

# Positive predictive value of first-time rheumatoid arthritis diagnoses and their serological subtypes in the Danish National Patient Registry

Asta Linauskas<sup>1,2</sup>Kim Overvad<sup>3</sup>Martin Berg Johansen<sup>4</sup>

Kristian

Stengaard-Pedersen<sup>1,5</sup>Annette de Thurah<sup>1,5</sup><sup>1</sup>Department of Rheumatology, Aarhus University Hospital, Aarhus, Denmark;<sup>2</sup>Department of Rheumatology, North Denmark Regional Hospital, Hjoerring, Denmark; <sup>3</sup>Section for Epidemiology, Department of Public Health, Aarhus University, Aarhus, Denmark; <sup>4</sup>Unit of Clinical Biostatistics, Aalborg University Hospital, Aalborg, Denmark; <sup>5</sup>Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

**Purposes:** To assess whether the positive predictive value (PPV) of first-time rheumatoid arthritis (RA) diagnosis registration in the Danish National Patient Registry increases when data are linked to the RA treatment codes and to assess the PPV of first-time RA diagnoses according to RA serological subtypes.

**Methods:** Participants from the Danish Diet, Cancer, and Health cohort with at least one RA diagnosis, registered at one of the Central Denmark Region hospitals in the Danish National Patient Registry during the period 1977–2016, were identified. Register-based RA diagnoses were verified by scrutinizing medical records against RA classification criteria or clinical case RA. PPVs for overall RA, seropositive RA, and other RA were calculated for two models: first-time RA diagnosis registration ever in the Danish National Patient Registry and first-time RA diagnosis registration ever where subsequently a prescription had been redeemed for a synthetic disease-modifying antirheumatic drug.

**Results:** Overall, 205 of 311 first-time register-based RA diagnoses were verified (PPV: 61.9%; 95% CI: 56.9–67.0). Regarding RA serological subtypes, 93 of 150 register-based seropositive RA (PPV: 62.0; 95% CI: 53.9–69.5) and 36 of 144 other RA (PPV: 25.0; 95% CI: 18.5–32.8) were confirmed. When register-based RA diagnosis codes were linked to RA treatment codes, the PPVs increased substantially: the PPV for overall RA was 87.7% (95% CI: 82.5–91.5), the PPV for seropositive RA was 80.2% (95% CI: 71.6–86.7), and the PPV for other RA was 41.1% (95% CI: 30.2–52.9).

**Conclusion:** The first-time RA diagnoses in the Danish National Patient Registry should be used with caution in epidemiology research. However, linking registry-based RA diagnoses to the subsequent RA treatment codes increases the probability of identifying true RA diagnoses, especially overall RA and seropositive RA.

**Keywords:** administrative database research, Danish National Prescription Registry, Anatomical Therapeutic Chemical codes, seropositive rheumatoid arthritis, Denmark

## Introduction

Rheumatoid arthritis (RA) is a systemic autoimmune inflammatory disease characterized by inflammation of synovial joints; it leads to irreversible joint damage, deformity, and functional impairment. Furthermore, RA patients are at higher risk of comorbidities and have a higher mortality rate than the general population.<sup>1,2</sup>

Large-scale registries have been widely used in research, contributing to knowledge regarding epidemiology, comorbidities, and mortality rates of RA patients.<sup>1–6</sup>

The Danish National Patient Registry is a Danish key health registry.<sup>7</sup> It was established for administrative purposes in 1977 as an extension of county-based

Correspondence: Asta Linauskas  
Department of Rheumatology, Aarhus University Hospital, Noerrebrogade 44, DK-8000 Aarhus C, Aarhus, Denmark  
Tel +45 9764 0990  
Fax +45 9764 0999  
Email asli@rn.dk

hospital registration systems. For many decades, the Danish National Patient Registry has been extensively used for epidemiological and clinical research. The registry contains comprehensive data regarding hospital contacts, including International Classification of Disease (ICD) codes.<sup>8</sup> However, the validity of both in- and outpatient diagnostic data is open to discussion. Pedersen et al<sup>9</sup> investigated the validity of 217 RA diagnoses registered in the Danish National Patient Registry during the period 1977–2001. They found that the positive predictive value (PPV) of RA diagnoses was 59%. Ibself et al<sup>10</sup> validated incident RA diagnoses recorded  $\geq 2$  times within 90 days in rheumatology departments during 2011. The PPV of RA diagnoses in their study was 79%. A study from the USA showed that the PPV of administrative registry-based RA diagnoses increased from 66% to 81% when data were linked to RA treatment codes.<sup>11</sup>

It is well known that seropositive and seronegative RA differ with respect to their phenotype, susceptibility genes, and the impact of smoking and obesity as risk factors.<sup>12–18</sup> Therefore, it is important to study these two serological subtypes separately.

The aims of our study were, first, to assess whether the PPV of first-time RA diagnosis registration in the Danish National Patient Registry increases when data are linked to the RA treatment codes and, second, to assess the PPV of first-time RA diagnoses according to serological subtypes.

## Methods

### Setting

The study explored the PPV of RA diagnoses in the Danish National Patient Registry during the period 1977–2016 among participants in the Danish Diet, Cancer, and Health cohort.<sup>19</sup>

### Data sources

#### The Danish Diet, Cancer, and Health cohort study

The cohort was recruited during the period 1993–1997. Overall, 80,996 men and 79,729 women were invited, and a total of 57,053 accepted the invitation. The eligibility criteria were the following: age 50–64 years, born in Denmark, living in the Copenhagen or Aarhus counties, and no cancer diagnosis registered in the Danish Cancer Registry.<sup>20</sup> A detailed description of the cohort has been published previously.<sup>19</sup>

All participants of the Danish Diet, Cancer, and Health cohort were linked to the Danish National Patient Registry<sup>7</sup> in February 2017 and subsequently linked to the Danish National Prescription Registry using the unique Civil Personal Registration number. This number is used in all

Danish registries, enabling data linkage at the individual level.

#### The Danish National Patient Registry

The registry contains data on all admissions to somatic hospitals since 1977 and on all visits to outpatient clinics since 1995,<sup>7</sup> including hospital information, dates of hospital admissions, ward types, referral and discharge diagnoses, dates of each attendance at outpatient clinics, type of clinics, and diagnoses recorded at each attendance. The diagnoses were classified in accordance with the Danish version of the ICD-8 until the end of 1993 and thereafter in accordance with the updated version, ICD-10.<sup>8</sup>

#### The Danish National Prescription Registry

The Danish National Prescription Registry is a redemption database that was established in 1994.<sup>21</sup> Every pharmacy in Denmark is equipped with an electronic accounting system to secure reimbursement from the National Health Service, which funds a variable proportion of the cost of prescribed medicine for every Danish citizen. Data are transferred to the Danish National Prescription Registry, which thus covers all reimbursed drugs at the level of the individual user.

The registry does not include information about drugs dispensed by hospital pharmacies directly to inpatients in relation to their hospital stays or to outpatients during their outpatient clinic attendances.

### Validation of registry-based RA diagnoses

#### Identification of RA diagnoses

We included participants from the Danish Diet, Cancer, and Health cohort with at least one RA diagnosis, registered at one of the Central Denmark Region<sup>22</sup> hospitals in the Danish National Patient Registry during the period 1977–2016.

The following ICD codes regarding RA diagnoses were retrieved from the Danish National Patient Registry: 712.39 (arthritis rheumatoides alia et non specificata), M05 (seropositive RA), and M06 (other RA) though not including M06.1 (adult-onset Still disease). Only RA diagnoses registered as the main or the secondary diagnoses were included in the analyses, and referral diagnoses were not taken into account.

#### Verification of RA diagnoses

RA diagnoses identified in the Danish National Patient Registry were verified by scrutinizing medical records as follows: first-time RA diagnosis registrations were retrieved through the Danish National Patient Registry as described previously; then, relevant medical records from public hospitals and out-

patient clinics were retrieved, and RA diagnoses were verified against the American College of Rheumatology (ACR) RA classification 1958 criteria,<sup>23</sup> the ACR 1987 criteria,<sup>24</sup> or the ACR/European League Against Rheumatism (ACR/EULAR) 2010 criteria.<sup>25</sup>

If no RA classification criteria were met, verification was undertaken against clinical case RA. Clinical case RA was defined as arthritis where an assessment of RA was based on a rheumatologist's expert opinion. If the RA diagnosis could not be confirmed, the likely alternative diagnosis was registered.

Register-based RA diagnoses with insufficient medical record information regarding satisfaction of the classification criteria or clinical case RA definition were categorized as nonverified RA diagnoses.

## Validation of registry-based RA diagnoses according to serological subtypes

### Identification of RA diagnoses according to serological subtypes

Considering that only the ICD-10 diagnosis codes distinguish between serological subtypes of RA,<sup>8</sup> verification of serological subtypes was exclusively carried out among first-time ICD-10 code registrations.

Diagnoses registered with one of the "M05" codes as the first registration in the Danish National Patient Registry were defined as registry-based seropositive RA, and diagnoses registered with one of "M06" codes were registered as other RA.

Only RA diagnoses registered as the main or the secondary diagnoses were included in the analyses, and referral diagnoses were not taken into account.

Patients with a previous ICD-8 RA code registration were not included in the analysis.

### Verification of RA diagnoses according to serological subtypes

Register-based RA diagnoses were verified against the RA classification criteria or clinical case definition by reviewing medical records, as described in the previous section. The criterion standard of RA seropositivity was defined as a positive blood test for IgM rheumatoid factor (RF) and/or anti-citrullinated protein antibody (ACPA) before, at or within 1 month after the first-time registration in the Danish National Patient Registry. The criterion standard of other RA was defined as a negative blood test for IgM RF and, if an ACPA test had been undertaken, a negative blood test

for ACPA before, at or within 1 month after the first-time registration in the Danish National Patient Registry.

Register-based RA diagnoses whose medical record information was insufficient in relation to satisfying the classification criteria, clinical case RA definition, or IgM RF/ACPA blood test results were categorized as nonverified RA diagnoses.

## Statistical analyses

### Verification of the registry-based RA diagnoses

Every first-time RA diagnosis in the Danish National Patient Registry was classified as a true positive RA diagnosis or a nonverified RA diagnosis using the RA classification criteria or clinical case definition based on medical records, as described earlier.

The PPV of registry-based RA diagnosis was estimated for two different models:

Model 1: first-time RA diagnosis registration ever in the Danish National Patient Registry;

Model 2: first-time RA diagnosis registration ever in the Danish National Patient Registry and where subsequently, at least once, a prescription had been redeemed for a synthetic disease-modifying antirheumatic drug (sDMARD). Information about prescribed drugs was available only from 1995. Therefore, subanalyses for RA diagnoses registered in the Danish National Patient Registry before and after 1995 were carried out using the second model. The following Anatomical Therapeutic Chemical (ATC) codes<sup>26</sup> from the Danish National Prescription Registry were retrieved: methotrexate (ATC: L01BA01, L04AX03), sulfasalazine (ATC: A07EC01), azathioprine (ATC: L04AX01), hydroxychloroquine (ATC: P01BA02), and leflunomide (ATC: L04AA13).

In Model 1, the PPV was determined as the fraction of true positive RA diagnoses among all first-time RA diagnoses recorded in the registry. In Model 2, the PPV was determined as the fraction of true positive RA among all first-time RA diagnoses recorded in the registry for cases, where subsequently at least once a prescription had been redeemed for a sDMARD. PPVs were presented with 95% CIs.

The PPV of RA diagnoses was subsequently explored according to patients' gender, age group at the time of diagnosis (<50, 50–59, 60–69, >70 years), type of department (rheumatology, internal medicine, or other), type of patient (inpatient, outpatient, or other), type of RA diagnosis (main or secondary), calendar periods, based on the development of RA classification criteria (1977–1987, 1988–2010, 2011–2016),<sup>23–25</sup> and ICD codes.<sup>8</sup>

## Verification of the registry-based RA diagnoses according to serological subtypes

Every first-time ICD-10 “M05” RA diagnosis in the Danish National Patient Registry was classified as a true positive seropositive RA or nonverified seropositive RA using the classification criteria or clinical case definition for RA and blood test results, as described earlier.

In the same manner, every first-time ICD-10 “M06” RA diagnosis in the registry was classified as a true positive other RA or nonverified other RA.

The PPVs of registry-based RA diagnoses, including serological subtypes, were estimated for two different models:

Model 1: first-time RA diagnosis registration ever in the Danish National Patient Registry with one of the ICD-10 RA codes;

Model 2: first-time RA diagnosis registration ever in the Danish National Patient Registry with one of ICD-10 RA codes and where subsequently, at least once, a prescription had been redeemed for a sDMARD.

For seropositive RA in Model 1, the PPV was calculated as the fraction of true positive seropositive RA diagnoses among all first-time seropositive RA diagnoses (ICD-10 “M05”) recorded in the Danish National Patient Registry. In Model 2, the PPV was ascertained as the fraction of true positive seropositive RA among all first-time seropositive RA diagnoses recorded in the registry and where subsequently, at least once, a prescription had been redeemed for a sDMARD.

The PPV of registry-based other RA diagnosis in Model 1 was ascertained as the fraction of true positive other RA diagnoses among all first-time other RA diagnoses (ICD-10 “M06”) recorded in the Danish National Patient Registry. In Model 2, the PPV was ascertained as the fraction of true positive other RA among all first-time other RA diagnoses recorded in the registry and, where subsequently, at least once, a prescription had been redeemed for a sDMARD.

Analyses were performed using Stata, version 14.2 (StataCorp LP, College Station, TX, USA).

## Results

A total of 331 participants in the Danish Diet, Cancer, and Health cohort were recorded as having RA at one of the Central Denmark Region hospitals in the period 1977–2016.

The characteristics of the registry-based RA diagnoses are presented in Table 1. As expected, the majority of persons recorded as having RA were women (70%). Most of the first-time RA diagnoses were recorded at a rheumatology or internal medicine department (78%) as the main diagnosis.

**Table 1** Characteristics of first-time rheumatoid arthritis diagnoses recorded in the Danish National Patient Registry among participants in the Danish Diet, Cancer, and Health cohort and derived from the Central Denmark Region hospitals

Characteristics	
Registered at least once with rheumatoid arthritis diagnosis during 1977–2016, n	331
Female, % (n)	70 (231)
Mean age, years (range)	65 (35–83)
Calendar periods, % (n)	
1977–1987	6 (19)
1988–2010	70 (232)
2011–2016	24 (80)
Department type, % (n)	
Rheumatology	58 (192)
Internal medicine	20 (65)
Surgery	16 (54)
Other	6 (20)
Patient type, % (n)	
Inpatient	40 (134)
Outpatient	57 (189)
Other	3 (8)
Diagnosis type, % (n)	
Main (A-diagnosis)	78 (257)
Secondary (B-diagnosis)	22 (74)
ICD-8 codes, % (n)	11 (37)
ICD-10 codes, % (n) <sup>a</sup>	89 (294)
M05, % (n)	45 (150)
M06, % (n)	44 (144)
ICD codes by patients' gender	
Men, n	100
ICD-8 codes, % (n)	12 (12)
ICD-10 M05 codes, % (n)	49 (49)
ICD-10 M06 codes, % (n)	39 (39)
Women, n	231
ICD-8 codes, % (n)	11 (25)
ICD-10 M05 codes, % (n)	44 (101)
ICD-10 M06 codes, % (n)	45 (105)
ICD-10 codes by patients age at the time of rheumatoid arthritis diagnosis registration	
M05, n	150
<50 years, % (n)	1 (2)
50–59 years, % (n)	22 (33)
60–69 years, % (n)	44 (65)
>70 years, % (n)	33 (50)
M06, n	144
<50 years, % (n)	1 (2)
50–59 years, % (n)	13 (19)
60–69 years, % (n)	46 (66)
>70 years, % (n)	40 (57)

**Note:** <sup>a</sup>ICD-10 code M05 – seropositive rheumatoid arthritis, ICD-10 code M06 – other rheumatoid arthritis.

**Abbreviation:** ICD, International Classification of Diseases.

## Verification of first-time RA diagnoses

For a total of 7% (23/331) of the diagnoses, no sufficient medical record information was available, and these diagnoses were therefore defined as nonverified RA.

Table 2 shows the PPVs of the first-time RA diagnoses registered in the Danish National Patient Registry. Using the Model 1 approach, RA diagnoses were confirmed for 62% (205/331) of the registry-based diagnoses (95% CI: 56.6–67.0). The majority of the confirmed diagnoses satisfied ACR or ACR/EULAR classification criteria (n=202), while others (n=3) met only the clinical case RA definition. ICD-8 coded diagnoses had the highest PPV (PPV: 78%; 95% CI: 61.5–89.2). Furthermore, a large proportion of the diagnoses was verified in the medical records if they were registered at a rheumatology department (PPV: 72%; 95% CI: 65.0–77.8), recorded as a main diagnosis (PPV: 68%; 95% CI: 62.1–73.5) or as a seropositive RA (PPV: 72%; 95% CI: 64.2–78.7).

When registry-based diagnosis codes were linked to the RA treatment codes in the Danish National Prescription Registry (Model 2), the PPV of RA diagnosis was

substantially higher (Table 2). Overall, 88% of RA diagnoses were confirmed (95% CI: 82.5–91.5). The PPV was higher for all categories compared with the PPV calculated using Model 1. The results were not different when stratified by periods (ie, before and after 1994 [establishment of the Danish National Prescription Registry], the PPV was 90% and 87%, respectively).

## Verification of first-time RA diagnoses according to serological subtypes

A total of 89% of the RA diagnoses (294/331) were registered with one of the ICD-10 RA codes, and these were almost equally distributed between “M05” (n=150) and “M06” (n=144) codes. Stratification by patient age at the diagnosis showed preponderance of seropositive RA diagnostic codes “M05” in younger patients (Table 1). Overall, for 8.7%

**Table 2** The positive predictive values of first-time overall rheumatoid arthritis diagnoses recorded in the Danish National Patient Registry in the period 1977–2016 among the participants in the Danish Diet, Cancer, and Health cohort and derived from the Central Denmark Region hospitals

Characteristics	Rheumatoid arthritis <sup>a</sup>			Rheumatoid arthritis + sDMARD <sup>b</sup>		
	Confirmed rheumatoid arthritis, n	Registered rheumatoid arthritis, n	Positive predictive value % (95% CI)	Confirmed rheumatoid arthritis, n	Registered rheumatoid arthritis, n	Positive predictive value % (95% CI)
Overall	205	331	61.9 (56.6–67.0)	185	211	87.7 (82.5–91.5)
Gender						
Men	61	100	61.0 (51.0–70.2)	55	63	87.3 (76.2–93.7)
Women	144	231	62.3 (55.9–68.4)	130	148	87.8 (81.4–92.2)
Age <sup>c</sup> (years)						
<50	13	19	68.4 (42.8–86.3)	10	11	90.1 (46.3–99.2)
50–59	49	72	68.1 (56.2–78.0)	44	50	88.0 (75.2–94.7)
60–69	76	133	57.1 (48.5–65.4)	68	82	82.9 (73.0–89.7)
>70	67	107	62.6 (52.9–71.4)	63	68	92.7 (83.2–97.0)
Department type						
Rheumatology	138	192	71.9 (65.0–77.8)	128	142	90.1 (84.0–94.1)
Internal medicine	38	65	58.5 (45.9–70.0)	33	39	84.6 (68.9–93.2)
Other	29	74	39.2 (28.6–51.0)	24	30	80.0 (60.8–91.2)
Patient type						
Inpatient	86	134	64.2 (55.6–71.9)	77	84	91.7 (83.3–96.0)
Outpatient	115	189	60.9 (53.7–67.6)	104	123	84.6 (76.9–90.0)
Other	4	8	50.0 (14.3–85.7)	4	4	100
Diagnosis type						
Main (A-diagnosis)	175	257	68.1 (62.1–73.5)	160	181	88.4 (82.8–92.3)
Secondary (B-diagnosis)	30	74	40.5 (29.8–52.3)	25	30	83.3 (64.3–93.3)
Calendar periods						
1977–1987	13	19	68.4 (42.8–86.3)	10	12	83.3 (45.7–96.7)
1988–2010	144	232	62.1 (55.6–68.1)	131	149	87.9 (81.6–92.3)
2011–2016	48	80	60.0 (48.7–70.3)	44	50	88.0 (75.2–94.7)
ICD-8 codes	29	37	78.4 (61.5–89.2)	25	27	92.6 (72.8–98.3)
ICD-10 codes <sup>d</sup>						
M05	108	150	72.0 (64.2–78.7)	102	111	92.0 (85.0–95.8)
M06	68	144	47.2 (39.1–55.5)	58	73	79.5 (68.4–87.4)

**Notes:** <sup>a</sup>Patients with first-time ever rheumatoid arthritis diagnosis registration in the Danish National Patient Registry; <sup>b</sup>patients with first-time ever rheumatoid arthritis diagnosis registration in the Danish National Patient Registry who subsequently, at least once, redeemed a prescription for a sDMARD; <sup>c</sup>patients' age at the time of rheumatoid arthritis diagnosis registration; <sup>d</sup>ICD-10 code M05 – seropositive rheumatoid arthritis, ICD-10 code M06 – other rheumatoid arthritis.

**Abbreviations:** ICD, International Classification of Diseases; sDMARD, synthetic disease-modifying antirheumatic drug.



(13/150) of the “M05” and 25% (36/144) of the “M06” diagnoses, no sufficient medical record information was available, including IgM RF and/or ACPA test results, and these cases were thus defined as nonverified RA.

Table 3 shows the PPVs of first-time RA diagnoses according to serological status, registered in the Danish National Patient Registry with one of the ICD-10 RA codes. Using the Model 1 approach, 62% of registry-based, seropositive RA diagnoses were confirmed (95% CI: 53.9–69.5). The PPV of other RA diagnoses was 25% (95% CI: 18.5–32.8). Stratification by age group at the time of diagnosis registration revealed the highest PPVs in the youngest age group (<50 years) for both seropositive (PPV: 100%) and other (PPV: 50%) RA. However, there were only a few RA diagnosis registrations in both groups (Table 3). After linking the data to the RA treatment codes (Model 2), the PPV increased for both seropositive (PPV: 80%; 95% CI: 71.6–86.7) and other RA (PPV: 41%; 95% CI: 30.2–52.9) diagnoses. The highest PPV for seropositive RA diagnosis was in the youngest age group (PPV: 100%), while the highest PPV for other RA diagnosis was in the oldest age group (PPV: 60%; 95% CI: 40.9–76.5) (Table 3).

Calculating the PPVs by gender showed that the PPVs for seropositive RA diagnoses were slightly higher for women than for men (using Model 1 – 63.4% vs 59.2%; Model 2 – 81.6% vs 77.1%), whereas the PPVs for other RA diagnoses

were higher for men than for women (using Model 1 – 28.2% vs 23.8%; Model 2 – 45.0% vs 39.6%) (Table 4).

Table 5 shows the PPVs calculated stratifying by calendar periods. The PPVs for seropositive RA diagnoses were similar for both periods. Though, the PPVs for other RA diagnoses after 2010 were higher than in the period before (using Model 1 – 40.0% vs 18.2%; Model 2 – 73.9% vs 26.0%).

Figure S1 and Tables S1 and S2 show available IgM RF and ACPA blood test results and their distribution among confirmed RA diagnoses. In almost 70% of verified RA diagnoses, ACPA test results were not available at the time of RA diagnosis registration.

The most likely other diagnosis among nonverified RA diagnoses was osteoarthritis. Table 6 shows the distribution of the most likely alternative diagnoses. Post hoc analyses were performed taking some of the most likely alternative diagnoses into account. The following additional criterion in determining RA cases was used: patients with their first RA diagnosis in the Danish National Patient Registry who had subsequently, at least once, redeemed a prescription for a sDMARD and who had no ICD code for psoriatic or enteropathic arthropathies (ICD-8: 696.0; ICD-10: M07), systemic connective tissue disorders (ICD-8: 446, 716, 734; ICD-10: M30–M35), inflammatory bowel diseases (ICD-8: 563; ICD-10: K50–K51), or sarcoidosis (ICD-8: 135.99;

**Table 3** Positive predictive values of first-time rheumatoid arthritis diagnoses according to serological subtypes registered in the Danish National Patient Registry in the period 1977–2016 among the participants in the Danish Diet, Cancer, and Health cohort and derived from the Central Denmark Region hospitals, overall and by age groups at the time of rheumatoid arthritis diagnosis

First-time diagnosis in the Danish National Patient Registry, ICD-10 codes	Rheumatoid arthritis <sup>a</sup>			Rheumatoid arthritis + sDMARD <sup>b</sup>		
	Confirmed rheumatoid arthritis, n	Registered rheumatoid arthritis, n	Positive predictive value % (95% CI)	Confirmed rheumatoid arthritis, n	Registered rheumatoid arthritis, n	Positive predictive value % (95% CI)
M05 – seropositive rheumatoid arthritis	93	150	62.0 (53.9–69.5)	89	111	80.2 (71.6–86.7)
<50 years <sup>c</sup>	2	2	100 <sup>d</sup>	2	2	100
50–59 years <sup>c</sup>	26	33	78.8 (60.6–90.0)	24	28	85.7 (66.0–94.9)
60–69 years <sup>c</sup>	34	65	52.3 (40.0–64.4)	33	43	76.7 (61.2–87.3)
>70 years <sup>c</sup>	31	50	62.0 (47.5–74.7)	30	38	79.0 (62.4–89.5)
M06 – other rheumatoid arthritis	36	144	25.0 (18.5–32.8)	30	73	41.1 (30.2–52.9)
<50 years <sup>c</sup>	1	2	50 <sup>d</sup>	0	0	0
50–59 years <sup>c</sup>	1	19	5.3 (0.6–33.8)	0	6	0
60–69 years <sup>c</sup>	15	66	22.7 (14.0–34.7)	12	37	32.4 (18.9–49.7)
>70 years <sup>c</sup>	19	57	33.3 (22.1–46.9)	18	30	60.0 (40.9–76.5)

**Notes:** <sup>a</sup>Patients with the first-time ever rheumatoid arthritis diagnosis registration in the Danish National Patient Registry; <sup>b</sup>patients with the first-time ever rheumatoid arthritis diagnosis registration in the Danish National Patient Registry who subsequently, at least once, redeemed a prescription for a sDMARD; <sup>c</sup>patients' age at the time of rheumatoid arthritis diagnosis registration. <sup>d</sup>Stata did not provide 95% CI in this case.

**Abbreviations:** ICD, International Classification of Diseases; sDMARD, synthetic disease-modifying antirheumatic drug.

**Table 4** Positive predictive values of first-time rheumatoid arthritis diagnoses according to serological subtypes registered in the Danish National Patient Registry in the period 1977–2016 among the participants in the Danish Diet, Cancer, and Health cohort and derived from the Central Denmark Region hospitals by gender

First-time diagnosis in the Danish National Patient Registry, ICD-10 codes	Rheumatoid arthritis <sup>a</sup>			Rheumatoid arthritis + sDMARD <sup>b</sup>		
	Confirmed rheumatoid arthritis, n	Registered rheumatoid arthritis, n	Positive predictive value % (95% CI)	Confirmed rheumatoid arthritis, n	Registered rheumatoid arthritis, n	Positive predictive value % (95% CI)
Men						
M05 – seropositive rheumatoid arthritis	29	49	59.2 (44.6–72.3)	27	35	77.1 (59.5–88.6)
M06 – other rheumatoid arthritis	11	39	28.2 (15.9–44.9)	9	20	45.0 (23.8–68.2)
Women						
M05 – seropositive rheumatoid arthritis	64	101	63.4 (53.4–72.3)	62	76	81.6 (71.0–88.9)
M06 – other rheumatoid arthritis	25	105	23.8 (16.5–33.0)	21	53	39.6 (27.1–53.7)

**Notes:** <sup>a</sup>Patients with the first-time ever rheumatoid arthritis diagnosis registration in the Danish National Patient Registry; <sup>b</sup>patients with the first-time ever rheumatoid arthritis diagnosis registration in the Danish National Patient Registry who subsequently, at least once, redeemed a prescription for a sDMARD.

**Abbreviations:** ICD, International Classification of Diseases; sDMARD, synthetic disease-modifying antirheumatic drug.

**Table 5** Positive predictive values of first-time rheumatoid arthritis diagnoses according to serological subtypes registered in the Danish National Patient Registry in the period 1977–2016 among the participants in the Danish Diet, Cancer, and Health cohort and derived from the Central Denmark Region hospitals by calendar periods

First-time diagnosis in the Danish National Patient Registry, ICD-10 codes	Rheumatoid arthritis <sup>a</sup>			Rheumatoid arthritis + sDMARD <sup>b</sup>		
	Confirmed rheumatoid arthritis, n	Registered rheumatoid arthritis, n	Positive predictive value % (95% CI)	Confirmed rheumatoid arthritis, n	Registered rheumatoid arthritis, n	Positive predictive value % (95% CI)
M05 – seropositive rheumatoid arthritis						
1988–2010	71	115	61.7 (52.4–70.3)	68	84	81.0 (70.9–88.1)
2011–2016	22	35	62.9 (45.1–77.7)	21	27	77.8 (57.0–90.2)
M06 – other rheumatoid arthritis						
1988–2010	18	99	18.2 (11.7–27.2)	13	50	26.0 (15.4–40.3)
2011–2016	18	45	40.0 (26.4–55.4)	17	23	73.9 (50.9–88.6)

**Notes:** <sup>a</sup>Patients with the first-time ever rheumatoid arthritis diagnosis registration in the Danish National Patient Registry; <sup>b</sup>patients with the first-time ever rheumatoid arthritis diagnosis registration in the Danish National Patient Registry who subsequently, at least once, redeemed a prescription for a sDMARD.

**Abbreviations:** ICD, International Classification of Diseases; sDMARD, synthetic disease-modifying antirheumatic drug.

ICD-10: D86) in the previous or following years in the Danish National Patient Registry. Using this additional criterion, 96% of register-based overall RA diagnoses were confirmed (95% CI: 91.7–98.1); 85% of seropositive RA diagnoses (95% CI: 76.0–90.6) and 48% of other RA diagnoses (95% CI: 34.2–62.1) were confirmed.

## Discussion

We found that, based on the ACR, ACR/EULAR classification criteria, or clinical case definition, only 62% of registry-based first-time RA diagnoses could be verified. Verification of RA diagnoses according to serological status was as low as 62% for seropositive RA and only 25% for other RA.

These findings are in line with those reported in previous Danish studies. Pedersen et al<sup>9</sup> also validated RA diagnoses in the Danish National Patient Registry, using data from two large population-based cohorts. They identified 217 RA diagnoses with available medical records registered at least once in the Danish National Patient Registry during 1977–2001. A total of 59% of RA diagnoses were verified using the clinical case RA definition as a reference, and only 46% of RA diagnoses were verified using the ACR 1987 classification criteria as a reference. Another Danish study validated first-time RA diagnoses recorded in the Danish National Patient Registry during 2011.<sup>10</sup> A total of 1,468 RA diagnoses were identified at a rheumatology department with

**Table 6** The most likely alternative diagnoses among patients, first-time registered with rheumatoid arthritis diagnosis in the Danish National Patient Registry during 1977–2016, but not confirmed by reviewing medical records

Alternative diagnoses	Number of nonconfirmed cases
Total number of nonconfirmed cases	106
Osteoarthritis	29% (n=31)
PMR or PMR with giant cell arteritis	9% (n=10)
Crystal arthropathy	9% (n=10)
Psoriatic arthritis	8% (n=9)
Connective tissue diseases	5% (n=5)
Tenosynovitis	3% (n=3)
Seronegative oligoarthritis	2% (n=2)
Enteropathic arthritis	1% (n=1)
Reactive arthritis	1% (n=1)
Juvenile idiopathic arthritis	1% (n=1)
Rotator cuff syndrome	1% (n=1)
Subacromial bursitis	1% (n=1)
Dupuytren's contracture	1% (n=1)
Arthralgia	1% (n=1)
Joint cyst	1% (n=1)
Myalgia	1% (n=1)
Unknown	26% (n=27)

**Abbreviation:** PMR, polymyalgia rheumatica.

one of the ICD-10 codes M05.9, M06.0, M06.8, or M06.9, and having an additional visit due to RA within 90 days. They verified 79% register-based RA diagnoses, which is in line with our findings of 72% verified RA diagnoses registered at a rheumatology department.

We included all register-based RA diagnoses in the analysis, even if medical records or IgM RF and ACPA blood test results were unavailable or inaccessible. Therefore, we estimated the lowest possible PPVs within the study population.

Our findings of relatively low overall PPVs may have several explanations. First of all, referral RA diagnoses may be carried forward as tentative diagnoses until conclusive determination of diagnoses can be made. Second, physicians are expected to record as many relevant diagnoses as possible during daily clinical practice; however, no routine validation takes place, and self-reported arthritis may be registered as RA. We found that the most likely alternative diagnoses were osteoarthritis, polymyalgia rheumatica, crystal arthropathy, psoriatic arthritis, and connective tissue diseases, a finding that supports these misclassifications pitfalls. Finally, the registry may contain self-reported RA as a secondary diagnosis if patients were treated in private rheumatology clinics, and these are not captured by the Danish National Patient Registry. Private practice medical records were not accessible for

this study. It is, however, unlikely that this would have had a substantial impact on the overall results since the number of diagnoses with no sufficient medical records was low (7%).

In our study, register-based RA diagnoses were almost equally distributed among seropositive and other RA. Several inception cohorts report that majority of RA patients are seropositive.<sup>27–29</sup> The discrepancy may have several explanations. Our study showed that register-based other RA diagnoses are misclassified more frequently than seropositive RA diagnoses. Moreover, in our study, both register-based and confirmed other RA diagnoses were observed more frequently among older patients (>60 years) than younger patients (<60 years). Thus, the mean age at the time of RA diagnosis in our study was in favor of other RA. Furthermore, incident RA cohorts may not precisely reflect the real RA population due to their inclusion and exclusion criteria. Lower PPVs were estimated for the other RA codes than the seropositive RA codes. Given that RA is a slowly progressing disease with clinical similarities to other joint diseases, a tendency to record other RA rather than seropositive RA may appear if the RA diagnosis is uncertain. This misclassification may also be explained by physicians' lack of awareness when selecting "M05" or "M06" RA codes in clinical practice.

The mean age at the time of RA diagnosis in our data was 65 years, which is higher than expected for the disease onset as indicated by inception cohorts.<sup>27,28</sup> This may partly be attributed to the eligibility criteria for the Danish Diet, Cancer, and Health cohort.<sup>19</sup> Furthermore, reduced functional status due to long-standing RA and concurrent comorbid conditions may have prevented potential attendees from participating in the Danish Diet, Cancer, and Health cohort.

We have observed that PPVs of other RA diagnoses registered during 2011–2016 were substantially higher than for diagnoses registered before 2011. It could partly be explained by availability of medical records, including IgM RF/ACPA test results.

PPVs increased substantially for both overall RA diagnoses and serological subtypes when diagnoses were linked to RA treatment codes. This is in line with the findings of an American study of RA diagnosis validation.<sup>11</sup> In that study, the PPV also increased substantially (from 66% to 81%) when the data were linked to RA treatment codes.

Our study has some limitations that merit discussion. First of all, we included RA diagnoses derived from only the Central Denmark Region hospitals, and therefore approximately 22% of the Danish population.<sup>22</sup> Although the Danish health care system is uniformly organized throughout



all regions, including the diagnosis registration procedure, regional differences may appear, and hospitals may vary in their registration practices, with higher PPVs anticipated for more accurate registrations.

Furthermore, the density of private rheumatology clinics differs between the regions,<sup>30</sup> and the PPVs would be expected to be lower in regions with a higher density of private practice rheumatologists due to inaccessibility to medical records. Still, our overall findings are in line with those reported by two previous Danish studies,<sup>9,10</sup> suggesting that such organizational differences had only a limited impact on the overall findings. It must, moreover, be taken into account that the structure of administrative registries differs between countries, and therefore our findings are not more broadly generalizable.

Unfortunately, we were unable to assess sensitivity, specificity, or negative predictive values due to the inaccessibility of medical records at private clinics. Furthermore, some RA patients may not seek medical advice, and thereby not be captured by any medical record or registry. These measures could have led to more precise assessment of the diagnostic accuracy of registry-based RA diagnoses.

## Conclusion

Our findings show that data for first-time RA diagnoses in the Danish National Patient Registry must be used with caution due to misclassification. However, linking registry-based RA diagnoses to the subsequent RA treatment codes increases the probability of identifying true RA diagnoses. The relatively high PPV of RA diagnosis registered with one of ICD-10 "M05" codes indicates a reliable means of coding seropositivity.

## Acknowledgments

The authors would like to thank the Danish Cancer Society for the funding of the Danish Diet, Cancer, and Health cohort study; the Steering Committee of the cohort study, for providing the opportunity to conduct the present study; programmer Katja Boll, for preparation of the dataset; and the participants of the cohort study. The present study was funded by the Danish Rheumatism Association, the Danish Heart Foundation, North Denmark Regional Hospital, Central Denmark Region, and Scandinavian Rheumatology Research Foundation, Grant 2014. The funders were not involved in the study design, data collection, analysis, interpretation of the results, manuscript writing, or decision to submit the paper for publication.

## Author contributions

KO, KSP, AT, and AL contributed to the conception and design of the study. KO and AL were involved in the acquisition of data. All authors contributed toward data analysis, drafting and critically revising the paper, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Dougados M, Soubrier M, Antunez A, et al. Prevalence of comorbidities in rheumatoid arthritis and evaluation of their monitoring: results of an international, cross-sectional study (COMORA). *Ann Rheum Dis*. 2014;73(1):62–68.
2. Myasoedova E, Davis JM, Crowson CS, Gabriel SE. Epidemiology of rheumatoid arthritis: rheumatoid arthritis and mortality. *Curr Rheumatol Rep*. 2010;12(5):379–385.
3. Neovius M, Simard JF, Askling J, ARTIS Study Group. Nationwide prevalence of rheumatoid arthritis and penetration of disease-modifying drugs in Sweden. *Ann Rheum Dis*. 2011;70(4):624–629.
4. Lindhardsen J, Ahlehoff O, Gislason GH, et al. The risk of myocardial infarction in rheumatoid arthritis and diabetes mellitus: a Danish nationwide cohort study. *Ann Rheum Dis*. 2011;70(6):929–934.
5. Lindhardsen J, Ahlehoff O, Gislason GH, et al. Risk of atrial fibrillation and stroke in rheumatoid arthritis: Danish nationwide cohort study. *BMJ*. 2012;344:e1257.
6. Hyldgaard C, Hilberg O, Pedersen AB, et al. A population-based cohort study of rheumatoid arthritis-associated interstitial lung disease: comorbidity and mortality. *Ann Rheum Dis*. 2017;76(10):1700–1706.
7. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health*. 2011;39(7 Suppl):30–33.
8. World Health Organization. *Classifications*. Geneva: World Health Organization. Available from: <http://www.who.int/classifications/icd/en/>. Accessed October 15, 2018.
9. Pedersen M, Klarlund M, Jacobsen S, Svendsen AJ, Frisch M. Validity of rheumatoid arthritis diagnoses in the Danish National Patient Registry. *Eur J Epidemiol*. 2004;19(12):1097–1103.
10. Ibfelt EH, Sørensen J, Jensen DV, et al. Validity and completeness of rheumatoid arthritis diagnoses in the nationwide DANBIO clinical register and the Danish National Patient Registry. *Clin Epidemiol*. 2017;9(9):627–632.
11. Singh JA, Holmgren AR, Noorbaloochi S. Accuracy of Veterans Administration databases for a diagnosis of rheumatoid arthritis. *Arthritis Rheum*. 2004;51(6):952–957.
12. Weyand CM, McCarthy TG, Goronzy JJ. Correlation between disease phenotype and genetic heterogeneity in rheumatoid arthritis. *J Clin Invest*. 1995;95(5):2120–2126.
13. Chibnik LB, Keenan BT, Cui J, et al. Genetic risk score predicting risk of rheumatoid arthritis phenotypes and age of symptom onset. *PLoS One*. 2011;6(9):e24380–24387.
14. Lahiri M, Morgan C, Symmons DP, Bruce IN. Modifiable risk factors for RA: prevention, better than cure? *Rheumatology*. 2012;51(3):499–512.
15. Ajeganova S, Huizinga TW. Rheumatoid arthritis: seronegative and seropositive RA: alike but different? *Nat Rev Rheumatol*. 2015;11(1):8–9.
16. Pedersen M, Jacobsen S, Klarlund M, et al. Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis Res Ther*. 2006;8(4):R133.

17. Wesley A, Bengtsson C, Elkan AC, et al. Association between body mass index and anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis: results from a population-based case-control study. *Arthritis Care Res.* 2013;65(1):107–112.
18. Lahiri M, Luben RN, Morgan C, et al. Using lifestyle factors to identify individuals at higher risk of inflammatory polyarthritis (results from the European Prospective Investigation of Cancer-Norfolk and the Norfolk Arthritis Register – the EPIC-2-NOAR Study). *Ann Rheum Dis.* 2014;73(1):219–226.
19. Tjønneland A, Olsen A, Boll K, et al. Study design, exposure variables, and socioeconomic determinants of participation in Diet, Cancer and Health: a population-based prospective cohort study of 57,053 men and women in Denmark. *Scand J Public Health.* 2007;35(4):432–441.
20. Gjerstorff ML. The Danish Cancer Registry. *Scand J Public Health.* 2011;39(7 Suppl):42–45.
21. Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. *Scand J Public Health.* 2011;39(7 Suppl):38–41.
22. Danske Regioner. Regional Denmark. København: Danske Regioner. Available from: <http://www.regioner.dk/services/in-english/regional-denmark>. Accessed June 18, 2018.
23. Ropes MW, Bennett GA, Cobb S, Jacox R, Jessar RA. 1958 Revision of diagnostic criteria for rheumatoid arthritis. *Bull Rheum Dis.* 1958;9(4):175–176.
24. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum.* 1988;31(3):315–324.
25. Aletaha D, Neogi T, Silman AJ, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum.* 2010;62(9):2569–2581.
26. WHO Collaborating Centre for Drug Statistics Methodology. *ATC. Structure and Principles*. Oslo, Norway: WHO Collaborating Centre for Drug Statistics Methodology. Available from: [https://www.whooc.no/atc/structure\\_and\\_principles/](https://www.whooc.no/atc/structure_and_principles/). Accessed June 18, 2018.
27. Hetland ML, Stengaard-Pedersen K, Junker P, et al. Combination treatment with methotrexate, cyclosporine, and intraarticular betamethasone compared with methotrexate and intraarticular betamethasone in early active rheumatoid arthritis: an investigator-initiated, multicenter, randomized, double-blind, parallel-group, placebo-controlled study. *Arthritis Rheum.* 2006;54(5):1401–1409.
28. Hørslev-Petersen K, Hetland ML, Junker P, et al. Adalimumab added to a treat-to-target strategy with methotrexate and intra-articular triamcinolone in early rheumatoid arthritis increased remission rates, function and quality of life. The OPERA Study: an investigator-initiated, randomised, double-blind, parallel-group, placebo-controlled trial. *Ann Rheum Dis.* 2014;73(4):654–661.
29. Sandberg MEC, Bengtsson C, Källberg H, et al. Overweight decreases the chance of achieving good response and low disease activity in early rheumatoid arthritis. *Ann Rheum Dis.* 2014;73(11):2029–2033.
30. sundhed.dk. København, Denmark: sundhed.dk. Available from: <https://www.sundhed.dk/borger/guides/find-behandler/?Informationskategori=Speciallæge&InformationsUnderkategori=Reumatolog> (gigtlæge). Accessed June 18, 2018.

## Supplementary materials

**Table S1** The blood test results of IgM RF and ACPA among confirmed ICD-10 M05 (seropositive) RA diagnoses in the Danish National Patient Registry in the period 1977–2016 among the participants in the Danish Diet, Cancer, and Health cohort and derived from the Central Denmark Region hospitals

Confirmed ICD-10 M05 diagnoses, n=108	IgM RF <sup>a</sup>				
		Positive, % (n)	Negative, % (n)	Not available, % (n)	Overall, % (n)
	ACPA <sup>a</sup>				
	Positive, % (n)	25 (27)	0	0	25 (27)
	Negative, % (n)	4.6 (5)	0.9 (1)	0	5.6 (6)
	Not available, % (n)	56.5 (61)	6.5 (7)	6.5 (7)	69.4 (75)
	Overall, % (n)	86.1 (93)	7.4 (8)	6.5 (7)	100 (108)

**Notes:** <sup>a</sup>Blood test results before, at or within 1 month after the first-time registration in the Danish National Patient Registry.

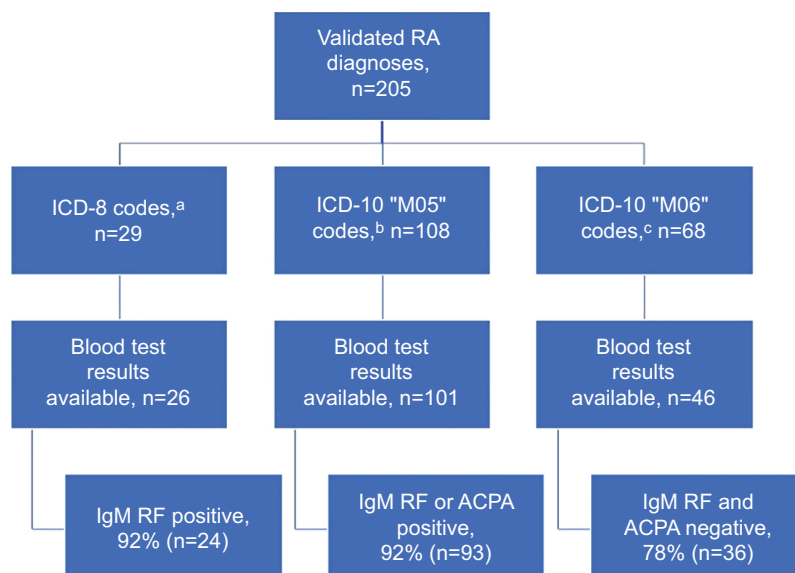
**Abbreviations:** ACPA, anti-citrullinated protein antibody; ICD, International Classification of Diseases; IgM RF, immunoglobulin M rheumatoid factor; RA, rheumatoid arthritis.

**Table S2** The blood test results of IgM RF and ACPA among confirmed ICD-10 M06 (other) RA diagnoses in the Danish National Patient Registry in the period 1977–2016 among the participants in the Danish Diet, Cancer, and Health cohort and derived from the Central Denmark Region hospitals

Confirmed ICD-10 M06 diagnoses, n=68	IgM RF <sup>a</sup>				
		Positive, % (n)	Negative, % (n)	Not available, % (n)	Overall, % (n)
	ACPA <sup>a</sup>				
	Positive, % (n)	2.9 (2)	1.4 (1)	0	4.3 (3)
	Negative, % (n)	(0)	26.5 (18)	0	26.5 (18)
	Not available, % (n)	10.3 (7)	26.5 (18)	32.4 (22)	69.2 (47)
	Overall, % (n)	13.2 (9)	54.4 (37)	32.4 (22)	100 (68)

**Notes:** <sup>a</sup>Blood test results before, at or within 1 month after the first-time registration in the Danish National Patient Registry.

**Abbreviations:** ACPA, anti-citrullinated protein antibody; ICD, International Classification of Diseases; IgM RF, immunoglobulin M rheumatoid factor; RA, rheumatoid arthritis.



**Figure S1** Serological status among confirmed RA diagnoses with available blood test results on IgM RF and/or ACPA before, at or within 1 month after the first-time registration in the Danish National Patient Registry in the period 1977–2016 among the participants in the Danish Diet, Cancer, and Health cohort and derived from the Central Denmark Region hospitals.

**Notes:** <sup>a</sup>Registered with the ICD-8 code "712.39" (Arthritis rheumatoides alia et non specificata) at their first registration in the Danish National Patient Registry. <sup>b</sup>Registered with one of ICD-10 "M05" (seropositive RA) codes at their first registration in the Danish National Patient Registry. <sup>c</sup>Registered with one of ICD-10 "M06" (other RA) codes at their first registration in the Danish National Patient Registry.

**Abbreviations:** ACPA, anti-citrullinated protein antibody; ICD, International Classification of Diseases; IgM RF, immunoglobulin M rheumatoid factor; RA, rheumatoid arthritis.

## Clinical Epidemiology

Dovepress

### Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification,

systematic reviews, risk and safety of medical interventions, epidemiology and biostatistical methods, and evaluation of guidelines, translational medicine, health policies and economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>