing them with specific study registers [8–15]. These studies showed a wide range in the estimated values for the validity of the data in national registers.

Less is known about the validity of incidence estimates using record linkage with national hospital discharge and mortality data for unstable angina pectoris and heart failure. These estimates may be more problematic, because the diagnoses of these diseases are more difficult to make. Only two studies investigated the validity of the diagnosis of heart failure in national HDRs [16, 17]. Both studies indeed showed lower values for the validity of this diagnosis in national registers.

In this study, we used the disease registry of the cardiovascular registry Maastricht (CAREMA) cohort study

to estimate the incidence rates of CHD, acute myocardial infarction, unstable angina pectoris, heart failure, and sudden cardiac arrest. We compared these incidence rates with the incidence rates estimated using hospital discharge data to assess the completeness and validity of the latter.

Materials and methods

Study population

The CAREMA cohort consists of participants of two large monitoring projects in the Netherlands living in the Maastricht region: the monitoring project on cardiovascular risk factors (PPHVZ) 1987–1991 [18] and the monitoring project on chronic disease risk factors (MORGEN Project) 1993–1997, [19] including the transition year (1992) between these projects. Each year, a random sample of people aged 20–59 years was selected from the municipal registries of Maastricht and surrounding communities: Eijsden, Margraten, Meerssen, and Valkenburg aan de Geul. Between 1987 and 1997, 21,662 men and women, born between 1927 and 1977, were included in this study, of whom 21,148 participants (97.6%) had given informed consent to retrieve information from the municipal population registries and from their general practitioner and specialist.

Follow up

Migration and mortality follow-up

A migration and mortality follow-up was performed by record linkage of the CAREMA cohort to the municipal population registries. During follow-up until 31 December 2003, 2,106 persons (10.0%) had migrated to a municipality outside the Maastricht region, 621 persons (2.9%) had emigrated, and 791 persons (3.7%) had died. Furthermore, 12 persons (0.1%) were lost to follow-up, of whom 9 persons appeared to have migrated out of the Netherlands just before their baseline study date.

Cardiologic follow-up

Cardiologic follow-up was performed by record linkage of the CAREMA cohort to several hospital registries of the University Hospital Maastricht (UHM). In April 2004, the cohort was linked to the hospital information system (HIS) of the UHM using a combination of date of birth, gender and the first four characters of the family name [20]. In the HIS, 20,632 cohort members (97.6%) could be found.

Subsequently, these subjects were linked to the cardiology information system (CIS) of the UHM department of cardiology using the personal identification number of the HIS as identifier. For all people that visited the UHM department of cardiology, the CIS contains all reports to the general practitioner and information from visits to the emergency ward or outpatient clinic for heart problems, hospital admissions for cardiologic diseases, physical examinations and treatments. Among the 20,632 persons, 4,694 (22.8%) were known in the CIS. The cardiologic history of these persons was abstracted and coded by trained registrars under guidance of a cardiologist (AG).

Furthermore, the CAREMA cohort was linked to the Maastricht circulatory arrest registry (MCAR) [21] of the UHM department of cardiology to identify people who suffered from a sudden cardiac arrest.

For participants who died, the cause of death was obtained from Statistics Netherlands. Causes of death have been coded according to the ninth revision of the international classification of diseases (ICD-9) until 1996, and thereafter according to the tenth revision (ICD-10). Among the 791 deceased cohort members, 276 persons (34.9%) had a cardiovascular disease as primary or secondary cause of death (ICD-9 390–459; ICD-10 I00–I99). The cause of death was unknown for 24 cohort members (3.0%) who died outside the Netherlands, while five persons (0.6%) could not be linked to the causes of death registry from Statistics Netherlands.

The following data was registered in the CIS-based registry: date of migration to a municipality outside the Maastricht region, date of emigration, date of death including cause of death, and the presence of a clinical diagnosis including date of diagnosis and several other characteristics of an acute or silent myocardial infarction, unstable or stable angina pectoris, heart failure, atrial fibrillation, ventricular fibrillation/tachycardia, several cardiologic treatments, and sudden cardiac arrest. All data were checked for completeness, possible errors, and inconsistencies.

In addition, the CAREMA cohort was linked to the HDR of the UHM to enlarge the completeness of the cardiologic follow-up. In the HDR, the discharge diagnoses of all admissions to the UHM have been registered using the ninth revision of the international classification of diseases (ICD-9-CM). By this linkage, only four participants were found with a discharge diagnosis of CHD (ICD-9 codes 410–414) in the HDR that were not linked to the CIS. After checking their medical history, they were no additional cases for the analyses.

Because some delay might have occurred in the registration of events in the hospital registries, the follow-up was censored at 31 December 2003 to ensure the completeness of the follow-up.

Statistical analyses

Incidence estimates

In the present study, CHD is defined as incident acute myocardial infarction, unstable angina pectoris, coronary artery bypass grafting (CABG), percutaneous transluminal coronary angioplasty (PTCA), or CHD death.

Incident cases were defined in two ways, i.e., based on causes of death and either the CIS or the HDR. Persons with cardiac diseases as primary or secondary cause of death according to Statistics Netherlands were defined as cases using the following ICD-codes: ICD-9 410–414 and ICD-10 I20–I25 for CHD; ICD-9 410 and ICD-10 I21–I22 for acute myocardial infarction; ICD-9 413 and ICD-10 I20 for unstable angina pectoris; ICD-9 428 and ICD-10 I50 for heart failure; and ICD-9 798 and ICD-10 I46, R96, and R98 for sudden cardiac arrest. In addition, cases were defined according to the clinical diagnosis of the disease, made by experienced cardiologists, as extracted from CIS for the CIS-based definition. This clinical diagnosis was mostly based on the diagnosis mentioned in the report to the general practitioner. Furthermore, additional information, such as enzyme levels, ECG and echo findings, was recorded in CIS and was used to check whether the patient's clinical signs and symptoms were in agreement with this diagnosis. For the HDR-based definition, cases were defined according to their hospital discharge diagnosis using the following ICD-9 codes: 410, 411.1, and 413.1 for CHD; 410 for acute myocardial infarction; 411.1 and 413.1 for unstable angina pectoris; and 428 for heart failure.

During follow-up, participants may have had multiple cardiac diseases. For each disease separately, incident cases were defined according to the first occurrence of that disease, irrespective of the occurrence of other diseases investigated in this study. For this reason, the sum of cases with an acute myocardial infarction and cases with unstable angina pectoris is higher than the total number of CHD cases.

For each cardiac disease separately, person time at risk was calculated from baseline until end of follow-up i.e., clinical diagnosis of the disease in case of the CIS-based definition and date of hospital admission in case of the HDR-based definition, migration to a municipality outside the Maastricht region, emigration, death or censoring at 31 December 2003, whichever occurred first. Incidence rates were calculated as the number of incident cases divided by the disease-specific person time at risk.

In the analyses, participants with a migration date to a municipality outside the Maastricht region before their baseline study date ($n = 26$) and participants lost to follow-up ($n = 12$) were excluded, leaving 21,110 cohort

members. In addition, cases with CHD at baseline ($n = 347$), based on self-report or diagnosis date before baseline in the CIS-based registry, were excluded in the analyses of CHD, acute myocardial infarction, and unstable angina pectoris. Cases with heart failure at baseline ($n = 7$), based on a diagnosis date before baseline in the CIS-based registry, were excluded in the analyses of heart failure. For sudden cardiac arrest, no prevalent cases were excluded.

Comparison between CIS-based and HDR-based definitions

In the analyses, the CIS-based registry was used as gold standard. A positive match between the registries was defined as a registration with the specific disease in both the CIS-based and HDR-based registry within a time frame of 6 months prior to or post diagnosis in CIS (true positives). Sensitivity was calculated as the number of cases with a positive match divided by the total number of cases in the CIS-based registry. Positive predictive value was calculated as the number of cases with a positive match divided by the total number of cases in the HDR-based registry. The 95% confidence intervals were calculated using the standard error of the estimate of the binomial distribution in the usual manner. Stratified analyses were performed for sex, age at diagnosis, and study period.

Results

Incidence estimates

During follow-up, 815 cases were registered with CHD in the CIS-based registry, 481 cases with acute myocardial infarction, and 390 cases with unstable angina pectoris (Table 1). The incidence rates per 100,000 person-years were 362.2 for CHD, 212.2 for acute myocardial infarction, and 171.8 for unstable angina pectoris (Table 2). In addition, 154 cases were registered with heart failure of whom 68 cases (44.2%) self-reported CHD at baseline or had been diagnosed with CHD prior to the diagnosis of heart failure in CIS (Table 1). Among the 152 cases with sudden cardiac arrest in the CIS-based registry, 57 cases (37.5%) self-reported CHD at baseline or had been diagnosed with CHD or heart failure prior to the diagnosis of sudden cardiac arrest in CIS. The incidence rates per 100,000 person-years were 66.4 for heart failure and 65.4 for sudden cardiac arrest (Table 2).

In the HDR-based registry, 656 cases were registered with CHD during follow-up, 417 cases with acute myocardial infarction, 269 cases with unstable angina pectoris, and 84 cases with heart failure. There were no cases with

sudden cardiac arrest as discharge diagnosis in the HDR. The incidence rates of these diseases derived from the HDR-based registry were lower than those derived from the CIS-based registry (Table 2). Especially in the older age categories (50–59 and 60–69 years) of both men and women, the incidence rates from the HDR-based registry were lower compared with those from the CIS-based registry, except for female cases with acute myocardial infarction. For both men and women, the estimated incidence rates per age category from the CIS-based and HDR-based registry are given in the [Appendix](#).

Validity of the HDR-based registry

For the HDR-based definition of CHD, the sensitivity and positive predictive value were 72 and 91%, respectively (Table 3). A positive match was found in 590 (70.8%) of the 833 cases with CHD in one or both of the registries. For 43 persons (5.2%), the time difference between the diagnosis in the HDR-based and CIS-based registry was longer than 6 months. Furthermore, 182 CHD cases (21.8%) were found only in the CIS-based registry, while 18 cases (2.2%) were found only in the HDR-based registry. Because normally CHD refers to the ICD-9 codes 410–414, the HDR-definition of CHD was extended to these ICD-9 codes in additional analyses. In doing so, the sensitivity increased from 72 to 85% while the positive predictive value decreased from 91 to 85%.

For the HDR-based definition of acute myocardial infarction, the sensitivity and positive predictive value were 84 and 97%, respectively (Table 3). For 404 cases (83.3%), a positive match was found between the HDR-based and the CIS-based registry. For eight cases (1.6%), the time difference between the diagnosis in the HDR-based and CIS-based registry was longer than 6 months. The remaining cases (15.1%) were found in either one of the registries.

The sensitivity and positive predictive value for the HDR-based definition of unstable angina pectoris (53 and 78%, respectively) were substantially lower (Table 3). Only 208 out of the 420 cases with unstable angina pectoris (49.5%) were found in both registries, while 185 cases (44.0%) were registered only in either one of the registries. For 27 cases (6.4%), the time between the diagnoses was more than 6 months.

The most important reasons to be only registered in the CIS-based registry for cases with CHD, acute myocardial infarction, and unstable angina pectoris were as follows: a different discharge diagnosis in the HDR-based registry (68, 59, and 75%, respectively), mostly ICD-9 codes 413.90 (other and unspecified angina pectoris) and 414.00 (coronary atherosclerosis), a diagnosis based on outpatient files (18, 6, and 21%, respectively), or a hospital admission in another Dutch hospital (1, 3, and 0%, respectively), or in a foreign hospital (4, 9, and 1%, respectively). All cases

Table 1 Baseline characteristics of the CAREMA cohort in The Netherlands, 1987–2003

	Total (<i>n</i> = 21,110) <i>N</i> (%)	Men (<i>n</i> = 9,935) <i>N</i> (%)	Women (<i>n</i> = 11,175) <i>N</i> (%)
Year of inclusion in the study			
1987–1991	12,485 (59.1)	5,953 (59.9)	6,532 (58.5)
1992	2,189 (10.4)	1,010 (10.2)	1,179 (10.6)
1993–1997	6,436 (30.5)	2,972 (29.9)	3,464 (31.0)
Age in years at baseline			
20–29	3,828 (18.1)	1,724 (17.4)	2,104 (18.8)
30–39	5,287 (25.0)	2,486 (25.0)	2,801 (25.1)
40–49	6,089 (28.8)	2,900 (29.2)	3,189 (28.5)
50–60	5,906 (28.0)	2,825 (28.4)	3,081 (27.6)
Mean age in years at baseline (range)	41.7 (20.1–60.9)	41.9 (20.1–60.2)	41.5 (20.1–60.9)
During follow-up (1987–2003)			
Migration out of study area	2,106 (10.0)	935 (9.4)	1,171 (10.5)
Emigration to foreign country	621 (2.9)	313 (3.2)	308 (2.8)
Deceased	791 (3.7)	500 (5.0)	291 (2.6)
Coronary heart disease ^a			
Number of incident cases ^b	815 (3.9)	595 (6.2)	220 (2.0)
Mean age at diagnosis (sd)	57.4 (7.6)	57.0 (7.7)	58.6 (7.4)
Acute myocardial infarction			
Number of incident cases ^b	481 (2.3)	372 (3.8)	109 (1.0)
Mean age at diagnosis (sd)	56.7 (7.7)	56.4 (7.7)	57.7 (7.5)
Unstable angina pectoris			
Number of incident cases ^b	390 (1.9)	269 (2.8)	121 (1.1)
Mean age at diagnosis (sd)	58.3 (7.4)	58.1 (7.3)	58.5 (7.6)
Heart failure			
<i>With history of coronary heart disease</i>			
Number of incident cases ^c	68 (0.3)	47 (0.5)	21 (0.2)
Mean age at diagnosis (sd)	62.4 (6.4)	62.0 (6.7)	63.5 (5.9)
<i>Without history of coronary heart disease</i>			
Number of incident cases ^c	86 (0.4)	50 (0.5)	36 (0.3)
Mean age at diagnosis (sd)	61.8 (7.7)	61.7 (8.1)	61.9 (7.3)
Sudden cardiac arrest			
<i>With history of cardiac disease</i>			
Number of incident cases	57 (0.3)	47 (0.5)	10 (0.1)
Mean age at diagnosis (sd)	60.6 (5.8)	60.9 (5.7)	59.0 (5.8)
<i>Without history of cardiac disease</i>			
Number of incident cases	95 (0.5)	67 (0.7)	28 (0.3)
Mean age at diagnosis (sd)	56.7 (8.8)	56.7 (8.5)	56.4 (9.8)

^a Defined as incident acute myocardial infarction, unstable angina pectoris, coronary artery bypass grafting (CABG), percutaneous transluminal coronary angioplasty (PTCA), or coronary heart disease as cause of death (ICD-9 codes 410–414; ICD-10 codes I20–I25)

^b Exclusion of persons with coronary heart disease at baseline (*n* = 347), based on self-report or a diagnosis date before baseline in the CIS-based registry

^c Exclusion of persons with heart failure at baseline (*n* = 7), based on a diagnosis date before baseline in the CIS-based registry

with CHD, acute myocardial infarction and unstable angina pectoris that were registered only in the HDR-based registry had a diagnosis of another (11, 100, and 40%, respectively) or no cardiovascular disease (89, 0, and 60%, respectively) according to the CIS-based registry.

For the HDR-based definition of heart failure, the sensitivity and positive predictive value were 43 and 80%, respectively (Table 3). A positive match was found in 66 cases (41.8%), while 79 cases (50.0%) were registered only in either one of the registries. For 13 cases (8.2%), the time

between the diagnoses was more than 6 months. The most important reasons for cases with heart failure to be registered only in the CIS-based registry were a diagnosis based on outpatient files (36%), a different discharge diagnosis in the HDR-based registry (36%), or a diagnosis of heart failure made during hospital admission for another cardiovascular disease, but not registered as discharge diagnosis (21%).

In the stratified analyses, slightly higher sensitivities were found in women compared with men, except for heart failure (data not shown). By contrast, the positive

Table 2 Estimated incidence rates from the HDR-based and CIS-based registry in the CAREMA cohort in 1987–2003

	HDR-based registry			CIS-based registry		
	Number of registered cases	Person years at risk ($\times 1,000$)	Incidence per 100,000 person-years at risk	Number of registered cases	Person years at risk ($\times 1,000$)	Incidence per 100,000 person-years at risk
Coronary heart disease ^a	656 ^b	226.2	290.0	815 ^b	225.0	362.2
Acute myocardial infarction	417 ^b	227.3	183.4	481 ^b	226.7	212.2
Unstable angina pectoris	269 ^b	227.8	118.1	390 ^b	227.0	171.8
Heart failure	84 ^c	232.3	36.2	154 ^c	232.0	66.4
Sudden cardiac arrest	–	–	– ^d	152	232.5	65.4

HDR hospital discharge register, CIS cardiology information system

^a Defined as incident acute myocardial infarction, unstable angina pectoris, coronary artery bypass grafting (CABG), percutaneous transluminal coronary angioplasty (PTCA), or coronary heart disease as cause of death (ICD-9 codes 410–414; ICD-10 codes I20–I25)

^b Exclusion of persons with coronary heart disease at baseline based on self-report or a diagnosis date before baseline in the registry ($n = 330$ in the HDR-based registry and $n = 347$ in the CIS-based registry)

^c Exclusion of persons with heart failure at baseline based on a diagnosis date before baseline in the registry ($n = 1$ in the HDR-based registry and $n = 7$ in the CIS-based registry)

^d It was not possible to estimate an incidence rate because there were no cases with sudden cardiac arrest as discharge diagnosis in the HDR

predictive value was slightly higher in men than in women. Furthermore, higher sensitivities were found in the age category <50 years compared with the age category ≥ 50 years, except for heart failure. For CHD and acute myocardial infarction, the positive predictive values were also higher in the age category <50 years, while these values were higher in the age category ≥ 50 years for unstable angina pectoris and heart failure. Both the sensitivities and positive predictive values were higher in the study period 1996–2003 than in the study period 1987–1995, except for the sensitivity of heart failure.

Discussion

In the comparison between the HDR-based and CIS-based registry, relatively high sensitivities and positive predictive values were found for CHD and acute myocardial infarction, while these values were considerably lower for unstable angina pectoris and heart failure. Furthermore, high percentages of the cases were only found in the CIS-based registry, varying from 14.2% for acute myocardial infarction to 47.5% for heart failure. These cases were missed or miscoded in the HDR-based registry. As a consequence, the incidence rates in the HDR-based registry were considerably lower than the incidence rates in the CIS-based registry, especially for unstable angina pectoris and heart failure.

Several reasons may have contributed to the differences found between the HDR-based and CIS-based registry. The diagnoses from CIS were abstracted and coded by trained registrars under guidance of a cardiologist (AG). Therefore, the diagnoses in the CIS-based registry are probably less susceptible to misclassification.

Furthermore, the CIS also contains information about visits to the outpatient clinic and emergency ward for heart problems. Cases diagnosed at these departments with cardiac diseases that do not warrant hospitalisation were still registered in the CIS-based registry. These cases were missed when only data from the HDR was used, leading to an underestimation of the incidence rates, especially for diagnoses that do not warrant hospitalisation, as can be seen from this study.

For the incidence estimates in this study, data were used from the University Hospital Maastricht (UHM). Because of the central and unique position of this hospital in the study region, the cardiologic follow-up is expected to be almost complete. Only a few cases will be missed, partly due to a diagnosis in another Dutch or in a foreign hospital. However, part of these cases may have visited the outpatient clinic of the UHM department of cardiology in a later stage, so that they were still registered in the CIS-based registry but not in the HDR-based registry. Nonetheless, cases diagnosed in another Dutch hospital would probably be found when data is used from the national HDR, while cases diagnosed in a foreign country would still be missed. In the Netherlands, however, record linkage to the national HDR is difficult, because of the limited number of identifying variables in this register.

Because the definition of CHD in the CIS-based registry was restricted to a clinical diagnosis of acute myocardial infarction, unstable angina pectoris, CABG or PTCA, we also narrowed the definition of CHD in the HDR-based registry to cases with ICD-9 codes 410, 411.1, and 413.1 as discharge diagnosis. When the definition in the HDR-based registry was extended to ICD-9 codes 410–414, the sensitivity increased from 72 to 85%, which can be explained by

Table 3 Sensitivity (Se) and positive predictive values (PPV) for the cardiovascular diseases in the HDR-based registry compared with the CIS-based registry in 1987–2003 within a time frame of 6 months prior to or post diagnosis in CIS

	Number of incident cases ^a	Se (95% CI)	PPV (95% CI)
<i>Coronary heart disease^b</i>			
Total	833	72 (69–75)	91 (88–93)
Age group			
<50 years	151	78 (71–85)	91 (86–96)
≥50 years	682	71 (68–75)	90 (88–93)
Study period			
1987–1995	275	71 (65–76)	82 (77–87)
1996–2003	558	73 (69–77)	95 (93–97)
<i>Acute myocardial infarction</i>			
Total	485	84 (81–87)	97 (96–99)
Age group			
<50 years	103	86 (80–93)	99 (97–100)
≥50 years	382	83 (80–87)	97 (95–99)
Study period			
1987–1995	168	82 (76–87)	94 (91–98)
1996–2003	317	85 (81–89)	99 (97–100)
<i>Unstable angina pectoris</i>			
Total	420	53 (48–58)	78 (74–83)
Age group			
<50 years	61	54 (41–67)	76 (62–89)
≥50 years	359	53 (48–59)	79 (74–84)
Study period			
1987–1995	119	53 (44–63)	69 (59–79)
1996–2003	301	53 (48–59)	83 (77–88)
<i>Heart failure</i>			
Total	158	43 (35–51)	80 (71–88)
Age group			
<50 years	11	30 (2–58)	60 (17–100)
≥50 years	147	44 (36–52)	81 (72–90)
Study period			
1987–1995	35	45 (28–62)	68 (49–88)
1996–2003	123	42 (33–51)	84 (74–93)

Prevalent cases based on self-report or a diagnosis date in one or both registries before baseline were excluded ($n = 349$ for coronary heart disease, acute myocardial infarction and unstable angina pectoris and $n = 7$ for heart failure)

^a Total number of incident cases with a diagnosis in either one or both the HDR-based or the CIS-based registry

^b Defined as incident acute myocardial infarction, unstable angina pectoris, coronary artery bypass grafting (CABG), percutaneous transluminal coronary angioplasty (PTCA), or coronary heart disease as cause of death (ICD-9 codes 410–414; ICD-10 codes I20–I25)

the larger range of ICD-9 codes used for specific CHD in the HDR. The positive predictive value, however, decreased from 91 to 85%, which can be explained by the inclusion of all ischemic heart diseases in the HDR-based registry, including stable angina pectoris and chronic CHD that were not included in the CIS-based registry.

The identification of incident cases in this study was partly performed by record linkage of the CAREMA cohort to the causes of death registry from Statistics Netherlands. Although we did not validate these cases, several studies in the Netherlands showed that the registration and coding of causes of death by Statistics Netherlands had a higher validity compared with other European countries [22, 23]. Because the cause of death registry was used in both the CIS-based and HDR-based registry for the identification of incident cases, this has led to an improvement of the

comparison between these registries. Additional analyses, in which incident cases identified by linkage to Statistics Netherlands were excluded, showed small decreases in positive predictive values but considerable decreases in the sensitivities. This was due to a relatively large increase in the number of cases registered only in the CIS-based registry after exclusion. This means that the record linkage with the causes of death registry was especially favourable for the completeness of the HDR-based registry.

In this study, clinical diagnoses in CIS were used for the identification of cases in stead of diagnostic criteria. However, 321 of the 417 cases (77.0%) with a clinical diagnosis of acute myocardial infarction in the CIS-based registry met the diagnostic criteria of the European Society of Cardiology and the American College of Cardiology [24]. The remaining 23.0% had incomplete data. Of the

321 cases that met the diagnostic criteria in the CIS-based registry, 291 cases (90.7%) were also registered with acute myocardial infarction in the HDR. Thus, even for a diagnosis of an acute myocardial infarction based on diagnostic criteria, a considerable part of the cases was not registered with this diagnosis in the HDR. During follow-up, more sensitive screening tests became available for the diagnosis of an acute myocardial infarction. Because of these tests, clinical decision-making may have changed during the follow-up period.

The estimates of the incidence rates from the HDR-based registry are comparable to those reported by Koek et al. [7] for the Netherlands as a whole. For comparison purposes, we calculated an expected incidence rate of acute myocardial infarction using the national incidence rates of Koek et al., age and gender-standardized to the CAREMA cohort. This expected incidence rate was higher than the incidence rate in the HDR-based registry (201.9 and 183.4 per 100,000 person-years, respectively), which may be explained by a lower incidence rate in the study population, [25] which is restricted to the Maastricht region, compared with the average Dutch population. Furthermore, regional differences in the coverage and validity of local hospital discharge registries may also explain this discrepancy. Conversely, the expected rate was lower than the incidence rate in the CIS-based registry (212.2 per 100,000 person-years).

Several studies have investigated the validity of hospital discharge and/or mortality data on acute myocardial infarction by comparing these data with specific study registers [8–15] or physician reviews [5, 26–28]. In these studies, a wide range of estimated values for the sensitivity and positive predictive value was found due to differences in case identification. However, most of the studies demonstrated that hospital discharge and/or mortality data underestimate the incidence of acute myocardial infarction in the population as was found in our study.

Furthermore, three Finnish validation studies found higher sensitivities and positive predictive values in men compared with women [9, 12, 15]. In our study, the positive predictive value was also higher in men, while the sensitivity was higher in women. In the stratified analyses, we also found higher sensitivities and positive predictive values in the study period 1996–2003 compared with the period 1987–1995. This implies an improvement of the validity of the HDR in time.

Only two studies investigated the validity of hospital discharge data on heart failure using the definition of heart failure by the European Society of Cardiology [29]. Ingelsson et al. [16] found a positive predictive value of 82% which is slightly higher than the value of 80% found in our study. A considerably lower value of 65% was found in the study by Khand et al. [17]. However, Khand et al. used a broader range of ICD-codes (including ICD-10 codes I25.5

and I42.9) which are probably less sensitive for a definite diagnosis of heart failure.

In many of the above mentioned studies, data from the WHO MONICA project were used [8, 9, 11–15]. This project is a multicenter study which monitors the incidence of myocardial infarction (MI) in several countries using study-specific MI registers. All events that occurred in the study population were registered according to previous defined diagnostic criteria. Although the registration in the CIS-based registry was based on a clinical diagnosis made by experienced cardiologists, a large number of cases with an acute myocardial infarction (77%) met the diagnostic criteria of the European Society of Cardiology and the American College of Cardiology as described earlier in this discussion. In addition, none of the registered cases with complete data did not fulfil these diagnostic criteria. In the CIS-based registry, however, not only diagnoses of an acute myocardial infarction were registered but also diagnoses of silent myocardial infarctions, unstable and stable angina pectoris, and heart failure. However, the registrations in the CIS-based registry were only made for people living in the Maastricht region. In the MONICA project, the centers did also not have national coverage [8]. Therefore, the estimated incidence rates of both the CIS-based registry and the MI registers of the MONICA project may not be generalised to a national level.

Because cardiovascular disease registries are lacking in most countries, record linkage with hospital discharge and mortality data is often used to estimate the incidence rates of CHD and other cardiologic diseases. However, this study and previous studies have shown that a considerable part of the cases is missed or miscoded using hospital discharge data. Therefore, incidence rates based on these data may underestimate the true incidence rates, especially for unstable angina pectoris and heart failure.

Furthermore, an accurate identification of cases is even more important in etiological studies in which risk estimates are based on the comparison between cases and non-cases. Case identification based on hospital discharge and mortality data may lead to biases in the results of these studies. Therefore, these data should be used with caution in epidemiological studies, especially in etiological studies. Although the CIS-based registry has several advantages over HDRs, some events may be still be missed. In etiological studies, it is important to keep in mind the potential weaknesses of such registries.

Acknowledgments The Monitoring Project on Cardiovascular Disease Risk Factors (PPHVZ) and the MORGEN study were financially supported by the Ministry of Public Health, Welfare and Sports of the Netherlands. The authors wish to thank D Kromhout for supervision of the PPHVZ project and the MORGEN-project steering committee consisting of (not mentioning co-authors of this article) H. B. Bueno de Mesquita, H. A. Smit, and J. C. Seidell (project leader).

Furthermore the authors thank epidemiologists and field-workers of the Municipal Health Services in Maastricht, Amsterdam, and Doetinchem for their contribution to baseline data collection and those involved in the logistic management (A. Jansen and J. Steenbrink) and the data management (A. Blokstra, A. van Kessel, P. Steinberger, E. den Hoedt, I. Toxopeus, J. van der Laan). The authors further wish to thank D. Jaspers, A. Hilton, V. Visser, P. Erkens, S. Philippens, J. Bremen, B. Bleijlevens, T. van Moergastel, S. van de Crommert for assistance in clinical data collection. Statistics Netherlands is acknowledged for providing data on causes of death. *Sources of financial support:* This study was funded by School for Public Health and Primary Care (Caphri).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Appendix

See Table 4.

Table 4 Age- and gender-specific incidence rates from the HDR-based and CIS-based registry in the CAREMA cohort, 1987–2003

	HDR-based registry				CIS-based registry			
	Men		Women		Men		Women	
	Number of incident cases	Incidence/100,000 person-years at risk	Number of incident cases	Incidence/100,000 person-years at risk	Number of incident cases	Incidence/100,000 person-years at risk	Number of incident cases	Incidence/100,000 person-years at risk
Coronary heart disease^a								
Total	471	449.4	185	152.4	595 ^b	572.9	220 ^b	181.6
20–29 years	0	0	0	0	0	0	0	0
30–39 years	11	53.3	4	17.3	11	53.3	4	17.3
40–49 years	84	280.3	24	71.5	106	354.8	28	83.4
50–59 years	193	633.1	73	206.2	249	825.4	87	246.2
60–69 years	169	1,087.1	71	363.1	213	1,413.4	89	458.4
Acute myocardial infarction^a								
Total	319	302.2	98	80.5	372	354.4	109	89.6
20–29 years	0	0	0	0	0	0	0	0
30–39 years	10	48.4	2	8.6	10	48.4	2	8.6
40–49 years	62	206.4	15	44.7	75	250.0	16	47.6
50–59 years	136	442.4	42	118.2	154	504.7	49	138.1
60–69 years	102	641.3	34	172.0	124	793.1	37	187.7
Unstable angina pectoris^a								
Total	175	165.0	94	77.2	269	255.2	121	99.5
20–29 years	0	0	0	0	0	0	0	0
30–39 years	2	9.7	2	8.6	2	9.7	2	8.6
40–49 years	24	79.6	11	32.8	37	122.9	16	47.7
50–59 years	67	216.4	36	101.4	112	364.1	43	121.3
60–69 years	77	480.0	37	187.5	110	700.8	52	264.6
Heart failure (total)^c								
Total	52	47.5	32	26.0	97	88.8	57	46.4
20–29 years	0	0	0	0	0	0	0	0
30–39 years	1	4.8	0	0	0	0	0	0
40–49 years	4	13.1	0	0	7	23.0	3	8.9
50–59 years	14	43.2	9	25.0	25	77.4	16	44.6
60–69 years	29	165.4	19	94.0	52	298.2	28	138.9
Heart failure (with history of CHD)^c								
Total	15	13.7	12	9.8	47	43.0	21	17.1
20–29 years	0	0	0	0	0	0	0	0
30–39 years	0	0	0	0	0	0	0	0

Table 4 continued

	HDR-based registry				CIS-based registry			
	Men		Women		Men		Women	
	Number of incident cases	Incidence/100,000 person-years at risk	Number of incident cases	Incidence/100,000 person-years at risk	Number of incident cases	Incidence/100,000 person-years at risk	Number of incident cases	Incidence/100,000 person-years at risk
40–49 years	1	3.3	0	0	2	6.6	0	0
50–59 years	5	15.4	3	8.3	13	40.2	5	13.9
60–69 years	8	45.6	8	39.6	25	143.4	12	59.5
Heart failure (without history of CHD) ^c								
Total	37	33.8	20	16.3	50	45.8	36	29.3
20–29 years	0	0	0	0	0	0	0	0
30–39 years	1	4.8	0	0	0	0	0	0
40–49 years	3	9.8	0	0	5	16.4	3	8.9
50–59 years	9	27.8	6	16.7	12	37.1	11	30.6
60–69 years	21	119.7	11	54.4	27	154.8	16	79.4
Sudden cardiac arrest (total)								
Total	–	–	–	–	114	104.1	38	30.9
20–29 years	–	–	–	–	0	0	0	0
30–39 years	–	–	–	–	3	14.5	2	8.6
40–49 years	–	–	–	–	17	55.8	6	17.8
50–59 years	–	–	–	–	36	111.1	16	44.5
60–69 years	–	–	–	–	55	312.6	10	49.3
Sudden cardiac arrest (with history of CVD) ^d								
Total	–	–	–	–	47	42.9	10	8.1
20–29 years	–	–	–	–	0	0	0	0
30–39 years	–	–	–	–	0	0	0	0
40–49 years	–	–	–	–	3	9.8	0	0
50–59 years	–	–	–	–	15	46.3	5	13.9
60–69 years	–	–	–	–	28	159.1	5	24.7
Sudden cardiac arrest (without history of CVD) ^d								
Total	–	–	–	–	67	61.2	28	22.8
20–29 years	–	–	–	–	0	0	0	0
30–39 years	–	–	–	–	3	14.5	2	8.6
40–49 years	–	–	–	–	14	46.0	6	17.8
50–59 years	–	–	–	–	21	64.8	11	30.6
60–69 years	–	–	–	–	27	153.5	5	24.7

HDR Hospital Discharge Register, CIS Cardiology Information System, “–” means not available. The table is censored at 70 years of age for all cardiovascular diseases because of the small number of cases in the age category ≥ 70 years

^a Exclusion of persons with coronary heart disease at baseline based on self-report or a diagnosis date before baseline in the registry ($n = 330$ in the HDR-based registry and $n = 347$ in the CIS-based registry)

^b Defined as incident acute myocardial infarction, unstable angina pectoris, coronary artery bypass grafting (CABG), percutaneous transluminal coronary angioplasty (PTCA), or coronary heart disease as cause of death (ICD-9 codes 410–414; ICD-10 codes I20–I25)

^c Exclusion of persons with heart failure at baseline based on a diagnosis date before baseline in the registry ($n = 1$ in the HDR-based registry and $n = 7$ in the CIS-based registry)

^d Refers to a history of coronary heart disease or heart failure prior to the diagnosis of sudden cardiac arrest in the registry

References

- Kesteloot H, Sans S, Kromhout D. Dynamics of cardiovascular and all-cause mortality in Western and Eastern Europe between 1970 and 2000. *Eur Heart J*. 2006;27(1):107–13. doi:10.1093/eurheartj/ehi511.
- Menotti A, Lanti M, Kromhout D, Blackburn H, Nissinen A, Dontas A, et al. Forty-year coronary mortality trends and changes in major risk factors in the first 10 years of follow-up in the seven countries study. *Eur J Epidemiol*. 2007;22(11):747–54. doi:10.1007/s10654-007-9176-4.
- McGovern PG, Jacobs DR Jr, Shahar E, Arnett DK, Folsom AR, Blackburn H, et al. Trends in acute coronary heart disease mortality, morbidity, and medical care from 1985 through 1997: the Minnesota heart survey. *Circulation*. 2001;104(1):19–24.
- Reitsma JB, Dalstra JA, Bonsel GJ, van der Meulen JH, Koster RW, Gunning-Schepers LJ, et al. Cardiovascular disease in the Netherlands, 1975 to 1995: decline in mortality, but increasing numbers of patients with chronic conditions. *Heart*. 1999;82(1):52–6.
- Kennedy GT, Stern MP, Crawford MH. Miscoding of hospital discharges as acute myocardial infarction: implications for surveillance programs aimed at elucidating trends in coronary artery disease. *Am J Cardiol*. 1984;53(8):1000–2. doi:10.1016/0002-9149(84)90625-8.
- Mahonen M, Salomaa V, Keskimaki I, Moltchanov V. The feasibility of routine mortality and morbidity register data linkage to study the occurrence of acute coronary heart disease events in Finland. The Finnish cardiovascular diseases registers (CVDR) project. *Eur J Epidemiol*. 2000;16(8):701–11. doi:10.1023/A:1026599805969.
- Koek HL, de Bruin A, Gast A, Gevers E, Kardaun JW, Reitsma JB, et al. Incidence of first acute myocardial infarction in the Netherlands. *Neth J Med*. 2007;65(11):434–41.
- 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):124–30. doi:10.1016/S0895-4356(02)00591-7.
- Pajunen P, Koukkunen H, Ketonen M, Jerkkola T, Immonen-Raiha P, Karja-Koskenkari P, et al. The validity of the Finnish hospital discharge register and causes of death register data on coronary heart disease. *Eur J Cardiovasc Prev Rehabil*. 2005;12(2):132–7. doi:10.1097/00149831-200504000-00007.
- Hammar N, Nerbrand C, Ahlmark G, Tibblin G, Tsipogianni A, Johansson S, et al. Identification of cases of myocardial infarction: hospital discharge data and mortality data compared to myocardial infarction community registers. *Int J Epidemiol*. 1991;20(1):114–20. doi:10.1093/ije/20.1.114.
- Boyle CA, Dobson AJ. The accuracy of hospital records and death certificates for acute myocardial infarction. *Aust N Z J Med*. 1995;25(4):316–23.
- Mahonen M, Salomaa V, Brommels M, Molarius A, Miettinen H, Pyorala K, et al. The validity of hospital discharge register data on coronary heart disease in Finland. *Eur J Epidemiol*. 1997;13(4):403–15. doi:10.1023/A:1007306110822.
- Rapola JM, Virtamo J, Korhonen P, Haapakoski J, Hartman AM, Edwards BK, et al. Validity of diagnoses of major coronary events in national registers of hospital diagnoses and deaths in Finland. *Eur J Epidemiol*. 1997;13(2):133–8. doi:10.1023/A:1007380408729.
- De Henauw S, de Smet P, Aelvoet W, Kornitzer M, De Backer G. Misclassification of coronary heart disease in mortality statistics. Evidence from the WHO-MONICA Ghent-Charleroi study in Belgium. *J Epidemiol Community Health*. 1998;52(8):513–9. doi:10.1136/jech.52.8.513.
- Mahonen M, Salomaa V, Torppa J, Miettinen H, Pyorala K, Immonen-Raiha P, et al. The validity of the routine mortality statistics on coronary heart disease in Finland: comparison with the FINMONICA MI register data for the years 1983–1992. Finnish multinational MONItoring of trends and determinants in CArdiovascular disease. *J Clin Epidemiol*. 1999;52(2):157–66. doi:10.1016/S0895-4356(98)00145-0.
- Ingelsson E, Arnlov J, Sundstrom J, Lind L. The validity of a diagnosis of heart failure in a hospital discharge register. *Eur J Heart Fail*. 2005;7(5):787–91. doi:10.1016/j.ejheart.2004.12.007.
- Khand AU, Shaw M, Gemmel I, Cleland JG. Do discharge codes underestimate hospitalisation due to heart failure? Validation study of hospital discharge coding for heart failure. *Eur J Heart Fail*. 2005;7(5):792–7. doi:10.1016/j.ejheart.2005.04.001.
- Van Leer EM, Seidell JC, Kromhout D. Dietary calcium, potassium, magnesium and blood pressure in the Netherlands. *Int J Epidemiol*. 1995;24(6):1117–23. doi:10.1093/ije/24.6.1117.
- Lean ME, Han TS, Seidell JC. Impairment of health and quality of life in people with large waist circumference. *Lancet*. 1998;351(9106):853–6. doi:10.1016/S0140-6736(97)10004-6.
- Van den Brandt PA, Schouten LJ, Goldbohm RA, Dorant E, Hunen PM. Development of a record linkage protocol for use in the Dutch cancer registry for epidemiological research. *Int J Epidemiol*. 1990;19(3):553–8. doi:10.1093/ije/19.3.553.
- Gorgels AP, Gijsbers C, de Vreede-Swagemakers J, Lousberg A, Wellens HJ. Out-of-hospital cardiac arrest—the relevance of heart failure. The Maastricht circulatory arrest registry. *Eur Heart J*. 2003;24(13):1204–9. doi:10.1016/S0195-668X(03)00191-X.
- Mackenbach JP, van Duyn WM. Registration and coding of various causes of death in the Netherlands and other EEC countries. *Ned Tijdschr Geneesk*. 1984;128(1):13–8.
- Mackenbach JP, Van Duyn WM, Kelson MC. Certification and coding of two underlying causes of death in The Netherlands and other countries of the European community. *J Epidemiol Community Health*. 1987;41(2):156–60. doi:10.1136/jch.41.2.156.
- Alpert JS, Thygesen K, Antman E, Bassand JP. Myocardial infarction redefined—a consensus document of The Joint European Society of Cardiology/American College of Cardiology Committee for the redefinition of myocardial infarction. *J Am Coll Cardiol*. 2000;36(3):959–69. doi:10.1016/S0735-1097(00)00804-4.
- Roedig A. Acuut hartinfarct 2001–2004: ziekenhuisopnamen, per GGD-regio, gecorrigeerd voor leeftijd en geslacht. The National Public Health Compass 2006; Available from: http://www.rivm.nl/vtv/object_map/o1626n25675.html. Cited 5 June 2008.
- Lindblad U, Rastam L, Ranstam J, Peterson M. Validity of register data on acute myocardial infarction and acute stroke: the Skaraborg hypertension project. *Scand J Soc Med*. 1993;21(1):3–9.
- Heliovaara M, Reunanen A, Aromaa A, Knekt P, Aho K, Suonen O. Validity of hospital discharge data in a prospective epidemiological study on stroke and myocardial infarction. *Acta Med Scand*. 1984;216(3):309–15.
- Pladevall M, Goff DC, Nichaman MZ, Chan F, Ramsey D, Ortiz C, et al. An assessment of the validity of ICD Code 410 to identify hospital admissions for myocardial infarction: the Corpus Christi heart project. *Int J Epidemiol*. 1996;25(5):948–52. doi:10.1093/ije/25.5.948.
- Guidelines for the diagnosis of heart failure. The task force on heart failure of the European society of cardiology. *Eur Heart J*. 1995;16(6):741–51.