Agreement between reported questionnaire data and medical records on diagnosis and COVID-19 symptoms at onset

Marta A Kisiel, Claes Kock, Josef Sultan, Helena Janols, Christer Janson & Ronnie Pingel

To cite this article: Marta A Kisiel, Claes Kock, Josef Sultan, Helena Janols, Christer Janson & Ronnie Pingel (2023) Agreement between reported questionnaire data and medical records on diagnosis and COVID-19 symptoms at onset, European Clinical Respiratory Journal, 10:1, 2282251, DOI: 10.1080/20018525.2023.2282251

To link to this article: https://doi.org/10.1080/20018525.2023.2282251

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 18 Nov 2023.

Submit your article to this journal

Article views: 152

View related articles

View Crossmark data
Agreement between reported questionnaire data and medical records on diagnosis and COVID-19 symptoms at onset

Marta A Kisiel, Claes Kock, Josef Sultan, Helena Janols, Christer Janson and Ronnie Pingel

Department of Medical Sciences: Occupational and Environmental Medicine, Uppsala University, Uppsala, Sweden; Department of Statistics, Uppsala University, Uppsala, Sweden; Medical School, Uppsala University, Uppsala, Sweden; Department of Medical Sciences: Section of Infectious Diseases, Uppsala University, Uppsala, Sweden; Department of Medical Sciences: Respiratory, Allergy and Sleep Research, Uppsala University, Uppsala, Sweden

ABSTRACT
The aim of this study was to assess whether there was agreement between self-reported data in a survey and medical records regarding diagnoses and symptoms at COVID-19 onset. The impact of sociodemographic factors on agreement between the two data sources was also assessed. Cross-sectional data were extracted from a Swedish longitudinal cohort study. In total, 401 non-hospitalized patients with a polymerase chain reaction-confirmed COVID-19 infection responded to a survey and agreed to a review of their electronic medical records. Agreement, estimated using the kappa statistic, sensitivity, and specificity were calculated for nine diagnoses and eleven symptoms. Differences between subgroups based on sociodemographic factors were assessed. The agreement between the self-reported data and medical records was at a substantial to moderate level for diagnoses such as diabetes mellitus (kappa 0.65, sensitivity 86%) and hypertension (kappa 0.59, sensitivity 56%) and at a fair level for more difficult-to-define conditions such as ongoing immunosuppressive treatment (kappa 0.27, sensitivity 25%). The agreement between the two data sources on symptoms was between fair and poor (kappa 0.36 for fever; kappa 0.05 for fatigue). Agreement for some diagnoses and symptoms varied across some sociodemographic subgroups, e.g., agreement in diabetes mellitus was significantly better in males (kappa 1.0) than females (kappa 0.52, homogeneity tests $p = 0.02$). In general, kappa values were lower for symptoms than diagnoses. The agreement between the two sources varied with diagnoses and symptoms and was also influenced by sociodemographic factors. This study illustrates that it is important to consider type of data used in the epidemiological studies as different information sources differ with quality and accuracy.

CONTACT Marta A Kisiel marta.kisiel@medsci.uu.se Department of Medical Sciences: Occupational and Environmental Medicine, Uppsala University, Uppsala, Sweden

ARTICLE HISTORY Received 7 August 2023 Accepted 7 November 2023

KEYWORDS Agreement between the two data sources; kappa statistic; sensitivity and specificity; symptoms at COVID-19; diagnosis prior to COVID-19

Introduction
Epidemiological studies are often based on multiple data sources. Collecting self-reported information from questionnaires and obtaining data from medical and administrative records are two commonly used methods in clinical research [1–3]. Under ideal circumstances, asking a patient and reviewing records should yield the same information. However, in reality, different data sources have different quality of data. Self-reported data may lack accuracy and may be affected by recall bias and differing willingness to report or differing understanding of questions [4]. On the other hand, collecting data from patient records can be more time- and resource-consuming and accuracy can be limited by incomplete documentation [5].

A number of studies have investigated the agreement between self-reported data and medical records. However, comparison of these studies is difficult due to heterogenous methodologies including type of population and data sources used [2,6]. Previous studies have found that the agreement between self-reported data and medical records may depend on the disease and the sociodemographic status of the responders [7,8]. There are some studies on symptom agreement for angina pectoris [9], cardiovascular disease [10], and cancer [1]. However, little is known about the agreement in symptoms of respiratory infections [6] and, thus far, no studies have been performed regarding coronavirus disease 2019 (COVID-19).

The aim of this study was to investigate the agreement between self-reported data and data from electronic medical records, including primary and hospital care, for diagnoses and symptoms at COVID-19 onset. We
also evaluated if agreement was related to specific social determinants of health such as age, sex, educational level, country of birth, marital status, and occupation.

**Material and method**

**Study design**

This was a cross-sectional study of data collected using a survey and from electronic patient records within the longitudinal “COMBAT post-COVID” project conducted at Uppsala University Hospital, Sweden. COMBAT, described elsewhere, aimed to follow up non-hospitalized patients with COVID-19 at the outpatient infectious disease department [11,12]. All the enrolled patients (≥18 years old) were symptomatic at onset and had a positive polymerase chain reaction for the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in nasopharyngeal swabs between 10 March and 30 July 2020, confirming COVID-19. Twelve months after infection, the patients answered a survey on their symptoms at onset and comorbidities, Figure 1. The survey was sent 45–51 weeks after the COVID-19 diagnosis, by e-mail via REDcap [13] and up to two reminders with an interval of 1 week. If no e-mail address was available, the survey was sent to home address by ordinary post, without any reminders, as it would not reach them during the relevant timeframe. The absolute time frame of the survey was between the 10th of March 2021 and the 15th of January 2022.

The patients gave written informal consent on collection of data on their symptoms and comorbidities from their electronic medical records. The study was approved by the institutional ethics committee of the University Hospital in Uppsala (EPN number 2020–05707) and conducted in accordance with the Helsinki Declaration.

**Sociodemographic data**

The survey collected several social determinants of health including: legal sex such as female/male, age, divided into ≤40 and >40 years old, marital status, divided into two groups including married/living with partner and single/divorced/widow/-er, educational level, divided into university education at least 3 years and education lower than university education, country of birth, divided into born in Sweden and born outside Sweden, occupation, divided into two groups: healthcare workers and other occupations.

**Outcomes**

**Survey questionnaire**

Self-reported data on diagnoses and symptoms at COVID-19 onset were based on participant responses to the following two questions:

![Figure 1. Flowcharts of responders. In total, 725 non-hospitalized COVID-19 patients were sent the questionnaire 12 months after their COVID-19 diagnosis. A total of 401 patients responded, making up the study population.](image-url)
Have you ever been told by a doctor that you have any of the following diagnoses/conditions? Hypertension, other heart disease (such as heart failure or acute myocardial infarction), hypo-/hyperthyroidism, diabetes mellitus type 1 or 2, lung disease (including asthma and chronic obstructive lung disease), a condition requiring immunosuppressive treatment (such as oral corticosteroids), cancer, with or without treatment, depression/anxiety, stroke.

Did you have any of the following symptoms at the onset of COVID-19?: Fever (at least 38°C), cough, sore throat, dyspnea, runny nose/nasal congestion, gastrointestinal complaints including vomiting/nausea, diarrhea and stomach pain, headache, loss of taste and smell, chest pain, muscle/joint pain, fatigue.

The diagnoses were selected as they were considered risk factors for severe COVID-19 infection and the symptoms were the most common symptoms at infection onset according to the World Health Organization/International Soil Reference and Information Centre (WHO/ISARIC) platform [14].

**Medical records**

A search of medical records from primary and secondary care from the period January 2010 to January 2022 was undertaken covering the same diagnoses as the survey. We based on the ICD-10 diagnostic codes. Symptoms at COVID-19 onset were also retrieved from medical records. The medical record review was blinded by the results of the survey.

**Statistical analysis**

Categorical variables are presented as proportions and continuous variables as means. The differences between the prevalence rates of diseases and symptoms in self-reported data and those in the medical records in the whole population were assessed using t-tests. Kappa coefficients were calculated in order to measure agreement corrected by random chance between the self-reported data and medical record information on diagnoses and symptoms at COVID-19 onset [15]. Kappa coefficients were interpreted as follows: ≤0: no agreement; 0.01–0.20: poor 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 agreement [16]. Once kappa coefficients had been calculated in the total population, the responders were divided into groups based on sociodemographic background. Differences across the strata were evaluated using a test of homogeneity (Donner’s goodness-of-fit test). As a supplement to kappa values, other measures of agreement were presented: total agreement (correctly reported positive medical records and survey or negative medical records and survey divided by total reported in survey and medical records), sensitivity (correctly reported positive in survey and medical records divided by all reported positive), and specificity (correctly reported negative in survey and medical records divided by all reported negative). For the whole population, true positive (TP), false positive (FP), false negative (FN), true negative (TN) values, positive predictive value (PPV), and negative predictive value (NPV) are also presented. P < 0.05 was considered statistically significant. All analyses were performed in R 4.1.1.

**Results**

**Basic characteristics**

Of 725 subjects, 401 answered the survey (74% women) and gave informal consent to data collection from their medical records, Figure 1. The mean age of responders was 44 years and 73% were healthcare workers. The characteristics of the population are described in Table 1.

**Agreement in diagnoses between medical records and self-reported data**

The agreement between self-reported data and medical records is presented in Figure 2 and Table 2. High levels of total agreement were observed (>78%) for all diagnoses. The highest level of kappa was found for diabetes mellitus. Moderate level of kappa between the data sources was found for hypertension, hypo-/hyperthyroidism, and cancer. The sensitivity for hypertension was 56% and the PPV was 74%, as significantly more people had this condition in the survey than in medical records. Lung disease, stroke, and other heart disease had a high fair kappa level between self-reported data and medical records. In depression/anxiety, kappa was 0.36 corresponding to a fair level of agreement. These conditions also had the lowest total agreement and lowest NPV among the studied comorbidities. The prevalence of hypertension, hypo/hyperthyroidism, and depression/anxiety was significantly higher in the survey compared to medical records. In contrary, the prevalence

**Table 1. Basic characteristics of the population, n = 401.**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age group</strong></td>
<td></td>
</tr>
<tr>
<td>18–40 years</td>
<td>170 (42)</td>
</tr>
<tr>
<td>&gt;40 years</td>
<td>231 (68)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>298 (74)</td>
</tr>
<tr>
<td>Male</td>
<td>103 (26)</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
</tr>
<tr>
<td>Married/living with partner</td>
<td>273 (68)</td>
</tr>
<tr>
<td>Single/divorced/widow(er)</td>
<td>127 (32)</td>
</tr>
<tr>
<td><strong>Country of birth</strong></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>319 (80)</td>
</tr>
<tr>
<td>Country other than Sweden</td>
<td>81 (20)</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
</tr>
<tr>
<td>No university degree</td>
<td>178 (44)</td>
</tr>
<tr>
<td>University degree</td>
<td>221 (55)</td>
</tr>
<tr>
<td><strong>Occupation</strong></td>
<td></td>
</tr>
<tr>
<td>Healthcare worker</td>
<td>293 (73)</td>
</tr>
<tr>
<td>Other occupation</td>
<td>108 (27)</td>
</tr>
</tbody>
</table>

Missing: education n=2, country of birth n=1, marital status n=1.
of heart disease and diabetes mellitus was higher in medical records in comparison to self-reported information.

**Agreement in symptoms between medical records and self-reported data**

The agreement between the survey data and the medical records regarding symptoms of COVID-19 ranged between fair and poor, Figure 2 and Table 2. The highest (fair) concordance between the data sources was found for fever (kappa 0.36) and dyspnea (kappa 0.31). A low fair agreement was shown for muscle/joint pain, gastrointestinal symptoms, and headache. The poorest agreement was found for fatigue, chest pain, and impaired taste and smell (kappa range 0.05–0.18). Fatigue was more than 3.5 times more often reported in the surveys than in the medical records, its NPV was low and its kappa was very low: 0.05. The same pattern of significant underreporting in medical records was seen for impaired smell and taste and chest pain, with
Table 2. The number and percent of subjects who indicated these symptoms and diagnoses in the questionnaire survey and the medical records. Agreement, sensitivity, and specificity in diagnoses and symptoms at COVID-19 onset between survey data and medical records in the study population.

<table>
<thead>
<tr>
<th>Survey N, %</th>
<th>Medical records N, %</th>
<th>P-value</th>
<th>Agreement</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td>50 (12)</td>
<td>38 (9)</td>
<td>0.03</td>
<td>92.0</td>
<td>0.56</td>
<td>0.97</td>
<td>0.74</td>
</tr>
<tr>
<td>Heart disease</td>
<td>14 (3)</td>
<td>24 (6)</td>
<td>0.03</td>
<td>94.5</td>
<td>0.57</td>
<td>0.96</td>
<td>0.33</td>
</tr>
<tr>
<td>Hypo/hyperthyroidism</td>
<td>33 (8)</td>
<td>21 (5)</td>
<td>0.01</td>
<td>94.5</td>
<td>0.48</td>
<td>0.99</td>
<td>0.76</td>
</tr>
<tr>
<td>Diabetes</td>
<td>14 (3)</td>
<td>22 (5)</td>
<td>0.02</td>
<td>97.0</td>
<td>0.86</td>
<td>0.97</td>
<td>0.58</td>
</tr>
<tr>
<td>Lung disease</td>
<td>45 (11)</td>
<td>57 (14)</td>
<td>0.10</td>
<td>86.5</td>
<td>0.52</td>
<td>0.98</td>
<td>0.42</td>
</tr>
<tr>
<td>Cancer</td>
<td>21 (5)</td>
<td>19 (5)</td>
<td>0.64</td>
<td>95.5</td>
<td>0.52</td>
<td>0.98</td>
<td>0.58</td>
</tr>
<tr>
<td>Immunosuppressive treatment</td>
<td>16 (4)</td>
<td>11 (3)</td>
<td>0.25</td>
<td>95.3</td>
<td>0.25</td>
<td>0.98</td>
<td>0.36</td>
</tr>
<tr>
<td>Stroke</td>
<td>3 (1)</td>
<td>7 (2)</td>
<td>0.10</td>
<td>98.5</td>
<td>0.67</td>
<td>0.99</td>
<td>0.29</td>
</tr>
<tr>
<td>Depression/ anxiety</td>
<td>100 (25)</td>
<td>74 (18)</td>
<td>0.01</td>
<td>78.1</td>
<td>0.43</td>
<td>0.9</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Note. The P-values show differences between the survey and the medical records. PPV-positive predictive value, NPV-negative predictive value.

Low NPV and a kappa value indicating poor agreement. Other symptoms such as cough, sore throat, and nasal congestion were significantly underestimated in the survey data compared with in medical records.

Participants answering on paper had a higher prevalence of dyspnea, otherwise only slight differences were found between those answering by e-mail or on paper, Supplementary table S1.

Social health determinants related to agreement in diagnoses and symptoms between medical records and self-reported data

Variation in agreement were observed for different social health determinants. The study participants ≤40 years had significantly higher agreement for depression/anxiety (kappa with CI: 0.44 (0.3–0.59)) compared to the group >40 years old (kappa with CI: 0.25 (0.1–0.4), p = 0.07). Then, the only significant difference in the agreement between the data sources regarding symptoms was seen for fatigue, where the group ≤40 years old had higher kappa value (kappa with CI: 0.1 (0.02–0.19)) than the group >40 years old (kappa with CI: 0.02 (−0.04–0.08), p = 0.01). The total agreement for fatigue was higher in the younger group (45.3%) compared to older (33.3%), Supplementary Table S2.

Men had significantly better agreement between self-reported data and medical records for stroke (kappa 1.0) and diabetes mellitus (kappa 1.0) than females (kappa with CI: 0.28–0.76 for diabetes mellitus), Supplementary Table S2. Males also tended to have slightly (not significantly) higher agreement in several other somatic diagnoses including hypertension, heart disease, and lung disease. However, females were more prone to have slightly (not significantly) higher kappa values in depression and anxiety and hypo/hyperthyroidism, as well as most symptoms at COVID-19 onset, Supplementary table S1.

There were no significant differences in the agreement between self-reporting and medical records in educational levels, Supplementary table S2. The subjects without a university degree tended to have slightly higher kappa values for the majority of the diagnoses, whereas those with a university degree tended to have better agreement in symptoms.

A significantly higher agreement between self-reported data and medical records in hypertension was seen in the responders born in Sweden (kappa with CI: 0.66 (0.53–0.79)) compared with those born in another country (kappa with CI: 0.34 (0.04–0.65), p = 0.04), Supplementary table S2. Among the symptoms, fatigue showed significantly higher kappa values and higher sensitivity in people born in Sweden (kappa with CI: 0.08 (0.02–0.13); sensitivity 23%) compared with the other subgroup (kappa with CI: 0.06 (−0.01–0.12); p = <0.001; sensitivity: 12%).

Among the diagnoses, lung diseases showed significantly better kappa values and better sensitivity between the data sources in married subjects or those living with partner (kappa with CI: 0.49 (0.34–0.64),
sensitivity 80%) compared with people living alone (kappa with CI: 0.17 (−0.05–0.40), \( p = 0.02 \); sensitivity 31%), Supplementary table S2. Among the symptoms, this was true of fatigue (kappa with CI: 0.09 (0.03–0.22) and sensitivity 22% for living with partner; kappa with CI: −0.03 (−0.11–0.06), \( p = 0.02 \); and sensitivity 17% for living alone).

Healthcare workers had significantly better agreement between self-reported data and medical records in several subjective symptoms at COVID-19 onset, including fatigue, headache, and muscle and joint pain compared with those in other occupations, Supplementary table S2. The agreement between the data sources was significantly higher for depression/anxiety, where healthcare workers had fair kappa value (kappa with CI: 0.44 (0.32–0.56)), whereas the other group had poor kappa (kappa with CI: 0.14 (−0.04–0.32), \( p = 0.01 \)).

**Discussion**

In this study, we evaluated agreement between self-reported data and medical records from primary and hospital care in a predominantly healthy and working non-hospitalized adult population after COVID-19 infection. This study was performed not to determine which data sources is 'more accurate,' but rather to see how closely they agree or disagree for common diagnoses and symptoms. We showed substantial or moderate agreement between the two data sources for diagnoses such as diabetes mellitus, hypertension, hypo-/hyperthyroidism, and cancer. Other diagnoses showed fair agreement, though depression/anxiety and immunosuppressive treatment had the lowest concordance. We also found an overall fair to poor concordance between symptoms reported in the survey and those reported in the medical records.

For symptoms, the highest agreement between the two data sources was found for fever and dyspnea, whereas the lowest was seen for fatigue. In addition, we demonstrated that the level of agreement between self-reported data and medical records in diagnoses and symptoms was modified by different sociodemographic and lifestyle factors. We observed the same agreement trends for both diagnoses and symptoms between sexes and similar agreement trends for symptoms between occupational groups. For example, males had higher agreement between the data sources in somatic conditions than females and healthcare workers had significantly better kappa values in several symptoms of a subjective character compared with other occupational groups. For other factors – including age, marital status, country of birth, and educational level – we found no specific directions of agreement.

Like previous studies, our study showed that well-known and prevalent diseases (such as diabetes mellitus and hypertension) had higher agreement, according to both kappa values and PPV, between self-reported data and medical records compared with diagnoses that are hard to characterize (such as heart disease) [17–20]. The discrepancy between the data sources was large, despite patients having access to their personal medical records in Sweden.

Although we found substantial agreement for diabetes mellitus, the prevalence rate of the disease varied between the two data sources and was significantly higher in medical records than in the survey. The explanation might be that patients tend to downplay the severity of their glucose intolerance even when their medical records indicate diabetes [2]. Also, heart disease was reported significantly more often in medical records than in self-reports in our study. This might be due to patients with heart disease often having few or no symptoms if they are adequately treated [21]. On the other hand, cardiac conditions share symptoms with pulmonary and kidney diseases and may be unrecognized in many patients [21].

Furthermore, we showed a substantially higher prevalence of hypertension in the survey than in medical records. The explanation might be that general practitioners have higher cut-off value for the diagnosis of hypertension in younger adults [22]. Another reason may be that some respondents were on antihypertensive medication despite the diagnosis not being clinically confirmed [22].

In line with a study on other respiratory infections, we found that symptoms like fever and dyspnea had better agreement between questionnaires and medical records than symptoms like fatigue [23]. Like us, Barbara et al. found that the overall agreement on symptoms between the two data sources was low [23]. The reason for rather low agreement in symptoms might be that healthcare personnel focus on recording symptoms that have a high diagnostic value and are more specific [24]. Patients generally report symptoms that have an impact on their functioning and daily life [24]. This was also true in our study where the frequency of subjective symptoms such as fatigue, loss of taste, and chest pain were significantly higher in the survey compared to the medical records. Information about subjective symptoms is important for the identifying of patients that need to be followed-up. Our results indicate that healthcare personal often miss this kind of information. Another explanation of the low agreement in symptoms might be that there was higher NPV than PPV for almost all symptoms and this might lower kappa values (whereas higher positive agreement would increase kappa values).
We observed that fever was significantly more prevalent in the survey than in the medical records. The explanation might be that data in the medical records documented one phase of infection and treatment, whereas the patients took the whole infection period into account. Furthermore, some patients interpreted fever as a body temperature higher than usual, not necessary at least 38°C [25].

When investigated whether sociodemographic factors influenced the agreement, we found significant higher agreement between the data sources for depression/anxiety in the younger group being up to 40 years old compared to the older one. Further, we found no differences between age groups in agreement between self-reported data and medical records regarding somatic comorbidities. It is difficult to compare our results with the literature studies since different studies use various age groups. For example, one study demonstrated that the older age group (60–69 years old) had more concordance reporting depression than the younger one (40–49 years old) [8]. Then, the other study showed that older patients (>65 years old) had a worse agreement between self-reported diagnosis and those in medical records than younger patients [2].

Then, we showed that males were more likely to have better agreement between the data sources in all cardio-/neurovascular conditions than females. The sex difference was particularly significant for stroke and diabetes mellitus. Our results are partially in line with those of Okura et al. indicating that females have better agreement for hypertension while males have a better agreement for diabetes mellitus and heart failure [2].

Like the previous study, we found that females were half as likely as males to report diabetes even when present in medical records [22]. In other studies, we have observed that the agreement and PPV for stroke is higher in males than in females, though this might relate to the lower prevalence of stroke in females [26]. Similar to the previous study, we found no significant gender difference in the agreement of two data sources regarding depression and anxiety [27].

Some previous studies have reported that educational level does not have a significant impact on the concordance of any diagnoses or symptoms [28,29]. The results in our study were in line with this. We found that healthcare workers had significantly better agreement in self-reported data and medical records in symptoms like fatigue, muscle and joint pain, and headache than other occupational groups. This is as expected, as the knowledge on diseases is probably higher among healthcare workers than among those in other occupations.

We observed that healthcare workers had significantly better agreement between the data sources regarding depression/anxiety. Females tended to have better agreement than males for these conditions. Overall agreement for depression was at a fair level, with 48% sensitivity, though lower than that for several somatic diagnoses. The previous study showed that in 40% of subjects, self-perception of depression and anxiety differed from that of doctors [30]. Underreporting may be linked to stigmatization [30].

To our knowledge, no previous studies have evaluated the agreement between questionnaire reports and medical records regarding symptoms at COVID-19 onset. This study on the agreement between data sources is important because adequate information on the comorbidities and symptoms at COVID-19 can help us identify the subject at risk for more severe infectious outcomes. In addition, few previous studies investigated weather specific social determinants of health influence on this agreement. The other strength of this study is that it was based on information on diagnoses from both primary and secondary care, as using only register-based hospital records might lead to missed conditions of hypertension and diabetes, which are usually managed in primary care in Sweden [31].

A limitation of our study was that the study population was limited to subjects collected at one COVID-19 testing outpatient clinic. The response frequency in the study survey was 59%, which may create bias. The self-reported symptoms at onset were collected 12 months after infection, increasing the risk of recall bias. On the other hand, most studies based on self-reported information ask participants to think back over several weeks and months [32]. Then, some of the symptoms at COVID-19 onset were tightly associated with COVID-19 at the time of the survey that may recall additional bias in reporting. The data collected from medical records were based on diagnostic codes from the last 11 years and only from Region Uppsala. Some patients, while living in Uppsala, might have been diagnosed in another place [6]. Another limitation was that since we based on the ICD-10 codes from the medical records, we cannot verify whether the study subjects fulfilled the gold standard criteria at the diagnosis. In a previous study, the agreement of patient data of individuals transferred between primary and secondary care was low [33]. However, this is probably not a major problem in the present study, as in the Uppsala Region, all levels of care use the same patient record system. All study participants were rather young and working at the time of infection and the majority were healthcare workers, hampering generalization of the outcomes. In addition, the study population was generally healthy; therefore, the prevalence rates of some diseases were rather low. The rather small power and the partially low prevalence
of some variables as well as the exploratory nature and numerous comparisons may lead to bias in the kappa values. Therefore, the interpretation of the level of kappa values and p-values indicating differences between subgroups must be done with caution [34]. Finally, since neither data source can be considered a gold standard, it is not possible to know whether it was the patient (overreporting) or the clinician (under-recording) that was responsible for the discrepancy between self-reported data and the medical record. This study was conducted to describe how self-reported information and medical records relationships differ in a specific cohort setting rather than to identify all biases that might affect the agreements of selected symptoms and diseases.

Our study showed that more well-known diagnoses and specific symptoms had generally higher overall agreement between the two data sources than lesser-known diagnoses and less specific symptoms. In addition, we found that different sociodemographic factors, in particularly sex and occupation, had an impact on the concordance of several diagnoses and/or symptoms. This study illustrates that it is important to consider type of data used in the epidemiological as different information sources differ with quality and accuracy.

**Acknowledgments**

Karin Ekberg, medical doctor at the Department of Rehabilitation Medicine, Uppsala University Hospital, is also acknowledged for help with data transfer from the paper survey to the database. The article language was revised by professional scientific editing service: Meena Strömqvist Strömqvist, Legal & Language Services AB, www.stromqvist.org.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

We acknowledge Åke Wiberg Stiftelsen (stipendium M22-0119 and M21-0080), Lars Hiertas Minne Stiftelsen (stipendium F02021-0144 and F02022-0098) and Sven och Dagmar Saléns Stiftelse (stipendium 2022) and Tore Nilsons Stiftelsen (stipendium 2021-00935) for support.

**Data availability statement**

Data will be available on demand.

**References**


