

SCIENTIFIC REPORTS



OPEN

Comparison of self-reported and register-based hospital medical data on comorbidities in women

Peh Joo Ho^{1,2}, Chuen Seng Tan², Shajedur Rahman Shawon³, Mikael Eriksson⁴, Li Yan Lim⁵, Hui Miao¹, Eileen Png¹, Kee Seng Chia², Mikael Hartman^{6,2}, Jonas F. Ludvigsson^{4,7}, Kamila Czene⁴, Per Hall^{4,8} & Jingmei Li^{1,4,5}

Breast cancer patients commonly present with comorbidities which are known to influence treatment decisions and survival. We aim to examine agreement between self-reported and register-based medical records (National Patient Register [NPR]). Ascertainment of nine conditions, using individually-linked data from 64,961 women enrolled in the Swedish KARolinska MAMmography Project for Risk Prediction of Breast Cancer (KARMA) study. Agreement was assessed using observed proportion of agreement (overall agreement), expected proportion of agreement, and Cohen's Kappa statistic. Two-stage logistic regression models taking into account chance agreement were used to identify potential predictors of overall agreement. High levels of overall agreement (i.e. $\geq 86.6\%$) were observed for all conditions. Substantial agreement (Cohen's Kappa) was observed for myocardial infarction (0.74), diabetes (0.71) and stroke (0.64) between self-reported and NPR data. Moderate agreement was observed for preeclampsia (0.51) and hypertension (0.46). Fair agreement was observed for heart failure (0.40) and polycystic ovaries or ovarian cysts (0.27). For hyperlipidemia (0.14) and angina (0.10), slight agreement was observed. In most subgroups we observed negative specific agreement of $>90\%$. There is no clear reference data source for ascertainment of conditions. Negative specific agreement between NPR and self-reported data is consistently high across all conditions.

Cancers including breast cancer commonly present with one or more additional medical conditions, hereafter referred to as comorbidities. Comorbidities are known to influence treatment decisions and ultimately survival of breast or other cancers^{1,2}. In breast cancer patients, presence of comorbidities was found to increase the likelihood of being diagnosed with advanced disease, for example, with distant metastasis³. These patients were more likely to be treated with less-than-standard therapy than patients with no comorbidities^{4,5}. In addition, delay and non-completion of adjuvant therapies were more frequent in this group of patients with comorbidities⁶. The risk of dying from breast cancer was also higher among breast cancer cases with a history of diabetes or myocardial infarction¹.

Information on comorbidities can be ascertained through multiple data sources including patient self-reports, medical record abstraction and disease registries⁷. National registers such as the National Patient Register (NPR) in Sweden are goldmines for epidemiological research due to their rich and long-term data on various health conditions and procedures⁸. The NPR was established in 1964, and achieved nationwide coverage for all in-patient visits in 1987⁹. It contains detailed information about the patient, geographical data, administrative data for both inpatient and outpatient visits and codes for related medical diagnosis and procedures. The register, however, does not yet contain primary care data. A validation study by Ludvigsson *et al.* using only inpatient records showed high positive predictive value (i.e. the probability of truly having the condition given that the inpatient records

¹Genome Institute of Singapore, 60 Biopolis Street, Genome, #02-01, Singapore, 138672, Singapore. ²Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore, Singapore. ³Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Old road campus, OX3 7LF, Oxford, UK. ⁴Karolinska Institutet, Department of Medical Epidemiology and Biostatistics, Box 281, 171 77, Stockholm, Sweden. ⁵Department of Surgery, University Surgical Cluster, National University Hospital, Singapore, Singapore. ⁶Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore. ⁷Department of Pediatrics, Örebro University Hospital, Örebro University, Örebro, Sweden. ⁸Department of Oncology, Södersjukhuset, 118 84, Stockholm, Sweden. Correspondence and requests for materials should be addressed to J.L. (email: lijm1@gis.a-star.edu.sg)

reports it) (85–95% in general) but lower sensitivity (i.e. the probability of inpatient records reporting the condition given that the condition is present) for many diagnoses in the NPR⁹.

Comorbid conditions such as hypertension and adulthood onset diabetes are almost exclusively managed in primary care and are therefore not well-captured in the register-based hospital records¹⁰. Ludvigsson *et al.* noted that sensitivity captured by inpatient records were especially low for hypertension and lipid disorders in the Swedish NPR (~10%)⁹. Other studies have highlighted the under-recording of mild medical conditions that do not require hospital care in hospital admission data^{11,12}. Therefore, self-reported data from the patients may bridge the gap in recording these conditions and be an important source of information¹³. Though the reliability of self-reported medical conditions varies considerably, it has been increasingly adopted in both research and clinical settings¹³. The accuracy of self-reported medical history is impacted by the individual's age, health status and formal education, with higher accuracy linked to younger age, better health and higher education^{14–17}. Potentially due to increased monitoring of heart disease and related conditions, higher body mass index (BMI) may also be associated with higher accuracy of self-reported medical history^{18,19}.

In practice, missed cases, under- and over-reporting are inevitable regardless of data source used. However, the key question remains as to how and under what conditions different data resources can be used for research. Detailed knowledge of the limitations of different data sources can help to critically interpret results and draw conclusions from large-scale population-based research and clinical studies. Therefore, the aim of this study is to compare self-reported and register-based hospital medical data in the Swedish NPR on comorbidities in a large breast cancer screening cohort of women in Sweden. We focused on common comorbidities such as hypertension, hyperlipidemia, heart failure, myocardial infarction, angina, stroke and type I or II diabetes. In addition, we examined concordance between self-reported and NPR data on several women's health problems which are less studied, such as preeclampsia and polycystic ovaries or ovarian cysts.

Methods

Study population. The KARolinska MAmnography Project for Risk Prediction of Breast Cancer (KARMA) study (<http://karmastudy.org/>) was set up to be a well-characterized breast cancer cohort²⁰. Participants of the prospective KARMA study comprise women attending mammography screening or clinical mammography at four hospitals in Sweden (Stockholm South General Hospital, Helsingborg Hospital, Skåne University Hospital, Lund, and Landskrona Hospital). Since 1994, all women in Sweden, aged 40–74 years, are invited for publicly-funded mammography screening every 18–24 months. Adherence to mammography screening is high – three in four eligible women attend screening regularly. All women who were invited for screening between January 2011 and March 2013, at the four hospitals were invited to participate in the KARMA study. Additionally, women who had a clinical mammography (i.e. woman being referred for a mammogram because of symptom noticed by her and/or her doctor) at any of the participating mammography units during the recruitment period were invited. Of 210,233 women who were invited to participate in the KARMA study, 70,877 (34%) were enrolled. These women answered a detailed web questionnaire (https://karmastudy.org/wp-content/uploads/2015/07/Karma_baseline_questionnaire_eng.pdf) on background and lifestyle risk factors. Consent was obtained for the retrieval of data from medical records and national registers. Ninety-two percent ($n = 65,231$) of the enrolled women completed the web questionnaire. The ethical review board in Stockholm approved the study (2010/958–31/1) and all study procedures were performed in accordance with relevant guidelines and regulations.

Data sources of medical history. In KARMA, self-reported data was collected for the following conditions: high blood pressure (hypertension), high blood cholesterol (hyperlipidemia), myocardial infarction, angina, heart failure, stroke, polycystic ovaries or ovarian cysts, preeclampsia, and diabetes. Given instructions to “choose all that apply”, the participants were asked - “Have you ever been diagnosed with [any of the medical conditions] by a medical doctor?” For each condition, participants would mark the corresponding checkbox (i.e. yes) if they have ever been diagnosed with the specified conditions, unmarked checkbox corresponds to not having the condition (i.e. no). Women who responded “Don't know/Refuse” to the question were excluded from further analysis ($n = 270$).

All participating women in KARMA study were electronically linked to the NPR (linkage date 1st October 2013) through unique personal identity numbers (i.e. personnummer)²¹. Inpatient/outpatient diagnoses (main and secondary) of the nine medical conditions studied were identified using International Classification of Diseases (ICD) diagnosis codes (see Supplementary Table 1). Diagnoses registered after the date of completion of questionnaire were excluded.

Other covariates. Self-reported information on age at time of survey, education level, BMI and smoking were derived from the KARMA web questionnaire.

Statistical analysis. The count and percent of diseases for each data source (i.e. self-reported, and NPR) and various combinations of each medical history status combination from the two data sources (i.e., “Self-reported No/ NPR No”, “Self-reported Yes/ NPR No”, “Self-reported No/ NPR Yes” and “Self-reported Yes/ NPR Yes”) were computed. In the comparison of self-reported and NPR we report the difference in prevalence instead of percentage difference. This was chosen to avoid the impression of either source is the gold standard. Both methods gave similar results. The “epi.kappa” function in the “epiR” package was used to compute observed proportion of agreement (overall agreement), expected proportion of agreement, prevalence index, bias index, prevalence and bias corrected kappa statistic and Cohen's Kappa statistic in R (version 3.4.2). The prevalence index $(\frac{[y/y] - [n/n]}{N})$, from the cells of a standard 2×2 matrix, Supplementary Method), an estimate of the difference in the probability of the condition being present and absent from the study population, ranges from -1 to 1 and equates to 0 when 50% of the study population has the condition. A larger absolute prevalence index value results in larger chance

agreement and smaller Kappa value²². A bias index $\frac{[y/n] - [n/y]}{N}$, the difference in the reported proportion of the condition being present between NPR and self-reported, ranges from -1 to 1 and large absolute values indicate bias increasing κ , whereas zero bias index indicates equal marginal proportions and no bias²². Kappa coefficients have the following interpretations: ≤ 0 : no agreement; 0.01 – 0.20 : slight agreement; 0.21 – 0.40 : fair agreement; 0.41 – 0.60 : moderate agreement; 0.61 – 0.80 : substantial agreement; 0.81 – 1.00 : almost perfect agreement²³. Proportion of positive specific agreement, the proportion that tests positive under both criteria compared with the average proportion that test positive under each criteria separately $\left(\frac{2 \times [y/y]}{N + [y/y] - [n/n]}\right)$ was expressed as percentage (100 denotes perfect agreement)²⁴. The positive specific agreement is the inverse transformed mean of the sensitivity and positive predictive values, i.e., sensitivity agnostic to which measure is the gold standard (Supplementary Method). Similarly negative specific agreement, the proportion that tests negative under both criteria compared with the average proportion that test negative under each criteria separately $\left(\frac{2 \times [n/n]}{N - [y/y] + [n/n]}\right)$ was expressed as percentage.

Two-stage logistic regression models for analysing agreement which takes into account chance agreement were used to identify potential predictors of overall agreement (i.e. “Self-reported No/ NPR No”, and “Self-reported Yes/NPR Yes” were coded as “1”, otherwise coded as “0”)²⁵. The offset term was determined by the variable(s) in the main model. The 95% confidence interval was obtained using the bootstrap approach, the 2.5 and 97.5 percentile of 2000 iterations. The following predictors were considered: age, education level, BMI, and reported ever smoked for one year or 100 cigarettes (smoking). Stratified analyses were carried out for each condition by age (<50 , 50 – 59 or ≥ 60 years), education level (elementary, intermediate or university), BMI (<25 or ≥ 25 kg/m²), and smoking (Yes or No). A subset of parous women (i.e. who had at least one full-term pregnancy) was used to study preeclampsia. Statistical significance threshold was set at $P < 0.05$.

Ethics Approval And Consent To Participate. All participants signed informed consent forms, and the ethical review board at Karolinska Institutet approved the study (2010/958-31/1). All study procedures were performed in accordance with relevant guidelines and regulations.

Results

A total of 64,961 women, aged between 21 to 87 years [mean (SD): 54.8 (10.0) years] were analysed. Descriptive statistics of participants are given in Table 1. Approximately half of the participants completed university education ($n = 29,400$; 45.3%), reported a BMI below 25 kg/m² ($n = 35,700$, 55.0%) and smoking status yes ($n = 34,274$, 52.8%) in the questionnaire.

According to the self-reported data, the five most commonly diagnosed conditions were hypertension (19.8%), hyperlipidemia (10.8%), polycystic ovaries or ovarian cysts (9.2%), preeclampsia (5.0%) and diabetes (2.8%) (Table 1). The remaining conditions (i.e. heart failure, myocardial infarction, angina and stroke) each affected $\sim 1\%$ of the study population. When comparing self-reported data to the NPR, the largest differences were observed for hypertension (11.2% more in self-reported data), angina (10.8% less in self-reported data), hyperlipidemia (9.6% more in self-reported data) and polycystic ovaries or ovarian cysts (4.1% more in self-reported data) (Table 1). Differences between self-reported and NPR data were minimal ($<1.0\%$) for heart failure, myocardial infarction, stroke, preeclampsia and diabetes.

Figure 1 shows estimates of Cohen’s Kappa for commonly diagnosed conditions. Relative cell counts, expected proportion of agreement, prevalence index, bias index, and Cohen’s Kappa statistic are presented in Supplementary Tables 2–6. Substantial agreement (Cohen’s Kappa) was observed for myocardial infarction (0.74), diabetes (0.71) and stroke (0.64). Moderate agreement was observed for preeclampsia (0.51) and hypertension (0.46). Fair agreement was observed for heart failure (0.40) and polycystic ovaries or ovarian cysts (0.27). For hyperlipidemia (0.14) and angina (0.10), slight agreement was observed between self-reported and NPR data. High levels of overall agreement (i.e. 86.6% or more) were observed for all included conditions (Fig. 2). The average agreement between self-reported and NPR data on absence of medical condition (percent negative specific agreement, range: 92.2–99.8%) was higher than for presence (percent positive specific agreement, range: 11.3–74.4) (Fig. 2).

In multivariate two-stage logistic regression analysis (OR [95% CI]), with age, education, BMI, and smoking status in the models, older age (age ≥ 60 vs age <50 : 1.54 [1.46–1.62]) and higher BMI (≥ 25 vs <25 : 1.06 [1.02–1.10]) were associated with higher agreement for hypertension (Table 2). Similarly, older age (age ≥ 60 vs age <50 : 1.15 [1.12–1.17]) and higher BMI (≥ 25 vs <25 : 1.07 [1.04–1.09]) were associated with higher agreement for angina (Table 2). Older age (age ≥ 60 vs age <50 : 1.15 [1.11–1.19]) was associated with higher agreement for hyperlipidemia (Table 2). Older age (age ≥ 60 vs age <50 : 0.74 [0.60–0.92]) was associated with lower agreement for diabetes (Table 2). No smoking was associated with higher agreement for diabetes (no vs yes: 1.23 [1.07–1.43]) but lower agreement for myocardial infarction (no vs yes: 0.75 [0.57–0.98]) (Table 2). Age, education, BMI, and smoking status were not associated with agreement for heart failure and stroke. For polycystic ovaries or ovarian cysts, older age (age ≥ 60 vs age <50 : 0.89 [0.85–0.93]) was associated with lower agreement, and higher BMI (≥ 25 vs <25 : 1.05 [1.01–1.09]) and no smoking (no vs yes: 1.04 [1.00–1.09]) was associated with higher agreement (Table 2). In the subset of parous women, older age (age ≥ 60 vs age <50 : 0.48 [0.44–0.52]) was associated with lower agreement, and higher education (university vs elementary: 1.15 [1.04–1.28]) and no smoking (no vs yes: 1.09 [1.01–1.17]) were associated with higher agreement for preeclampsia (Table 2).

Fair agreement (0.23) was observed for the number of conditions between self-reported and NPR data. Similar with the more common conditions, hypertension and hyperlipidemia, the associations of older age, higher BMI and lower education were associated with higher agreement, using multinomial modelling for the number of conditions (data not shown).

Characteristics			
Total number of women, <i>n</i> (%)	64,961 (100.0)		
Age at recruitment in years, mean (SD)	54.8 (10.0)		
Education level, <i>n</i> (%)			
Elementary	7,793 (12.0)		
Intermediate	17,721 (27.3)		
University	29,400 (45.3)		
Self-reported BMI at recruitment, <i>n</i> (%)			
<25 kg/m ²	35,700 (55.0)		
≥25 kg/m ²	28,800 (44.3)		
Smoking, <i>n</i> (%)			
Yes, <i>n</i> (%)	34,274 (52.8)		
No, <i>n</i> (%)	30,380 (46.8)		
Prevalence of medical conditions, <i>n</i> (%)	Self-reported	NPR	Difference in prevalence (Self-reported – NPR)
Hypertension	12,880 (19.8)	5,582 (8.6)	11.2
Hyperlipidemia	7,033 (10.8)	819 (1.3)	9.6
Heart failure	373 (0.6)	287 (0.4)	0.1*
Myocardial infarction	482 (0.7)	426 (0.7)	0.1*
Angina	585 (0.9)	7,611 (11.7)	–10.8
Stroke	696 (1.1)	942 (1.5)	–0.4
Polycystic ovaries or ovarian cysts	5,980 (9.2)	3,317 (5.1)	4.1
Preeclampsia ^a	2,930 (5.0)	3,456 (5.9)	–0.9
Diabetes	1,719 (2.6)	1,380 (2.1)	0.5

Table 1. Characteristics of 64,961 women attending mammography units in the KARMA study. BMI: Body mass index; NPR: National patient register; SD: Standard deviation. ^aIncludes women who had at least 1 pregnancy (*n* = 58,936). *Discrepancy due to rounding errors.

Discussion

Misclassification of conditions may result in confounded studies. In the study of survival, the misclassification of comorbid conditions that are associated with the higher risk of death would lead to over-emphasis of the risk of death from the disease of interest. In addition, studies looking at treatment outcomes may be potentially confounded by comorbidities. To systematically examine the appropriateness of using self-reported and register-based hospital medical (NPR) data to identify comorbidities, we compared prevalence and agreement between these two data sources in a large population-based breast cancer cohort in Sweden. Both data sources have their respective strengths and shortcomings. However, the focus of this study is not on whether self-reported personal medical history is “more correct” than NPR records and vice versa, but rather how closely they agree or disagree for various medical conditions.

Few studies have looked at the concordance of preeclampsia between self-reported and hospital data^{26,27} and to the best of our knowledge, ours is the first study investigating the same for polycystic ovaries and ovarian cysts. The Swedish NPR has been used to identify women with preeclampsia and polycystic ovaries and ovarian cysts in epidemiological studies previously^{28,29}. Self-reported data provided more cases of polycystic ovaries or ovarian cysts than the NPR did and there was fair agreement between the sources. The combined classification of polycystic ovarian syndrome and ovarian cysts may have resulted in the higher than expected self-reported occurrence in the older age group. However, the prevalence estimates for preeclampsia were found to be similar for both self-reported and NPR data with moderate amount of agreement. The use of self-reported pre-eclampsia in the older generations of women may be of concern as preeclampsia may have been referred to as “toxemia of pregnancy” in the earlier period. However, in our population we observed similar concordance across the three age groups. These conditions could be potential risk factors or confounders for breast cancer risk, for example, preeclampsia has been shown to be associated with reduced risk of breast cancer³⁰. Therefore, in order to identify women with these conditions reliably, both self-reported and NPR (hospital) data should be explored, if available.

Our study showed that the prevalence of hypertension and hyperlipidemia were highly under-represented in the NPR data when comparing it to self-reported data due to the fact that primary care outpatient records are not included in the NPR data. This is in agreement with previous studies which showed that medical conditions typically treated in primary care settings (not often leading to hospital admission) are not recognised or recorded in the hospital admission data. For example, a concordance study of self-reported and administrative hospital data in Australian Longitudinal Study on Women’s Health showed under-recording of hypertension in the hospital data¹².

For life-threatening conditions like heart failure, myocardial infarction and stroke, the differences in prevalence from self-reported and NPR data were minimal. In contrast to all the comorbidities we have studied in this paper, angina was less reported in the self-reported data, for example, absolute difference in prevalence was ~10%. This might be due to the fact that angina is not a well-defined disease and many people misclassify it because its symptoms are similar to other disease (e.g. myocardial infarction) and it is perceived as a symptom,

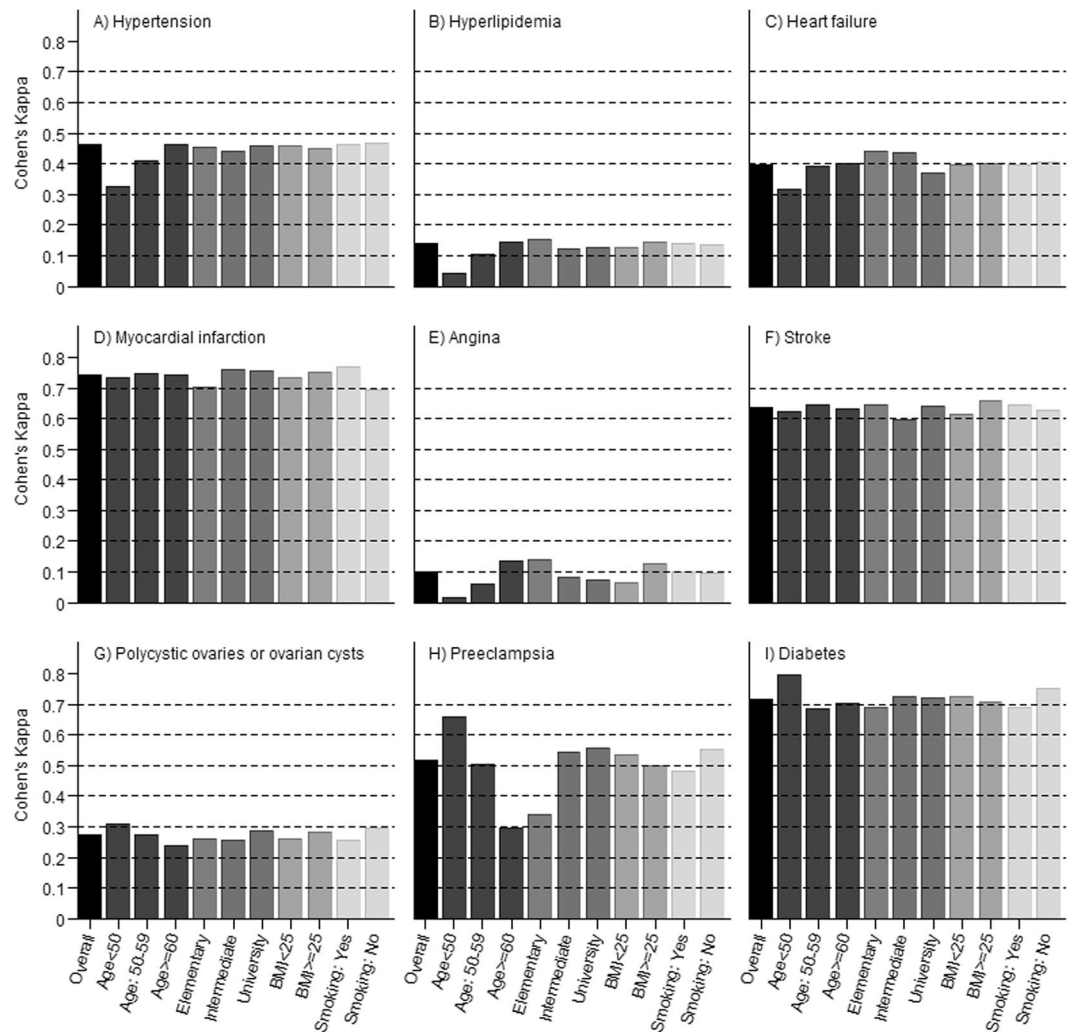


Figure 1. Cohen's Kappa for commonly diagnosed conditions.

not a disease³¹. Subsequently, the agreement between self-reported angina and NPR recorded angina was poor in our study.

In spite of heterogeneous methodology and comparisons, we observed common findings among previously published studies – higher agreement for medical conditions that are widely recognized and easily diagnosed (e.g. diabetes, hypertension) or require hospital care (e.g. myocardial infarction, stroke), and lower agreement for poorly defined diseases (e.g. heart failure, angina), conditions perceived as symptoms (e.g. angina) and conditions that may not require hospitalization (e.g. hyperlipidemia). Okura *et al.* measured the agreement between self-reported cardiovascular disease and extensive medical records with high completeness (including hospital inpatient or outpatient care, office visits, emergency room and nursing home care and death certificate and autopsy information) and long archival period for ~2,000 participants from the Olmsted County in Minnesota and found substantial agreement for diabetes, hypertension, myocardial infarction and stroke (Kappa values ranging from 0.71 to 0.80)¹⁵. Moderate agreement was observed for heart failure (Kappa 0.46)¹⁵. Hamood *et al.* assessed agreement for self-reported medical history and electronic medical records (including primary and hospital care) for 119 breast cancer patients and found almost perfect agreement for diabetes (Kappa 0.93), moderate agreement for stroke (Kappa 0.79), hypertension (Kappa 0.55) and hyperlipidemia (Kappa 0.46)¹⁸. Agreement between self-reported and primary care data presented by Hansen *et al.* based on the MultiCare Cohort Study ($n = 3,189$) was found to be substantial for diabetes (Kappa 0.80), moderate for hypertension (Kappa 0.56) and stroke (Kappa 0.55) and fair for hyperlipidemia (Kappa 0.36)³². Huerta *et al.* compared self-reported diabetes, hypertension and hyperlipidemia with biometric data (levels of blood glucose and lipids and blood pressure) and found substantial (Kappa 0.78), moderate (Kappa 0.51) and fair agreement (Kappa 0.27) for the three conditions, respectively. Nonetheless, as low prevalence may result in high chance agreement, and consequently, low Kappa, caution should be exercised when interpreting statistics for less common conditions.

Overall agreement is a common measure of agreement between self-reported and hospital data^{15,33}. Based on overall agreement, self-reported and NPR were concordant for 86.6% or more of the participants for all nine comorbidities studied. The high overall agreement observed in our study is mainly driven by the high negative

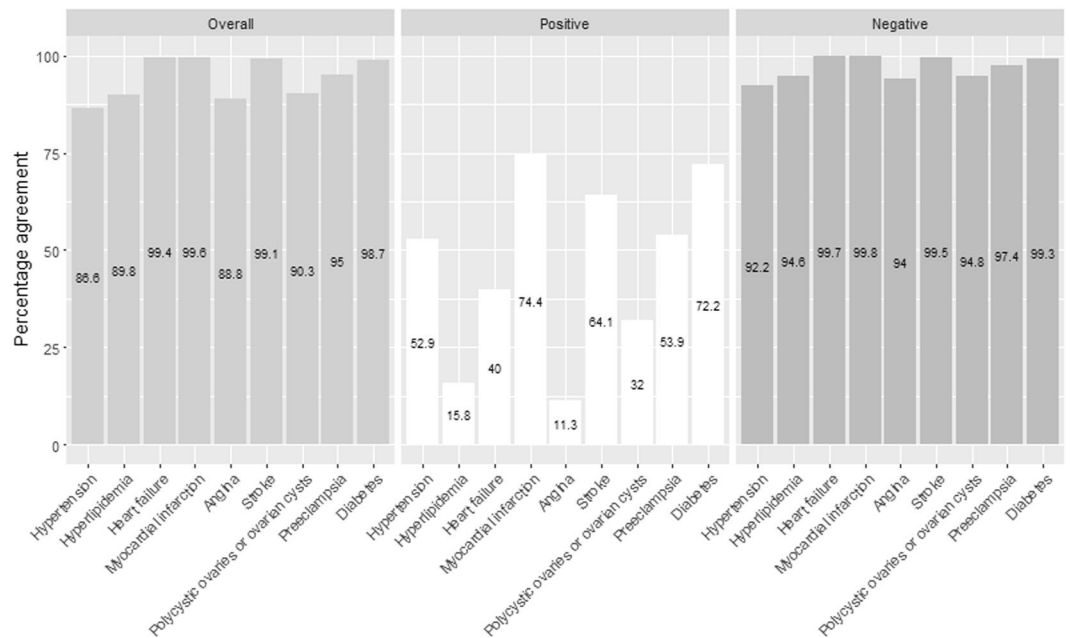


Figure 2. Percentage overall, positive and negative specific agreement between self-reported and register-based hospital medical data.

Condition	Age 50–59 (n = 18,797) vs Age <50 (n = 23,154)	Age ≥60 (n = 23,010) vs Age <50 (n = 23,154)	Intermediate (n = 17,721) vs Elementary (n = 7,793)	University (n = 29,400) vs Elementary (n = 7,793)	BMI ≥ 25 (n = 28,800) vs BMI < 25 (n = 35,700)	Smoking No (n = 30,380) vs Smoking Yes (n = 34,274)
<i>Unadjusted</i>						
Hypertension	1.24 (1.18–1.31)	1.59 (1.51–1.67)	0.86 (0.80–0.91)	0.87 (0.82–0.92)	1.11 (1.07–1.15)	0.98 (0.94–1.02)
Hyperlipidemia	1.08 (1.04–1.11)	1.16 (1.13–1.20)	0.94 (0.90–0.98)	0.95 (0.91–0.98)	1.03 (1.00–1.06)	0.99 (0.97–1.02)
Heart failure	1.12 (0.80–1.51)	1.15 (0.85–1.49)	0.98 (0.76–1.25)	0.88 (0.71–1.09)	1.01 (0.86–1.19)	1.01 (0.87–1.20)
Myocardial infarction	1.05 (0.59–1.80)	1.05 (0.62–1.65)	1.24 (0.88–1.79)	1.20 (0.86–1.68)	1.07 (0.81–1.37)	0.75 (0.57–0.98)
Angina	1.05 (1.03–1.07)	1.17 (1.15–1.19)	0.92 (0.88–0.95)	0.91 (0.87–0.94)	1.08 (1.06–1.10)	0.99 (0.97–1.01)
Stroke	1.07 (0.81–1.42)	1.06 (0.82–1.35)	0.87 (0.70–1.08)	0.97 (0.79–1.19)	1.13 (0.96–1.33)	0.95 (0.81–1.10)
Polycystic ovaries or ovarian cysts	0.95 (0.90–1.00)	0.89 (0.85–0.94)	1.00 (0.93–1.06)	1.04 (0.97–1.10)	1.04 (1.00–1.08)	1.06 (1.01–1.10)
Preeclampsia ^a	0.67 (0.62–0.73)	0.45 (0.42–0.49)	1.50 (1.35–1.67)	1.53 (1.39–1.68)	0.95 (0.88–1.01)	1.18 (1.10–1.27)
Diabetes	0.66 (0.52–0.83)	0.71 (0.58–0.87)	1.09 (0.90–1.32)	1.07 (0.89–1.27)	0.98 (0.84–1.14)	1.26 (1.09–1.45)
<i>Multivariable-adjusted, all covariates shown</i>						
Hypertension	1.24 (1.17–1.31)	1.54 (1.46–1.62)	0.99 (0.93–1.06)	1.00 (0.94–1.06)	1.06 (1.02–1.10)	1.01 (0.97–1.05)
Hyperlipidemia	1.07 (1.04–1.11)	1.15 (1.11–1.19)	0.97 (0.93–1.01)	0.97 (0.94–1.01)	1.02 (0.99–1.05)	1.00 (0.97–1.03)
Heart failure	1.13 (0.80–1.52)	1.16 (0.85–1.50)	1.00 (0.77–1.29)	0.89 (0.71–1.12)	1.00 (0.84–1.18)	1.02 (0.87–1.22)
Myocardial infarction	1.02 (0.57–1.75)	1.07 (0.63–1.70)	1.26 (0.88–1.84)	1.24 (0.88–1.77)	1.08 (0.82–1.39)	0.75 (0.57–0.98)
Angina	1.04 (1.03–1.07)	1.15 (1.12–1.17)	0.97 (0.93–1.01)	0.96 (0.92–0.99)	1.07 (1.04–1.09)	1.01 (0.99–1.03)
Stroke	1.05 (0.79–1.39)	1.00 (0.77–1.28)	0.87 (0.69–1.08)	0.99 (0.80–1.22)	1.13 (0.96–1.33)	0.95 (0.81–1.11)
Polycystic ovaries or ovarian cysts	0.95 (0.90–1.00)	0.89 (0.85–0.93)	0.95 (0.89–1.01)	1.00 (0.93–1.06)	1.05 (1.01–1.09)	1.04 (1.00–1.09)
Preeclampsia ^a	0.69 (0.63–0.75)	0.48 (0.44–0.52)	1.12 (1.00–1.25)	1.15 (1.04–1.28)	1.02 (0.95–1.09)	1.09 (1.01–1.17)
Diabetes	0.68 (0.53–0.85)	0.74 (0.60–0.92)	1.05 (0.86–1.28)	1.01 (0.84–1.22)	1.00 (0.85–1.17)	1.23 (1.07–1.43)

Table 2. Odds ratio and corresponding 95% confidence intervals for overall agreement in each medical condition. Significant associations ($P < 0.05$) are denoted in bold. BMI: body mass index. ^aIncludes women who had at least 1 full-term pregnancy ($n = 58,936$).

specific agreement (>92%) for all comorbidities studied. In addition, conditions with higher proportion of positive specific agreement had higher Kappa. This may be an indication that we might be limited in identifying comorbidities when we use only one source of information. Factors associated with overall agreement tend to have similar association with positive specific agreement (Supplementary Table 7).

Previously, Ye *et al.* argued that the number of comorbidities increases with age, leading to lower precision between self-reported and medical records³³. We observed fair agreement between self-reported and NPR data.

However, we observed lower overall agreement with increased age for polycystic ovaries or ovarian cysts and preeclampsia after accounting for other factors such as education, BMI, smoking and breast cancer history, better overall agreement was observed for hypertension, hyperlipidemia and angina. Our results suggest that relationship between overall agreement and age is likely due to the length of time between disease diagnosis and study entry, as polycystic ovaries or ovarian cysts and preeclampsia are typically diagnosed at much younger ages than hypertension, hyperlipidemia or angina.

Similar to the work of Ye *et al.*³³, we did not find education level to be a predictor of overall agreement in general, with preeclampsia being the only exception. This finding should be interpreted in the light of education in Sweden being mandatory for all children between ages 7 to age 16. In addition, higher education is available at no cost for Swedish citizens. It is unclear why better overall agreement was observed for preeclampsia. Nonetheless, women with higher education may be privileged with higher health literacy, which in turn puts them in a better position to understand information conveyed to them by physicians.

Short *et al.* hypothesized that higher BMI is correlated with lower agreement between self-reported values of healthcare utilization and administrative claims¹⁴. It was suggested that there might be a tendency for people with higher BMI to use more healthcare services, making it less likely for them to accurately recall and report doctor visits and inpatient hospital admissions¹⁴. However, in our study, self-reported diagnoses for several diseases were more likely to be confirmed by NPR data in women with higher BMI. A possible explanation may be related to the high education among women in general and also greater health consciousness of women enrolled in KARMA; they may have been more aware of the risk of chronic diseases associated with obesity. Other studies have also shown that better health status is associated with better agreement¹⁴. This is supported by the higher agreement observed for (polycystic ovaries or ovarian cysts, preeclampsia, and diabetes) non-smokers in our results.

The main strengths of our study include a women-only cohort, the large sample size and resulting statistical power. An electronic linkage with NPR provided complete follow-up for virtually every woman in the cohort. Although our study base comprises of women attending screening or clinical mammography, the publicly funded health care system in Sweden means that all residents have access to health care and socioeconomic bias in hospital admission is very unlikely. Nonetheless, a number of limitations warrant discussion. For example, register-based diagnoses can be complemented by information from the Swedish Drug Prescription Register (e.g. beta blockers to indicate hypertension and lipid lowering agents to indicate hyperlipidemia)³⁴. However, while the Swedish Drug Prescription Register contains information regarding drug utilization and expenditures for dispensed prescribed drugs in the entire Swedish population, it was established fairly recently in July 2005 (i.e. too young to be used)³⁴. There are other conditions that may be of interest, however we were limited to those in the KARMA questionnaire. Our study consisted of a highly educated population that is well-served by a mainly government-funded and decentralized health care system. It is unclear whether the results can be generalized to other populations with different health-seeking behaviour and access to healthcare. In addition, two inherent disadvantages with agreement measures must be taken into account when interpreting the results. Firstly, we acknowledge that there is no clear reference standard for the ascertainment of the medical conditions. Secondly, when the disease prevalence in the population is very high or low, the value of Cohen's Kappa may indicate poor reliability even with a high observed proportion of agreement^{35–37} (i.e. agreement results are dependent on the disease prevalence in the study population). We have thus reported multiple measures of agreement to take into account bias, prevalence and possible imbalance in each 2 × 2 table's marginal totals to address this paradox of the Kappa statistic.

Conclusions

An increasing number of breast cancer cohort studies^{38,39} are including self-reported comorbidities in the data collection forms, prompting an investigation into how well the data from self-reported questionnaires correspond to register-based hospital medical data such as the Swedish NPR. Our study confirmed that on comorbidities of stroke and myocardial infarction, there is substantial overall agreement between registry data and self-reported data, regardless of age, education, and BMI. Older age was associated with better overall agreement on comorbidities of hypertension, hyperlipidemia and angina, but poorer overall agreement for polycystic ovaries or ovarian cysts, and preeclampsia. In most subgroups, negative specific agreement between registry data and self-reported data is >90%, which suggests that both sources can confidently identify individuals without the conditions studied in this subgroups.

Data Availability

The datasets used and/or analysed during the current study are available through an application for Karma Data Access (<https://karmastudy.org/data-access/>) on reasonable request.

References

1. Wu, A. H. *et al.* Diabetes and other comorbidities in breast cancer survival by race/ethnicity: the California Breast Cancer Survivorship Consortium (CBCSC). *Cancer Epidemiol Biomarkers Prev* **24**, 361–368, <https://doi.org/10.1158/1055-9965.EPI-14-1140> (2015).
2. Koroukian, S. M., Bakaki, P. M., Schluchter, M. D. & Owusu, C. Treatment and survival patterns in relation to multimorbidity in patients with locoregional breast and colorectal cancer. *Journal of geriatric oncology* **2**, 200–208, <https://doi.org/10.1016/j.jgo.2011.02.004> (2011).
3. Gurney, J., Sarfati, D. & Stanley, J. The impact of patient comorbidity on cancer stage at diagnosis. *British Journal of Cancer* **113**, 1375–1380, <https://doi.org/10.1038/bjc.2015.355> (2015).
4. Yood, M. U. *et al.* Mortality impact of less-than-standard therapy in older breast cancer patients. *J Am Coll Surg* **206**, 66–75, <https://doi.org/10.1016/j.jamcollsurg.2007.07.015> (2008).
5. Giordano, S. H., Duan, Z., Kuo, Y. F., Hortobagyi, G. N. & Goodwin, J. S. Use and outcomes of adjuvant chemotherapy in older women with breast cancer. *J Clin Oncol* **24**, 2750–2756, <https://doi.org/10.1200/JCO.2005.02.3028> (2006).

6. Gold, H. T., Do, H. T. & Dick, A. W. Correlates and effect of suboptimal radiotherapy in women with ductal carcinoma *in situ* or early invasive breast cancer. *Cancer* **113**, 3108–3115, <https://doi.org/10.1002/cncr.23923> (2008).
7. Muggah, E., Graves, E., Bennett, C. & Manuel, D. G. Ascertainment of chronic diseases using population health data: a comparison of health administrative data and patient self-report. *BMC Public Health* **13**, 16, <https://doi.org/10.1186/1471-2458-13-16> (2013).
8. Webster, P. C. Sweden's health data goldmine. *CMAJ* **186**, E310, <https://doi.org/10.1503/cmaj.109-4713> (2014).
9. Ludvigsson, J. F. *et al.* External review and validation of the Swedish national inpatient register. *BMC Public Health* **11**, 450, <https://doi.org/10.1186/1471-2458-11-450> (2011).
10. Brunstrom, M. Hypertension, the Swedish Patient Register, and Selection Bias. *JAMA Intern Med* **176**, 862–863, <https://doi.org/10.1001/jamainternmed.2016.1556> (2016).
11. Anwar, H. *et al.* Assessment of the under-reporting of diabetes in hospital admission data: a study from the Scottish Diabetes Research Network Epidemiology Group. *Diabetic Medicine* **28**, 1514–1519, <https://doi.org/10.1111/j.1464-5491.2011.03432.x> (2011).
12. Navin Cristina, T. J., Stewart Williams, J. A., Parkinson, L., Sibbritt, D. W. & Byles, J. E. Identification of diabetes, heart disease, hypertension and stroke in mid- and older-aged women: Comparing self-report and administrative hospital data records. *Geriatrics & Gerontology International* **16**, 95–102, <https://doi.org/10.1111/ggi.12442> (2016).
13. Qureshi, Z. P., Ganz, P. A. & Bennett, C. L. Improving the Evidence Base for Delivery of High-Quality Cancer Care. *JAMA Oncol* **3**, 1029–1031, <https://doi.org/10.1001/jamaoncol.2016.6722> (2017).
14. Short, M. E. *et al.* How accurate are self-reports? Analysis of self-reported health care utilization and absence when compared with administrative data. *J Occup Environ Med* **51**, 786–796, <https://doi.org/10.1097/JOM.0b013e3181a86671> (2009).
15. Okura, Y., Urban, L. H., Mahoney, D. W., Jacobsen, S. J. & Rodeheffer, R. J. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. *J Clin Epidemiol* **57**, 1096–1103, <https://doi.org/10.1016/j.jclinepi.2004.04.005> (2004).
16. Raina, P., Torrance-Rynard, V., Wong, M. & Woodward, C. Agreement between Self-reported and Routinely Collected Health-care Utilization Data among Seniors. *Health Services Research* **37**, 751–774, <https://doi.org/10.1111/1475-6773.00047> (2002).
17. Mackenbach, J. P., Looman, C. W. & van der Meer, J. B. Differences in the misreporting of chronic conditions, by level of education: the effect on inequalities in prevalence rates. *Am J Public Health* **86**, 706–711 (1996).
18. Hamood, R., Hamood, H., Merhasin, I. & Keinan-Boker, L. A feasibility study to assess the validity of administrative data sources and self-reported information of breast cancer survivors. *Isr J Health Policy Res* **5**, 50, <https://doi.org/10.1186/s13584-016-0111-6> (2016).
19. Jackson, J. M. *et al.* Validity of diabetes self-reports in the Women's Health Initiative. *Menopause* **21**, 861–868, <https://doi.org/10.1097/gme.000000000000189> (2014).
20. Gabrielson, M. *et al.* Cohort profile: The Karolinska Mammography Project for Risk Prediction of Breast Cancer (KARMA). *Int J Epidemiol*, <https://doi.org/10.1093/ije/dyw357> (2017).
21. Ludvigsson, J. F., Otterblad-Olausson, P., Pettersson, B. U. & Ekblom, A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol* **24**, 659–667, <https://doi.org/10.1007/s10654-009-9350-y> (2009).
22. Petersen, H. H., Enøe, C. & Nielsen, E. O. Observer agreement on pen level prevalence of clinical signs in finishing pigs. *Preventive Veterinary Medicine* **64**, 147–156, <https://doi.org/10.1016/j.prevetmed.2004.05.002> (2004).
23. McHugh, M. L. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)* **22**, 276–282 (2012).
24. Hripcsak, G. & Heitjan, D. F. Measuring agreement in medical informatics reliability studies. *Journal of Biomedical Informatics* **35**, 99–110, [https://doi.org/10.1016/S1532-0464\(02\)00500-2](https://doi.org/10.1016/S1532-0464(02)00500-2) (2002).
25. Lipsitz, S. R., Parzen, M., Fitzmaurice, G. M. & Klar, N. A two-stage logistic regression model for analyzing inter-rater agreement. *Psychometrika* **68**, 289–298, <https://doi.org/10.1007/bf02294802> (2003).
26. Coolman, M. *et al.* Medical record validation of maternally reported history of preeclampsia. *Journal of Clinical Epidemiology* **63**, 932–937, <https://doi.org/10.1016/j.jclinepi.2009.10.010> (2010).
27. Klemmensen, Å. K., Olsen, S. F., Østerdal, M. L. & Tabor, A. Validity of Preeclampsia-related Diagnoses Recorded in a National Hospital Registry and in a Postpartum Interview of the Women. *American Journal of Epidemiology* **166**, 117–124, <https://doi.org/10.1093/aje/kwml39> (2007).
28. Ludvigsson, J. F. *et al.* A Population-based Cohort Study of Pregnancy Outcomes Among Women With Primary Sclerosing Cholangitis. *Clinical Gastroenterology and Hepatology* **12**, 95–100.e101, <https://doi.org/10.1016/j.cgh.2013.07.011> (2014).
29. Lovvik, T. S. *et al.* Pregnancy and perinatal outcomes in women with polycystic ovary syndrome and twin births: a population-based cohort study. *BJOG* **122**, 1295–1302, <https://doi.org/10.1111/1471-0528.13339> (2015).
30. Terry, M. B. *et al.* Preeclampsia, Pregnancy-related Hypertension, and Breast Cancer Risk. *American Journal of Epidemiology* **165**, 1007–1014, <https://doi.org/10.1093/aje/kwk105> (2007).
31. Barbara, A. M., Loeb, M., Dolovich, L., Brazil, K. & Russell, M. Agreement between self-report and medical records on signs and symptoms of respiratory illness. *Prim Care Respir J* **21**, 145–152, <https://doi.org/10.4104/pcrj.2011.00098> (2012).
32. Hansen, H. *et al.* Agreement between self-reported and general practitioner-reported chronic conditions among multimorbid patients in primary care - results of the MultiCare Cohort Study. *BMC Fam Pract* **15**, 39, <https://doi.org/10.1186/1471-2296-15-39> (2014).
33. Ye, F. *et al.* Comparison of Patient Report and Medical Records of Comorbidities: Results From a Population-Based Cohort of Patients With Prostate Cancer. *JAMA Oncol* **3**, 1035–1042, <https://doi.org/10.1001/jamaoncol.2016.6744> (2017).
34. Wettermark, B. *et al.* The new Swedish Prescribed Drug Register—opportunities for pharmacoepidemiological research and experience from the first six months. *Pharmacoepidemiol Drug Saf* **16**, 726–735, <https://doi.org/10.1002/pds.1294> (2007).
35. Sim, J. & Wright, C. C. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* **85**, 257–268 (2005).
36. Cicchetti, D. V. & Feinstein, A. R. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* **43**, 551–558 (1990).
37. Lantz, C. A. & Nebenzahl, E. Behavior and interpretation of the kappa statistic: resolution of the two paradoxes. *J Clin Epidemiol* **49**, 431–434 (1996).
38. Kwan, M. L. *et al.* The Pathways Study: a prospective study of breast cancer survivorship within Kaiser Permanente Northern California. *Cancer causes & control: CCC* **19**, 1065–1076, <https://doi.org/10.1007/s10552-008-9170-5> (2008).
39. Flesch-Janys, D. *et al.* Risk of different histological types of postmenopausal breast cancer by type and regimen of menopausal hormone therapy. *International journal of cancer* **123**, 933–941, <https://doi.org/10.1002/ijc.23655> (2008).

Acknowledgements

We thank Hanis Mariyah Mohd Ishak for proof-reading our manuscript. KARMA was financed by the Märkt and Hans Rausing's Initiative Against Breast Cancer and the Kamprad Family Foundation. JL is a recipient of a Singapore National Research Foundation Fellowship (NRF-NRFF2017-02). MH is a recipient of the National Medical Research Council Clinician Scientist Award (Senior Investigator Category) [NMRC/CSA-SI/0015/2017], National University Cancer Institute Singapore Centre Grant Programme [CGAug16M005], and Breast Cancer Prevention Programme funded under SSHSPH-Res-Prog.

Author Contributions

J.L., K.C. and P.H. conceived and designed the study. J.L. and P.J.H. performed the analyses and wrote the manuscript. P.J.H. and C.S.T. provided statistical expertise and were involved in data interpretation. S.R.S., J.F.L., E.P., H.M., M.H., M.E., L.Y.L., K.S.C., P.H. and K.C. provided clinical and epidemiological expertise. All authors critically reviewed and approved the manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-019-40072-0>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019