Validity and underrecording of diagnosis of COPD in the Danish National Patient Registry

Reimar W. Thomsen\textsuperscript{a,*}, Peter Lange\textsuperscript{b}, Birthe Hellquist\textsuperscript{c}, Ejvind Frausing\textsuperscript{b}, Paul D. Bartels\textsuperscript{d}, Birgitte R. Krog\textsuperscript{d}, Anne-Marie S. Hansen\textsuperscript{d}, Daniel Buck\textsuperscript{a}, Anette E. Bunk\textsuperscript{d}

\textsuperscript{a}Department of Clinical Epidemiology, Institute of Clinical Medicine, Aarhus University Hospital, Aarhus, Denmark
\textsuperscript{b}Department of Cardiology and Respiratory Medicine, Hvidovre University Hospital, Copenhagen, Denmark
\textsuperscript{c}Department of Respiratory Diseases & Allergy, Aarhus University Hospital, Aarhus, Denmark
\textsuperscript{d}The Danish National Indicator Project (DNIP) Secretariat, Regionshuset, Aarhus, Denmark

Received 3 August 2010; accepted 21 January 2011
Available online 12 February 2011

Summary

Introduction: We examined the positive predictive value of diagnoses of acute exacerbation of chronic obstructive pulmonary disease (COPD) in the Danish National Patient Registry. We also examined the negative predictive value of acute pneumonia or respiratory failure discharge diagnoses for absence of underlying COPD.

Methods: We identified all patients aged 30 years or older with acute hospital admission in Denmark from January 1st to December 31st 2008. Physicians at 34 Danish hospitals retrieved and reviewed medical records for 1581 patients with a discharge diagnosis of COPD, and for 1546 patients with a discharge diagnosis of either pneumonia or respiratory failure but no COPD diagnosis. Presence of COPD was assessed based on medical history, clinical symptoms and findings, and spirometry results.

Results: The overall positive predictive value for COPD was 92% (95% confidence interval [CI] = 91–93%). Among patients coded with pneumonia or respiratory failure but not COPD, 19% (95% CI = 17–21%) had COPD, corresponding to a negative predictive value for COPD of 81% (95% CI = 79–83%).

Conclusions: The positive predictive value of acute COPD discharge diagnoses in the Danish National Patient Registry is high. At the same time, there is a substantial underrecording of COPD during hospitalizations with other acute respiratory disorders like pneumonia and respiratory failure.

© 2011 Elsevier Ltd. All rights reserved.

* Corresponding author. Department of Clinical Epidemiology, Aarhus University Hospital, Sdr. Skovvej 15, DK-9000 Aalborg, Denmark. Tel.: +45 9932 6904; fax: +45 9932 6914.
E-mail address: r.thomsen@rm.dk (R.W. Thomsen).

0954-6111/S - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.rmed.2011.01.012
Introduction

COPD is a major cause of morbidity and the fourth leading cause of death in the world. In Denmark with a population of 5.4 million, approximately 10% of persons aged 45 or older have COPD, and the disease is the cause of approximately 25,000 hospitalizations every year.

Use of administrative health care data could be a valuable and cost-efficient method of collecting data on COPD, both for health care quality monitoring and for epidemiological research. Administrative registries have the advantage of readily available data, and high completeness of registration of patients in the target population. However, the quality of routinely collected data for administrative purposes may be questionable for use in research and in quality measurements, and validation studies are a prerequisite for using registry-based data for these purposes. Although one Danish study has suggested that COPD is often under-diagnosed in patients admitted to hospital with acute respiratory conditions, in general there is a lack of data on the validity and completeness of COPD diagnosis in acute patient admissions. In the present nationwide prospective validation study, we examined the positive predictive value (PPV) of acute COPD discharge diagnoses for presence of underlying COPD in the Danish National Patient Registry (DNPR). We also examined the negative predictive value (NPV) of acute pneumonia or respiratory failure discharge diagnoses for absence of COPD.

Methods

This prospective nationwide validation study was conducted as part of the Danish National Indicator Project (DNIP). The DNIP started in 2000 and is a nationwide initiative to document, monitor and improve the quality of treatment and care provided by the Danish health care system. There is mandatory participation for all Danish hospitals in DNIP. COPD is one of currently 10 specific diseases monitored in the DNIP by clinical indicators and quality standards (see: http://www.nip.dk/about-the-danish-national-indicator-project). COPD was included in 2008 as the first monitored disease in the DNIP for which patient data are collected exclusively through administrative diagnosis codes and codes for medical procedures in the Danish DNPR, rather than through primary data collection into a separate quality database. Therefore, it was decided to validate a large nationwide sample of COPD codes in the DNPR.

The Danish National Patient Registry (DNPR)

The DNPR covers all individuals within the entire Danish population of 5.4 million who have been in contact with any Danish hospital. The Registry has tracked 99.4% of discharges from all Danish acute-care, non-psychiatric hospitals since 1977 and all hospital outpatient and emergency department visits since 1995. Records include patients’ unique civil registration numbers, admitting hospital and department, dates of admission and discharge, and up to 20 discharge diagnoses coded by physicians in charge of the patient admission. Diagnosis coding has been classified according to the International Classification of Diseases, 10th revision (ICD-10) since 1994.

Identification of patients coded with COPD, pneumonia, and respiratory failure in the DNPR

The present study covered all 34 hospitals in Denmark who reported to care for patients with acute COPD exacerbations in the beginning of 2008. Eligible were all patients who had a discharge date in the DNPR during the 12-month period from January 1st through December 31st, 2008.

For the study of the PPV of COPD diagnosis, we retrieved all acute hospitalization episodes with an ICD-10 hospital discharge diagnosis code of COPD (J44) among patients aged 30 years or older (n = 19,672 patients with COPD) (group 1). We included COPD diagnoses that were coded either as primary (first-listed) diagnosis or as a secondary diagnosis combined with a primary diagnosis code for acute respiratory failure (J96) or pneumonia (J13–J18) during the same hospitalization.

To obtain an estimate of the NPV of absence of any COPD diagnosis among patients hospitalized with related acute respiratory disorders, we retrieved all acute hospitalization episodes with a primary ICD-10 hospital discharge diagnosis code for either acute respiratory failure (J96) or pneumonia (J13–J18) without presence of any diagnosis code of COPD (J44) during the same hospitalization episode (n = 1713 patients with respiratory failure and n = 16,944 patients with pneumonia) (group 2).

Hospital record review

From the DNPR, we selected for each involved hospital a random sample of up to 60 acute COPD patient admissions from group 1, and up to 60 acute non-COPD patient admissions from group 2 (up to 15 admissions per quarter of the year 2008, if available, were randomly selected). In this manner, the total validation cohort included approximately 10% of all patients acutely admitted with COPD diagnosis during the study period (n = 1988), and a random sample of approximately 10% of all patients acutely admitted with a diagnosis of respiratory failure or pneumonia without any diagnosis of COPD (n = 1986). We then sent standardized registration forms to all involved hospitals, containing the up to 120 selected patients’ civil registration numbers, diagnosis codes, and discharge dates, and asked for a clinical audit of the patients’ hospital records. Hospital records were retrieved and reviewed by senior physicians at the involved pulmonary and general medical departments. The review was based on all the available information from actual and previous hospitalizations in the medical records, including spirometry results in stable condition, case notes, radiology reports, and results from laboratory tests. A diagnosis of COPD was made by the reviewing physician based on clinical judgment, i.e. when the patient’s medical history and typical clinical symptoms and signs were in agreement with COPD, supported by findings from spirometry within the last three years if available in the hospital record (proportion of the forced vital capacity exhaled in the first second (FEV1/FVC ratio) < 70%).
The data validity of COPD diagnosis, expressed as the positive predictive value (PPV) of the COPD discharge diagnosis, was calculated as a proportion (i.e., the numerator containing the number of patients with confirmed COPD diagnoses after review of medical records, and the denominator containing the total number of patients registered with COPD diagnosis). The data validity of absence of COPD, expressed as the negative predictive value (NPV) for presence of COPD in cases coded with pneumonia or respiratory failure (without mentioning COPD), was also calculated as a proportion (i.e., the numerator containing the number of patients without COPD diagnoses after review of medical records, and the denominator containing the total number of patients registered with pneumonia or respiratory failure (without COPD diagnoses)). Furthermore, data were stratified based on gender, administrative region in Denmark, and type of admitting hospital department (specialized pulmonary medicine versus general medicine; only 2% of patients had been admitted to non-medical departments). PPVs were also calculated separately for COPD as primary (first-listed) and secondary diagnosis, and restricted to COPD confirmed by spirometry with FEV1/FVC ratio < 70% within the last 3 years according to available data in hospital records. NPVs for absence of COPD were calculated separately for diagnoses of pneumonia and respiratory failure. To estimate 95% confidence intervals (CI) and compare proportions, an approximation to the binomial distribution was assumed.

Results

Descriptive data

Table 1 shows characteristics of the 19,672 patients acutely admitted with COPD diagnosis in Denmark and recorded in the DNPR database in 2008 who were eligible for sampling for this validation study, as well as the sampled validation group of 1986 COPD patients. There were slightly more women than men, and half of the cohort was older than 74 years. More than half of the patients had previously hospital-diagnosed comorbidity other than COPD as evidenced from the Charlson comorbidity index7,8 prior to admission. The cumulative 30-day mortality after acute COPD admission was 9–10%. Table 1 also shows characteristics of the sampled validation group of 1986 patients admitted with pneumonia or respiratory failure without COPD diagnosis. Compared with patients diagnosed with COPD, these patients were of comparable age (median 75 years), slightly more were men (52%), and less had comorbidity recorded.

Validation of COPD

Among the nationwide subgroup of 1986 patients with a COPD discharge diagnosis randomly sampled for the validation study, physicians at the involved hospitals retrieved and reviewed hospital records and returned the filled data forms for 1581 hospitalization episodes (80%). The primary reasons for not returning filled data forms were logistic difficulties and/or lack of human resources at a number of hospitals, who therefore did not participate in the validation study (a nationwide nurses’ strike took place in Denmark between April 16 and June 15, 2008). Among the 1581 validated patient episodes, the overall PPV for COPD was 92% (95% confidence interval [CI] 91–93%) (Table 2). Stratification by gender revealed a PPV of 92% (90–94%) in both men and women. When we stratified the data by the five administrative regions of Denmark, the PPVs for COPD were uniformly high, ranging from 87% to 94%. The PPV for COPD among episodes with a primary (first-listed) COPD diagnosis was 93% (95% CI 92–95%), whereas the PPV for COPD in episodes with a secondary COPD diagnosis combined with a primary diagnosis of acute respiratory failure or pneumonia was slightly lower at 87% (95% CI 84–91%). When we stratified the data on the type of admitting hospital department, the PPV for COPD was 95% (92–97%) in the specialized pulmonary medical departments and 92% (90–93%) in the general medical departments. The PPV for COPD reported to be confirmed by spirometry (FEV1/FVC ratio < 70%) based on availability of data in hospital records within the last 3 years was substantially lower at 61% (95% CI 59–64%) overall and 75% (70–80%) in the specialized pulmonary medical departments.

Validation of absence of COPD

Among the nationwide sampled subgroup of 1986 patients who had been discharged with a diagnosis of respiratory failure or pneumonia without any mention of COPD, physicians retrieved and reviewed hospital record and returned filled data forms for 1546 hospitalization episodes (78%). Reasons for non-participation of hospitals were the same as for the patients with COPD diagnoses mentioned above. In this subgroup, 19% (95% CI 17–21%) had COPD, corresponding to a PPV of 98% (95% CI 97–99%) (Table 3). Stratification on gender revealed a PPV of 83% (95% CI 80–85%) in men and 78% (95% CI 75–81%) in women. The PPVs varied from 87% (95% CI 81–92%) in the North Denmark Region to 72% (95% CI 68–77%) in the Sealand Region. Among the patients with a diagnosis of pneumonia without COPD code, 18% (95% CI 16–20%) of the patients had COPD (NPV = 82% (95% CI 80–84%)), whereas among the patients coded with acute respiratory failure without COPD code, an even higher percentage of 41% (95% CI 32–51%) had COPD (NPV = 59% (95% CI 49–68%)). When we stratified the data based on the type of hospital department, the NPVs tended to be slightly lower in the specialized pulmonary medical departments than in the general medical departments. The NPV for COPD confirmed with FEV1/FVC ratio < 70% within the last 3 years was considerably higher at 91% (95% CI 89–92%) overall.

Discussion

The present study is the first nationwide study of the validity of the population-based DNPR with regard to COPD diagnoses. We found that the positive predictive value of acute COPD diagnoses in the registry is high. At the same time, there seems to be a substantial underrecording of COPD among patients hospitalized with pneumonia and in particular among patients hospitalized with respiratory failure.
The strengths of our study included the access to medical records with detailed clinical data for two entire nationwide validation cohorts. The large sample size resulted in an improved statistical precision compared to previous studies, and made it possible to perform subgroup analysis. We only received filled data forms for approximately 80% of the hospitalization episodes, yet this was largely explained by non-participation of a few specific hospitals due to logistic difficulties, and we believe that the patients and the hospitals included in this study are representative of the Danish hospitals and the DNPR as a whole. Because the Danish National Health Service provides free universal tax-supported health care for all inhabitants, disadvantaged population groups are not underrepresented in our study, and there are no financial incentives for hospitals not to provide a similar diagnostic workup for all patients.9

This study also has limitations. Physicians used medical records as the gold standard when assessing the predictive value of the diagnoses recorded in the DNPR. Although this is the conventional approach used in most validation studies, medical records are not perfect, and it can be difficult to assess the quality of the information they contain.9,14 The medical records were reviewed by a large number of different physicians, with the possibility of inter-observer variation in the evaluations. Furthermore, the reviewers were not blinded to the diagnosis codes, since presence or absence of the COPD register diagnosis was by definition given for all the examined patients in the two groups. Thus, knowledge of the COPD diagnosis code could have influenced the results of the physicians’ assessment.

COPD is a complex disease characterized by chronic, not fully reversible airflow limitation, a range of pathological changes in the lung, significant extrapulmonary effects, and frequent comorbidities.1 The characteristic symptoms of COPD are chronic and progressive dyspnea, cough, and sputum production. According to the GOLD criteria, an accurate COPD diagnosis requires spirometry.1,2 However, in real-world clinical practice in acutely hospitalized COPD patients, an updated spirometry result is often not available in the hospital record. Thus, in recent audits of hospitalizations for COPD in Norway/Sweden and the UK, respectively, 19% and 47% of COPD patients had no accessible spirometry data before or during admission.14,15 Therefore, in our study we pragmatically defined COPD as presence of a medical history of COPD together with clinical

### Table 1
Characteristics of all patients acutely admitted with COPD in Denmark recorded in the DNIP database 2008, and of the two groups sampled for the validation study: patients admitted with COPD, and patients hospitalized with pneumonia or respiratory failure without COPD diagnosis.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Patients with COPD in the DNIP database (n = 19,672)</th>
<th>Patients with COPD in the validation sample (n = 1988)</th>
<th>Patients without COPD in the validation sample (n = 1986)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age (quartiles) in years</td>
<td>74 (66–80)</td>
<td>74 (66–81)</td>
<td>75 (62–83)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>10,901 (55%)</td>
<td>1082 (54%)</td>
<td>944 (48%)</td>
</tr>
<tr>
<td>Male</td>
<td>8771 (45%)</td>
<td>906 (46%)</td>
<td>1042 (52%)</td>
</tr>
<tr>
<td>Charlson comorbidity index score (COPD excluded)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>7603 (39%)</td>
<td>861 (43%)</td>
<td>1185 (60%)</td>
</tr>
<tr>
<td>1</td>
<td>5195 (26%)</td>
<td>476 (24%)</td>
<td>261 (13%)</td>
</tr>
<tr>
<td>2</td>
<td>3012 (15%)</td>
<td>270 (14%)</td>
<td>261 (13%)</td>
</tr>
<tr>
<td>3+</td>
<td>3862 (20%)</td>
<td>381 (19%)</td>
<td>279 (14%)</td>
</tr>
<tr>
<td>Assisted ventilation during hospitalization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-invasive ventilation</td>
<td>1535 (8%)</td>
<td>124 (6%)</td>
<td>a</td>
</tr>
<tr>
<td>Ventilator</td>
<td>540 (3%)</td>
<td>32 (2%)</td>
<td>a</td>
</tr>
<tr>
<td>Cumulative 30-day mortality</td>
<td>1841 (9%)</td>
<td>193 (10%)</td>
<td>a</td>
</tr>
</tbody>
</table>

a Data on ventilation and mortality not recorded for non-COPD patient sample.

### Table 2
Positive predictive values of COPD diagnoses for 1581 acutely admitted patients in the Danish National Patient Registry.

<table>
<thead>
<tr>
<th>Discharge diagnoses in the Danish National Patient Registry</th>
<th>n</th>
<th>COPD verified</th>
<th>PPV (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COPD overall</td>
<td>1581</td>
<td>1456</td>
<td>92 (91–93)</td>
</tr>
<tr>
<td>COPD, primary diagnosis</td>
<td>1223</td>
<td>1143</td>
<td>93 (92–95)</td>
</tr>
<tr>
<td>COPD, secondary diagnosis</td>
<td>358</td>
<td>313</td>
<td>87 (84–91)</td>
</tr>
<tr>
<td>COPD, men</td>
<td>737</td>
<td>680</td>
<td>92 (90–94)</td>
</tr>
<tr>
<td>COPD, women</td>
<td>844</td>
<td>776</td>
<td>92 (90–94)</td>
</tr>
<tr>
<td>COPD, specialized pulmonary medicine department</td>
<td>279</td>
<td>265</td>
<td>95 (92–97)</td>
</tr>
<tr>
<td>COPD, general medicine department</td>
<td>1292</td>
<td>1185</td>
<td>92 (90–93)</td>
</tr>
</tbody>
</table>
symptoms and signs of COPD, supported by presence of a spirometry with FEV$_1$/FVC ratio < 70% within the last three years, if available. This definition is open to some subjective interpretation of symptoms and signs by the reviewing physicians. It may be considered a limitation of our validation study that physicians were not asked for date and result of the most recent FEV$_1$/FVC measurement for each patient regardless of the measured value. Thus, we do not know the exact reasons behind the fact that 39% of the reviewed patients had no recent spirometry with FEV$_1$/FVC ratio < 70% recorded. Different scenarios are possible, i.e.; the patients had not been examined with spirometry at all, patients did have spirometry performed in another clinical setting with results unavailable for the reviewing physician, patients did have a spirometry performed in another clinical setting within more than 3 years ago, or patients did have a recent spirometry with FEV$_1$/FVC ratio > 70% suggesting another diagnosis than COPD. Based on clinical experience in our setting and on the written comments we received from many clinicians involved in the validation study, for example: "This patient has his lung function parameters controlled in another outpatient clinic" or "This patient is normally followed in another hospital for his COPD and was only here for intensive care", etc., we find it likely that simple unavailability of any updated spirometry results in the reviewed hospital record constitute the vast majority of cases where COPD was defined on the basis of clinical history.

In our literature search, we were unable to find validation studies similar to ours in order to compare our present findings with earlier results. Previous studies focusing on underdiagnosing and underreporting of COPD comprise either cohorts of the general population,$^{11,16}$ patients in general practice$^7$ or patients with heart disease.$^{18}$ In particular, there has been focus on COPD diagnostic problems among patients with heart failure,$^{19}$ but no studies on underreporting of COPD among patients with pneumonia or respiratory failure, although COPD is an important risk factor for both conditions.$^{20}$

It is reassuring that the diagnosis of COPD in acutely admitted patients has high validity when assessed through clinical criteria by experienced physicians. We observed higher validity of the COPD diagnosis given in specialized pulmonary than general medical departments, as we would expect in a specialist setting. At the same time and more unexpected, NPVs for COPD tended to be lower in the specialist setting, corresponding to more false negative COPD patients in pulmonary than general medical departments. This fact may have to do with a more frequent and correct use of pneumonia and acute respiratory failure as primary diagnostic codes in pulmonary medical departments, with potential simultaneous underregistration of COPD as secondary diagnosis during the same episode.

The implications of problems relating to the data quality of registry-based COPD discharge codes differ between different purposes and study designs.$^5,6$ Thus, in studies of the incidence of acute hospitalizations with COPD, use of codes J44 without further validation would only overestimate the COPD incidence slightly per se, since the codes are highly valid. In contrary, exclusive use of J44 codes may lead to substantial underestimation of acute COPD hospitalization incidences, since there seems to be substantial underreporting of COPD during hospitalizations with other acute respiratory disorders. Furthermore, some patients with true COPD may have been wrongly coded with codes that we did not examine, including J42 (unspecified chronic bronchitis) or J43 (emphysema), potentially leading to further underestimation of acute COPD hospitalization incidences.$^4$ In analytical epidemiological studies, the misclassification of COPD diagnoses would tend to bias the examined associations toward the null hypothesis, as long as the COPD sensitivity and specificity is not strongly associated with exposure. Importantly, in analytic database studies investigators would usually aim to screen the patients’ entire previous hospitalization history for COPD diagnosis codes, likely leading to increased completeness of COPD exposure as compared to assessment from a single hospitalization episode only. For health care quality measurement such as the DNIP, the rather similar COPD data quality in Danish regions suggests that quality of care indicator results can be validly compared between regions.

**Conclusion**

In conclusion, we found a high predictive value (>90%) of the diagnosis of COPD for acutely admitted patients when we used clinical COPD criteria with or without spirometry. At the same time, there is a substantial underreporting of COPD during hospitalizations with other acute respiratory disorders, like pneumonia or respiratory failure, which may lead to substantial underestimation of the incidence of acute admissions with COPD on the population level.

**Financial support**

None.
Conflict of interest statement

Peter Lange has received honoraria for speaking and consulting and/or financial support for attending meetings from AstraZeneca, Boehringer Ingelheim, GSK, Novartis, Norpharma, Nycomed, Pfizer, and UCB. No funding has been provided for this article. All other authors have no conflicts of interest to declare.

Acknowledgments

We sincerely thank the participating staff at all hospital departments in Denmark for collecting the data. We thank Frank Mehnert and Morten Madsen for contributions to the statistical analyses.

References


