BMJ Open

# Improving accuracy of self-reported diagnoses of rheumatoid arthritis in the French prospective E3N-EPIC cohort: a validation study

Yann Nguyen ,[1,2] Carine Salliot,[1,3] Gaëlle Gusto,[1] Elise Descamps,[2] Xavier Mariette,[2,4] Marie-Christine Boutron-Ruault,[1,5] Raphaèle Seror[2,4]

**To cite:** Nguyen Y, Salliot C, Gusto G, *et al*. Improving accuracy of self-reported diagnoses of rheumatoid arthritis in the French prospective E3N-EPIC cohort: a validation study. *BMJ Open* 2019;**9**:e033536. doi:10.1136/bmjopen-2019-033536

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2019-033536).

M-CB-R and RS are joint senior authors.

Received 09 August 2019
Revised 03 October 2019
Accepted 28 October 2019



For numbered affiliations see end of article.

**Correspondence to**
Dr Marie-Christine Boutron-Ruault;
marie-christine.BOUTRON@gustaveroussy.fr

## ABSTRACT

**Objectives** The French E3N-EPIC (*Etude Epidémiologique auprès des femmes de la Mutuelle générale de l'Education Nationale*-European Prospective Investigation into Cancer and Nutrition) cohort enrolled 98 995 women aged 40 to 65 years at inclusion since 1990 to study the main risk factors for cancer and severe chronic conditions in women. They were prospectively followed with biennially self-administered questionnaires collecting self-reported medical, environmental and lifestyle data. Our objective was to assess the accuracy of self-reported diagnoses of rheumatoid arthritis (RA) and to devise algorithms to improve the ascertainment of RA cases in our cohort.

**Design** A validation study.

**Participants** Women who self-reported an inflammatory rheumatic disease (IRD) were asked to provide access to their medical record, and to answer an IRD questionnaire. Medical records were independently reviewed.

**Primary and secondary outcome measures** Positive predictive values (PPV) of self-reported RA alone, then coupled with the IRD questionnaire, and with a medication reimbursement database were assessed. These algorithms were then applied to the whole cohort to ascertain RA cases.

**Results** Of the 98 995 participants, 2692 self-reported RA. Medical records were available for a sample of 399 participants, including 305 who self-reported RA. Self-reported RA was accurate only for 42% participants. Combining self-reported diagnoses to answers to a specific IRD questionnaire or to the medication reimbursement database improved the PPV (75.6% and 90.1%, respectively). Using the devised algorithms, we could identify 964 RA cases in our cohort.

**Conclusion** Accuracy of self-reported RA is poor but adding answers to a specific questionnaire or data from a medication reimbursement database performed satisfactorily to identify RA cases in our cohort. It will subsequently allow investigating many potential risk factors of RA in women.

## Strengths and limitations of this study

- Two algorithms were devised and tested to improve accuracy of self-reported diagnosis of rheumatoid arthritis in a large population-based cohort.
- A large sample of medical records was available and independently reviewed to test the devised algorithm.
- Nearly 1000 cases of rheumatoid arthritis were identified, which will subsequently allow investigating many potential risk factors of rheumatoid arthritis in this cohort.
- The control population was women who self-reported another rheumatic disease and not healthy women.
- The sample of medical records was not provided at random.

## INTRODUCTION

Rheumatoid arthritis (RA) is the most common autoimmune inflammatory rheumatic disease (IRD) in adults, and is a major cause of functional alteration and handicap. RA is a complex multifactorial autoimmune disease in which both genetic and environmental factors interact in the pathogenesis of the disease to trigger autoimmunity.[1]

Little is known about environmental factors that may contribute to the disease, except smoking, which has been reproducibly reported as associated with an increased risk of anti-citrullinated protein autoantibody (ACPA)-positive RA, particularly in individuals carrying the HLA-DRB1-shared epitope alleles.[2–6] The role of other environmental factors has been suggested but results were rarely reproducible. Only epidemiological studies, such as case-control studies or cohort studies can appropriately address the question. The main advantage of case-control studies is that cases are easily ascertained, with detailed phenotypes and easy availability of biological data, but their main limits are a retrospective collection of environmental factors, the risk of hindsight and recall bias and a potentially biassed control population. Cohort studies offer the advantage of having

a prospective collection of environmental factors before disease onset and a non-biassed non-cases population. However, collected information about disease phenotypes is usually limited, and in large population-based cohorts, diagnoses are often self-reported.

The diagnostic accuracy of self-reported RA has been studied in various populations, and varies considerably, between 7% and 96%.[7–15] One of the evocated reasons is the confusion between RA and other forms of arthritis, mainly osteoarthritis (OA), the prevalence of which being higher than RA in general populations.[16] If the accuracy of self-reported diagnosis is poor, using self-reported RA alone as case definition might create an ascertainment bias, because of the high rate of false-positive cases.

To overcome this lack of accuracy, some studies have used a linkage with national patient registries, primary healthcare records and/or hospital discharge databases usually based on International Classification of Diseases codes.[17–21] However, such registries are not always available, and these methods can also lack specificity.[22] Other studies have ascertained self-reported RA through linkage with a medical record review, or even with clinical examination of all suspected cases.[23–25] However, in large cohorts, medical record screening is time-consuming, expensive and subject to difficulties in obtaining patients' consents and medical charts.[12] These difficulties underscore the need for increasing accuracy of RA case definition based on self-reported and/or other available information.

Our primary objective was to evaluate the accuracy of self-reported diagnoses of RA in a French population-based cohort and to determine if the use of additional information obtained from a dedicated questionnaire and from a medication reimbursement database could improve their accuracy. A secondary objective was to use the devised algorithms to identify RA cases in this large cohort for subsequent epidemiological studies.

## MATERIAL AND METHODS

### The E3N-EPIC cohort study

The E3N cohort study (*Etude Epidémiologique auprès des femmes de la Mutuelle générale de l'Education Nationale*) is a French prospective cohort study including 98 995 women living in France and covered by a national health insurance scheme primarily involving teachers.[26] This study is also the French component of the European Prospective Investigation into Cancer and Nutrition (EPIC). It was initiated in France in 1990 to study the main risk factors for cancer and severe chronic conditions in women. Participants ages were 40 to 65 at inclusion. After the baseline questionnaire (Q1), participants were biennially mailed questionnaires (Q2 to Q12) to update their health-related information and newly diagnosed diseases. The last questionnaire to date (Q12) was sent in 2018, but corresponding data are not yet available. In addition, a drug-reimbursement claims database has been available since 2004 for all cohort women from their medical insurance records (*Mutuelle Générale de l'Éducation Nationale* (MGEN)). The average follow-up rate per questionnaire has been 83% and, overall, the total proportion of patients lost to follow-up since 1990 was <3% in 2014. All women gave written informed consent, and approvals were obtained from the French National Commission for Data Protection and Individual Freedom (327346-V14) and the French Advisory Committee on Information Processing in Material Research in the Field of Health (13.794).

### Participants

In three follow-up questionnaires (Q9, Q10 and Q11, sent in 2007, 2011 and 2014, respectively), study participants self-reported a diagnosis of IRD (RA and/or spondyloarthritis (SpA)) by answering the following questions: 'Do you have RA?' (yes/no) at Q9, Q10 and Q11, and 'Do you have ankylosing spondylitis' (yes/no) at Q10 and Q11, together with the date of IRD diagnosis. In addition, women were asked at each questionnaire from baseline if they had been hospitalised since the last questionnaire, and if so, they had to specify the reasons for those admissions. All women who self-reported RA or SpA in questionnaires and/or in hospitalisation reasons were eligible to participate in the validation study, those who self-reported SpA serving as a control population.

### IRD questionnaire design

A specific IRD questionnaire was designed to ascertain diagnoses of RA and SpA (online supplementary appendix 1). The questionnaire was adapted from a telephone questionnaire designed by Guillemin *et al*, with reference to the signs, symptoms and epidemiological criteria for RA (American College of Rheumatology 1987).[27 28] In this IRD questionnaire, women had the possibility to confirm or retract their self-reported diagnosis (online supplementary appendix 1, Q0, Q1). We included additional questions: if a physician confirmed the diagnosis (only a general practitioner, a rheumatologist and/or an internist), date of diagnosis, date of first symptoms, presence of ACPA and current and past treatments.

All eligible women were sent this specific IRD questionnaire with an information letter and were asked to send back the questionnaire and their medical chart comprising all relevant medical documents in relation with their rheumatic condition, including medical reports, laboratory findings, hand and foot radiographs and results of rheumatoid factors (RF) and ACPA testing, when available. A first mailing was sent on June 2017, and a reminder was sent in December 2017 to those who did not answer the first one.

### RA ascertainment algorithm from IRD questionnaire

Based on data from the IRD questionnaire, a decision algorithm aimed at improving the accuracy of self-reported RA was devised by a consensus of rheumatologists (RS, XM and ED). We considered as RA cases women who confirmed having RA in the IRD specific questionnaire,

and self-reported at least one of the following: (1) RA diagnosis confirmed by a rheumatologist and/or another physician (internal medicine specialist or general practitioner), (2) taking or having taken any of the RA conventional synthetic disease modifying anti-rheumatic drugs (DMARDs) or biological DMARDs (listed in online supplementary appendix 1, Question 34), (3) having positive RF or ACPA or (4) at least four of the seven 1987 American College ofRheumatology (ACR) criteria (listed in online supplementary appendix 1, Questions 8,9,11,14–18).

### RA ascertainment algorithm from medication reimbursement database

The MGEN medication reimbursement database included, for all E3N participants, all medications delivered by community-based pharmacies since 2004. Thus, medications only delivered by hospital pharmacies (ie, intravenous infusions), and medications used before 2004 were not available.

Using this medication reimbursement database, we devised a second algorithm: women were considered as RA cases if they self-reported having RA, and had had reimbursements for any conventional synthetical or biological DMARD used in the treatment of RA, including methotrexate, leflunomide, any subcutaneous tumour necrosis factor alpha (TNF-α) inhibitor and subcutaneous abatacept or tocilizumab. Oral steroids, being widely used for other reasons, were not considered specific enough to be included in this definition. This algorithm had been previously used to ascertain RA cases in our cohort.[29] All algorithms are reported in detail in online supplementary table 1.

### RA cases ascertainment: medical chart review

Medical records were obtained from the IRD questionnaire mailing for a subset of women and included medical reports from hospitalisation and/or from outpatient medical visits, laboratory findings and/or bone X-rays. They were independently reviewed by two trained rheumatologists (YN and RS), blinded to the self-reported diagnoses and confirmed cases or not according to the RA identification algorithm. Classification was based on reviewer's expertise, and not on strict ACR 1987 criteria or ACR/European League against Rheumatism 2010 criteria,[28 30] and was used as the reference to assess the accuracy of self-reported diagnosis of RA alone and associated with additional information from the specific IRD questionnaire and from the medication reimbursement database. If the provided medical data were enough to confirm a diagnosis, reviewers classified women as RA, or not RA (including alternate diagnoses, such as OA, SpA or other). Disagreements between the two reviewers were resolved by consensus. If diagnosis could not be ascertained by medical chart review, cases were considered as uncertain and were not used to determine the accuracy of the algorithms.

### Identification of RA cases in the E3N cohort

Since we expected that the accuracy of self-reported RA diagnoses alone would not be sufficient, we used the devised algorithms to identify RA cases in our cohort (including women who did not provide their medical records). For women who answered the IRD questionnaire, we used the algorithm based on this questionnaire, and for those who self-reported RA in Q9, Q10 and/or Q11 but did not answer the specific IRD questionnaire, were deceased or lost to follow-up, we subsequently used the algorithm based on the medication reimbursement database. Women with available medical record who were identified as RA cases by these algorithms were reassessed as non-cases if their diagnosis was invalidated by medical chart review (false-positive cases).

### Statistical analysis

To assess the accuracy of self-reported diagnosis alone, and the two algorithms based on the IRD questionnaire and/or the medication reimbursement database, we used the classification based on medical chart review as the reference standard. Thus, this assessment was performed on the subset of participants with an available medical chart and for whom its review allowed to classify them as case or non-case. The level of agreement between each algorithm and the chart review diagnoses was assessed by the kappa statistic with 95% CIs. Positive predictive value (PPV) and negative predictive value (NPV), sensitivity and specificity of each algorithm were calculated.

Finally, a descriptive analysis of demographic characteristics was performed on all women enrolled in the E3N study, on women who self-reported RA, on those who self-reported RA and provided their medical charts, on chart-reviewed confirmed RA and on RA cases identified by combining self-report to the IRD questionnaire and/or the medication reimbursement database. All analyses were carried out using the SAS software, V.9.4 (SAS Institute Inc, Cary, North Carolina, USA).

### Patient and public involvement

Patients were involved in this validation study. Our validation study relied on a self-completed patient questionnaire adapted from a previous questionnaire not designed to be sent by mail. We modified the questionnaire for this purpose and added some questions on X-rays, and on ACPA and RF testing. To make sure that the revised questionnaire could be clearly understandable by patients, a patients' association (Association Française des Polyarthrites et rhumatismes inflammatoires chroniques (AFPric)) helped us to review the contents and wording of the questionnaire. The findings from this study will be shared with E3N participants through the next newsletter.

## RESULTS

### IRD case identification

Among the 98 995 participants, 3230 women self-reported RA and/or SpA and were eligible to participate in the validation study: 2692 self-reported RA, 637 self-reported

Table 1 Baseline characteristics of the study population

| N | All women (n=98 995) | Self-reported RA (n=2692) | Self-reported RA with available medical records (n=305) | Confirmed RA after chart review (n=129) | Identified RA with devised algorithms (n=964) |
|---|---|---|---|---|---|
| Age at Q1 (years) | 49.4 (6.7) | 51.1 (6.7) | 49.6 (5.6) | 48.5 (5.2) | 50.2 (6.3) |
| Year of birth | | | | | |
| <1930 | 7808 (7.9) | 278 (10.3) | 13 (4.3) | 2 (1.6) | 59 (6.1) |
| (1930–1940) | 31 529 (31.9) | 1114 (41.4) | 112 (36.7) | 37 (28.7) | 380 (39.4) |
| (1940–1950) | 56 647 (57.2) | 1247 (46.3) | 177 (58.0) | 88 (68.1) | 509 (52.8) |
| ≥1950 | 3011 (3.0) | 53 (2.0) | 3 (1.0) | 2 (1.6) | 16 (1.7) |
| Body mass index at Q1 (kg/m²) | 22.6 (3.2) | 23.2 (3.4) | 23.0 (2.9) | 22.9 (2.9) | 23.0 (3.4) |
| Smoking status | | | | | |
| Not available | 945 (1.0) | 17 (0.6) | 0 (0) | 0 (0) | 7 (0.7) |
| Current smoker | 14 755 (14.8) | 420 (15.6) | 40 (13.1) | 16 (12.4) | 158 (16.4) |
| Non smoker | 53 130 (53.7) | 1465 (54.4) | 176 (57.7) | 75 (58.1) | 504 (52.3) |
| Former smoker | 30 165 (30.5) | 790 (29.4) | 89 (29.2) | 38 (29.5) | 295 (30.6) |
| Passive smoking in childhood | 12 854 (13.0) | 398 (14.8) | 48 (15.7) | 19 (14.7) | 158 (16.4) |
| Education level | | | | | |
| Not available | 4277 (4.3) | 136 (5.1) | 14 (4.6) | 5 (3.9) | 55 (5.7) |
| <High school | 16 185 (16.4) | 597 (22.2) | 61 (19.9) | 19 (14.7) | 186 (19.3) |
| Up to 2 years after high school | 44 986 (45.4) | 1186 (44.1) | 131 (43.0) | 57 (44.2) | 432 (44.8) |
| ≥3 years after high school | 33 547 (33.9) | 773 (28.6) | 99 (32.5) | 48 (37.2) | 291 (30.2) |
| Socio-professional category | | | | | |
| Not available | 15 800 (16.0) | 337 (12.5) | 25 (8.2) | 11 (8.5) | 106 (11.0) |
| Teacher | 62 013 (62.6) | 1632 (60.6) | 198 (64.9) | 86 (66.7) | 609 (63.2) |
| Higher managerial and professional occupations | 2499 (2.5) | 83 (3.1) | 9 (3.0) | 3 (2.3) | 28 (2.8) |
| Intermediate occupations | 15 340 (15.5) | 495 (18.4) | 58 (19.0) | 27 (20.9) | 179 (18.6) |
| Unemployed | 2602 (2.6) | 106 (3.9) | 10 (3.3) | 1 (0.8) | 28 (2.8) |
| Other | 741 (0.8) | 39 (1.5) | 5 (1.6) | 1 (0.8) | 14 (1.5) |
| Deprivation index | −0.3 (1.0) | −0.2 (1.0) | −0.1 (1.0) | −0.2 (0.9) | −0.3 (1.1) |

Results are presented as n (%) for categorical variables and mean (SD) for continuous variables.
RA, rheumatoid arthritis.

SpA and 109 women self-reported both RA and SpA. Demographic characteristics of the whole cohort, and of women who self-reported RA is described in table 1.

### RA cases ascertainment: medical chart review

Mailings were sent to 2924 of the eligible women (306 women could not be contacted because of death or withdrawn consent), with a recall letter for those who failed to answer. The specific IRD questionnaire was sent back by 2182 eligible women (74.6%), including 1833 women who self-reported RA (84%). Medical charts were sent by 594 women (20.3%). Among them, 195 (32.8%) could not be classified because of insufficient provided medical data and were therefore excluded from the performance study. Thus, 399 women provided sufficient medical data to ascertain their diagnosis. Among them, 129 (32.3%) were classified as RA cases, 60 (15.0%) as SpA cases and 210 (52.6%) as having another diagnosis (ie, osteoarthritis or other diagnosis). All 399 women completed the IRD questionnaire and had available medication reimbursement data on the MGEN database. The accuracy of the different diagnosis algorithms has been assessed on this subset of 399 women. Among the 399 women, 305 had self-declared RA. The demographic characteristics of these 305 women are described in table 1.

### Determination of accuracy of self-reported diagnosis and validation algorithms

Accuracy of the validation algorithms compared with medical chart review is described in table 2. Of the 305 women who self-reported RA with an available medical chart, only 125 (41%) were confirmed by chart review,

Table 2 Agreement between self-reported rheumatic disease and medical chart review

| Self-reported diagnosis | N | Available medical chart, n | Confirmed cases, n | Agreement between self-report and medical chart review, n (%) |
|---|---|---|---|---|
| RA | 2692 | 305 | 129 | 125 (40.9) |
| RA only | 2583 | 290 | 129 | 122 (42.1) |
| SpA | 637 | 90 | 60 | 48 (53.3) |
| SpA only | 528 | 75 | 60 | 42 (56.0) |
| RA and SpA | 109 | 15 | 0 | 0 (0.0) |
| Total | 3230 | 399 | | |

RA, rheumatoid arthritis; SpA, spondylarthritis.

leading to a PPV and specificity of self-report of 41% and 33%, respectively. Concordance between self-reported RA alone and medical chart review was low (kappa statistic=0.2).

The addition of the IRD questionnaire dramatically improved PPV and specificity (table 3). When combining self-reported RA with the IRD questionnaire algorithm (any of the four definitions), PPV was 72%, sensitivity 94% and specificity 83%, with a kappa statistic of 0.7. The combination associated with the best performances (highest PPV, sensitivity and specificity) was self-reported RA plus use of any specific RA medication; the one with the lowest specificity was self-reported RA plus confirmation by a rheumatologist of another physician. The combinations of self-reported RA with positive RF and/or ACPA or with the ACR criteria were specific but had the lowest sensitivities. Alternate diagnoses for the false-positive cases detected by this algorithm are reported in table 4.

Using medication reimbursement data from the MGEN database also improved PPV and sensitivities of self-report alone (table 3). If women self-reported RA and had at least one reimbursement of any RA specific medication, PPV was 90%, sensitivity 71%, specificity 87% and kappa coefficient 0.7. With this algorithm, 10 women were detected by the medication reimbursement database but did not have RA (false-positive cases, table 4). All of them had received methotrexate. Also, 38 women were not detected by this algorithm but had RA (false-negative): 21 received methotrexate before 2004, thus before the onset of the MGEN reimbursement database, five received intravenous biological DMARDs not available in the database and 27 received treatments which were not specific enough of RA (online supplementary table 2).

Combining self-report to both IRD questionnaire and medication reimbursement database improved PPV (98%) but considerably lowered sensitivity (67%), with no amelioration of the kappa value (table 3).

### Identification of RA cases in the E3N cohort

Finally, we used both algorithms to identify RA cases in our cohort. Among the 1833 women who answered the IRD questionnaire and self-declared RA, 904 RA cases (49.3%) were confirmed by the algorithm based on the IRD questionnaire (self-reported RA and any of the four definitions). Among them we excluded the 47 (5.2%) false-positive cases (based on medical chart review) and 34 (3.8%) RA cases without diagnosis date, thus not allowing to know whether they were incident or prevalent. Finally, 823 (44.9%) RA cases were identified by this algorithm. The second algorithm based on the MGEN reimbursement database was used on the 859 remaining eligible women who self-reported RA but did not answer the questionnaire, and identified 141 (16.4%) RA cases. Overall, 964 RA cases were detected by one of the two algorithms, including 698 incident cases and 266 prevalent cases, during a mean follow-up of 25.2 years (figure 1). In addition, 65.1% of our identified cases have been identified by at least two methods, and 16.4% and 21% have even been validated by three or four methods, respectively (online supplementary table 3). Demographic characteristics of the identified RA cases are shown in table 1.

## DISCUSSION

In this large prospective cohort of French adult women, we examined the accuracy of self-reported diagnoses of RA and provided interesting information regarding the way to validate these diagnoses. As expected, in our study, the accuracy of self-reported diagnoses of RA was poor. But, combining self-report to a specific IRD questionnaire providing addition self-reported data and/or to a medication reimbursement database, dramatically improved accuracy of RA diagnoses, with high sensitivity, specificity and PPV. Using these algorithms, we could detect nearly 1000 RA cases in this cohort.

The accuracy of self-reported RA diagnoses has previously been evaluated in other cohorts.[7–9 12 13 15 23 24] Reliability, sensitivity and specificity of self-reported RA varied widely, depending on how the question was phrased, and on the confirmation method (diagnostic registries, chart review, use of ACR criteria and/or clinical evaluation). When compared with chart review, PPV varies between 7% and 35%.[8 9 15 24 31] In the Nurses' Health Study,[23] Karlson *et al* only confirmed 7% of the original self-reported RA, by reviewing the medical charts to look if women fulfilled the ACR criteria. In our cohort, self-reported diagnoses

Table 3 Agreement between self-report of RA alone, combined to the IRD questionnaire and to the medication reimbursement database with chart review

| | Chart review (reference standard) | | | Positive predictive value, % | Negative predictive value, % | Sensitivity, % | Specificity, % | Kappa coefficient (95% CI) |
|---|---|---|---|---|---|---|---|---|
| | Yes | No | Total | | | | | |
| **Self-report of RA** | | | | | | | | |
| Yes | 125 | 180 | 305 | 41.0 | 95.7 | 96.9 | 33.3 | 0.22 (0.17 to 0.28) |
| No | 4 | 90 | 94 | | | | | |
| Total | 129 | 270 | 399 | | | | | |
| **Self-report of RA + IRD questionnaire** | | | | | | | | |
| *1. Confirmation by a rheumatologist or an internal medicine specialist* | | | | | | | | |
| Yes | 120 | 43 | 166 | 72.3 | 96.1 | 93 | 83 | 0.71 (0.65 to 078) |
| No | 9 | 224 | 233 | | | | | |
| Total | 129 | 270 | 399 | | | | | |
| *2. RA medication* | | | | | | | | |
| Yes | 118 | 11 | 129 | 91.5 | 95.9 | 91.5 | 95.9 | 0.87 (0.82 to 0.93) |
| No | 11 | 259 | 270 | | | | | |
| Total | 129 | 270 | 399 | | | | | |
| *3. Positive RF and/or ACPA* | | | | | | | | |
| Yes | 72 | 3 | 75 | 96.0 | 82.4 | 55.8 | 98.9 | 0.61 (0.53 to 0.70) |
| No | 57 | 267 | 324 | | | | | |
| Total | 129 | 270 | 399 | | | | | |
| *4. ACR criteria* | | | | | | | | |
| Yes | 63 | 7 | 70 | 90.0 | 79.9 | 48.8 | 97.4 | 0.52 (0.43 to 0.61) |
| No | 66 | 263 | 329 | | | | | |
| Total | 129 | 270 | 399 | | | | | |
| *Any of these four definitions* | | | | | | | | |
| Yes | 121 | 47 | 168 | 72.0 | 96.5 | 93.8 | 82.6 | 0.71 (0.64 to 0.78) |
| No | 8 | 223 | 231 | | | | | |
| Total | 129 | 270 | 399 | | | | | |
| **Self-report of RA + medication reimbursement database** | | | | | | | | |
| Yes | 91 | 10 | 101 | 90.1 | 87.3 | 70.5 | 87.3 | 0.71 (0.63 to 0.78) |
| No | 38 | 260 | 298 | | | | | |
| Total | 129 | 270 | 399 | | | | | |
| **Self-report of RA + IRD questionnaire + medication reimbursement database** | | | | | | | | |
| Yes | 86 | 2 | 88 | 97.7 | 86.2 | 66.7 | 99.3 | 0.72 (0.64 to 0.79) |
| No | 43 | 268 | 311 | | | | | |
| Total | 129 | 270 | 399 | | | | | |

ACPA, anti-citrullinated protein autoantibody; ACR, American College of Rheumatology; IRD, inflammatory rheumatic disease; RA, rheumatoid arthritis; RF, rheumatoid factors.

of RA were accurate for ~40% of the cases. Comparison with other studies, mainly involving English language questionnaires, might be difficult. Indeed, our higher rate of accurate diagnoses could be partially explained by language differences, RA and osteoarthritis being phonetically close in English, but not in French.

Nevertheless, this accuracy was not sufficient. Thus, to improve the accuracy of RA diagnosis, we used self-reported data from an IRD questionnaire, derived from a validated questionnaire designed to validate RA and SpA cases by phone interviews in a population of patients of 10 French university hospital rheumatology units.[27] We adapted it with the help of a patients' association that reviewed the wording and phrasing to make it clearly understandable to general population subjects, and we added questions about the presence or absence of RF and/or ACPA and on RA medication. Using this questionnaire, self-report of RA combined to a self-reported use of RA medication had the excellent accuracy, with both high sensitivity and specificity. Although very specific,

**Table 4** Alternate diagnoses for false-positive cases detected by the algorithms

| | Alternate diagnosis |
|---|---|
| False-positive cases detected by self-report + IRD questionnaire, n=39 | Osteoarthritis (n=24)<br>Scapulohumeral periarthritis (n=5)<br>Polymyalgia rheumatica (n=3)<br>Primary Sjögren's syndrome (n=3)<br>Systemic lupus erythematosus (n=2)<br>Osteoporosis (n=1)<br>Lumbar sciatic (n=1) |
| False-positive cases detected by self-report + reimbursement database, n=10 | Psoriatic arthritis (n=7)<br>Systemic lupus erythematosus (n=2)<br>Osteoarthritis associated with inflammatory bowel disease (n=1) |

IRD, inflammatory rheumatic disease.

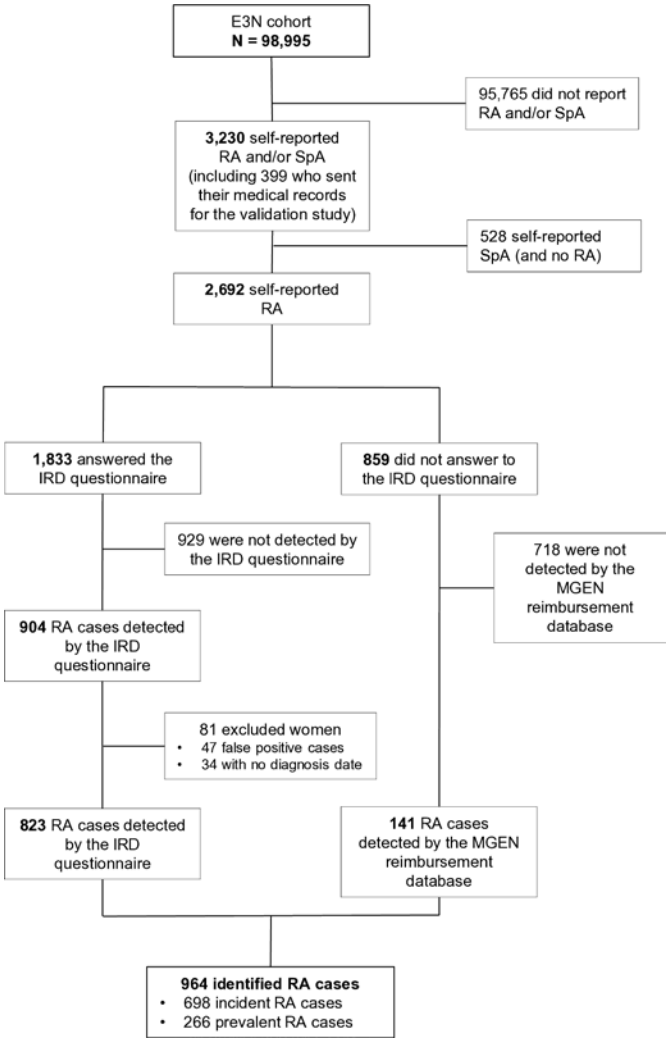


**Figure 1** Flow chart of the identification of RA cases in the E3N cohort. E3N,'Etude Epidémiologique auprèsdes femmes de la Mutuelle générale de l'Education Nationale'; IRD, inflammatory rheumatic disease; MGEN, 'Mutuelle Générale del'Education Nationale'; RA, rheumatoid arthritis; SpA, spondylarthritis.

and useful for further disease phenotyping, a self-report of positive RF and/or ACPA resulted in a low sensitivity and using this definition might miss RA cases. Using the ACR criteria in the IRD questionnaire resulted in a low sensitivity, because those criteria were not designed to be used in self-reported questionnaires, nevertheless they were highly specific. Our results demonstrate that the use of a limited list of items, particularly focusing on specific medications, in a dedicated questionnaire could drastically improve self-report accuracy.

We also assessed the performance of the algorithm using the medication reimbursement database. This method had been used to identify RA cases in the first study on RA in the E3N cohort study.[29] As expected, the algorithm has an excellent specificity and PPV, but underestimates the number of RA cases. Indeed, the database included all medications delivered by community-based pharmacies since 2004 and we only considered methotrexate, leflunomide, subcutaneous TNF-α inhibitors and subcutaneous abatacept or tocilizumab; therefore we could not detect RA cases treated before 2004 and no longer treated with those drugs, those only treated by intravenous biologics delivered by hospital pharmacies only, and those with other treatments (eg, hydroxychloroquine). Thus, if an exhaustive medication reimbursement database was available, using this algorithm could probably lead to both high specificities and high sensitivities.

Using both algorithms, we detected nearly 1000 RA cases, mainly incident cases. Since a proper evaluation with the reference standard (ie, medical chart review) was not available for all women, there might be some false-positive RA cases among them. But given the number of methods used to limit their number and their accuracy, this rate might be small.

We acknowledge some limitations to the present study. First, it was not designed to estimate the number of unreported RA cases in our cohort. Our population of non-cases were women who did not self-report RA but self-reported another IRD, which could bias our results. Ideally, we would have analysed medical records from women who did not report any IRD to determine the proportion of cases missed. Thus, reported sensitivities and NPVs should be interpreted with caution. However, our main concern was to avoid false-positive cases that is, to ascertain detected cases, rather than to avoid missing a few cases. Therefore, there may be a few undetected RA cases in the control group, but the number of these cases is likely to be small, and, given the large number of non-cases in our cohort, the risk of bias induced by the false-negative cases is negligible. Also, our validation study relies on an additional questionnaire. Answers to this questionnaire were not obtained for all women, which might have created a response bias. However, such bias was limited by using the medication reimbursement database for women who did not answer to the IRD questionnaire.

Another limitation could be the representativeness of the sample of women who provided their medical

records, sent on a voluntary basis, thus not at random. This could have introduced a selection bias toward more severe disease, inflating the accuracy. However, medical chart review confirmed the diagnosis of RA in only 41% of them, showing that both cases and non-cases provided medical chart. Also, women who provided their medical charts did not differ from other women who self-reported IRD in terms of age or education level, which may limit the bias.

Finally, the algorithms we devised to improve accuracy of self-reported RA diagnoses could prove useful to validate RA diagnoses in other population-based cohorts. However, they could be difficult to transpose from the French care setting to another one; thus, all data potentially available for validation (medication database, national patient registries, primary care records and/or hospital discharge databases) must be considered.

## CONCLUSIONS

To conclude, our study highlights the poor accuracy of self-reported RA diagnoses, even among educated women. We demonstrated that this accuracy could be improved using medication reimbursement data and/or other self-reported data from a specific questionnaire. Even if ascertaining RA diagnoses with a complete medical chart review might probably be one of the best option, it appears that obtaining other information, particularly on RA specific treatment, either from the patients themselves or from health insurance databases can be a reasonably good alternative, sparing the difficulties of obtaining complete medical charts, and the time and cost of medical chart review. Even much less sensitive, obtaining confirmation of ACPA or RF positivity from patients was also highly specific, and offer the advantage of giving a key phenotypic characteristic, particularly important when studying RA risk factors. Our results could help other teams that aim at ascertaining RA cases in large epidemiological studies. Also, the validation of almost 1000 RA cases in our cohort will serve as a basis to future epidemiological studies, since the design and the long follow-up of participants of our cohort will be used to investigate many potential RA risk factors.

**Author affiliations**
[1]Centre for Research in Epidemiology and Population Health, (CESP), INSERM U1018, Université Paris-Saclay, Université Paris-Sud, Villejuif, France
[2]Rheumatology, Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpitaux universitaires Paris-Sud – Hôpital Bicêtre, Le Kremlin-Bicetre, France
[3]Rheumatology, Centre Hospitalier Régional d'Orléans, Orléans, France
[4]Université Paris-Sud, Center for Immunology of Viral Infections and Auto-immune Diseases (IMVA), INSERM U1184, Université Paris-Saclay, Le Kremlin-Bicêtre, France
[5]Gustave Roussy Institute, Villejuif, France

**Acknowledgements** The authors are indebted to all participants for their continued participation. The authors would like to thank Pascale Gerbouin-Rerolle, Maxime Valdenaire and Roselyn Rima Gomes for their help on data management. They also acknowledge the AFPric patients' association that helped to review the wording and phrasing of the validation questionnaire, particularly Patricia Preiss and Angelique Hochodé.

**Contributors** All authors contributed to the manuscript. YN, CS, ED, XM, MCB and RS were responsible for conception and design. YN, GG, MCB and RS were responsible for collection of data and analysis. All authors were responsible for the interpretation of data. YN and RS wrote the first version of the manuscript. All authors critically revised and approved the final version of the manuscript.

**Funding** The present work was performed using data from the Inserm E3N cohort and support from the MGEN, Gustave Roussy and the Ligue contre le Cancer for setting up and maintaining the cohort. The study was also supported by a state grant ANR-10-COHO-0006 from the Agence Nationale de la Recherche within the Investissement d'Avenir program, and by by a research grant from FOREUM Foundation for Research in Rheumatology. In addition, this study was conducted thanks to the help of an unrestricted grant from the Société Française de Rhumatologie.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** This study was approved by the French authorities ('Comité Consultatif sur le Traitement de l'information en matière de Recherche dans le domaine de la Santé' and 'Commission Nationale de l'Informatique et des Libertés'). An informed consent was obtained from all patients.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request.



**ORCID iD**
Yann Nguyen http://orcid.org/0000-0002-0866-3824

## REFERENCES

1 Klareskog L, Padyukov L, Rönnelid J, *et al*. Genes, environment and immunity in the development of rheumatoid arthritis. *Curr Opin Immunol* 2006;18:650–5.
2 Karlson EW, Chang S-C, Cui J, *et al*. Gene–environment interaction between *HLA-DRB1* shared epitope and heavy cigarette smoking in predicting incident rheumatoid arthritis. *Ann Rheum Dis* 2010;69:54–60.
3 Linn-Rasker SP, van der Helm-van Mil AHM, van Gaalen FA, *et al*. Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients who carry HLA-DRB1 shared epitope alleles. *Ann Rheum Dis* 2006;65:366–71.
4 Lee H-S, Irigoyen P, Kern M, *et al*. Interaction between smoking, the shared epitope, and anti–cyclic citrullinated peptide: a mixed picture in three large North American rheumatoid arthritis cohorts. *Arthritis Rheum* 2007;56:1745–53.
5 Bang S-Y, Lee K-H, Cho S-K, *et al*. Smoking increases rheumatoid arthritis susceptibility in individuals carrying the HLA-DRB1 shared epitope, regardless of rheumatoid factor or anti-cyclic citrullinated peptide antibody status. *Arthritis Rheum* 2010;62:369–77.
6 Willemze A, van der Woude D, Ghidey W, *et al*. The interaction between HLA shared epitope alleles and smoking and its contribution to autoimmunity against several citrullinated antigens. *Arthritis & Rheumatism* 2011;63:1823–32.
7 Kvien TK, Glennås A, Knudsrød OG, *et al*. The validity of self-reported diagnosis of rheumatoid arthritis: results from a population survey followed by clinical examinations. *J Rheumatol* 1996;23:1866–71.
8 Star VL, Scott JC, Sherwin R, *et al*. Validity of self-reported rheumatoid arthritis in elderly women. *J Rheumatol* 1996;23:1862–5.
9 Ling SM, Fried LP, Garrett E, *et al*. The accuracy of self-report of physician diagnosed rheumatoid arthritis in moderately to severely disabled older women. women's health and aging Collaborative Research Group. *J Rheumatol* 2000;27:1390–4.
10 Barlow JH, Turner AP, Wright CC. Comparison of clinical and self-reported diagnoses for participants on a community-based arthritis self-management programme. *Rheumatology* 1998;37:985–7.
11 Cooper GS, Wither J, McKenzie T, *et al*. The prevalence and accuracy of self-reported history of 11 autoimmune diseases. *J Rheumatol* 2008;35:2001–4.

12 Walitt BT, Constantinescu F, Katz JD, *et al*. Validation of self-report of rheumatoid arthritis and systemic lupus erythematosus: the women's health Initiative. *J Rheumatol* 2008;35:811–8.
13 Formica MK, McAlindon TE, Lash TL, *et al*. Validity of self-reported rheumatoid arthritis in a large cohort: results from the black women's health study. *Arthritis Care Res* 2010;62:NA–41.
14 Gill TK, Hill CL. The accuracy of self-report in rheumatic diseases. *J Rheumatol* 2017;44:1109–11.
15 Videm V, Thomas R, Brown MA, *et al*. Self-Reported diagnosis of rheumatoid arthritis or ankylosing spondylitis has low accuracy: data from the Nord-Trøndelag health study. *J Rheumatol* 2017;44:1134–41.
16 O'Rourke JA, Ravichandran C, Howe YJ, *et al*. Accuracy of self-reported history of autoimmune disease: a pilot study. *PLoS One* 2019;14:e0216526.
17 Linauskas A, Overvad K, Berg Johansen M, *et al*. Positive predictive value of first-time rheumatoid arthritis diagnoses and their serological subtypes in the Danish national patient registry. *Clin Epidemiol* 2018;10:1709–20.
18 Linauskas A, Overvad K, Symmons D, *et al*. Body fat percentage, waist circumference, and obesity as risk factors for rheumatoid arthritis: a Danish cohort study. *Arthritis Care Res* 2019;71:777–86.
19 Sundstrom B, Johansson I, Rantapaa-Dahlqvist S. Interaction between dietary sodium and smoking increases the risk for rheumatoid arthritis: results from a nested case-control study. *Rheumatology* 2015;54:487–93.
20 Fisher BA, Cartwright AJ, Quirke A-M, *et al*. Smoking, Porphyromonas gingivalis and the immune response to citrullinated autoantigens before the clinical onset of rheumatoid arthritis in a southern European nested case-control study. *BMC Musculoskelet Disord* 2015;16:331.
21 Muller S, Hider SL, Raza K, *et al*. An algorithm to identify rheumatoid arthritis in primary care: a clinical practice research Datalink study. *BMJ Open* 2015;5:e009309.
22 Singh JA, Holmgren AR, Noorbaloochi S. Accuracy of Veterans administration databases for a diagnosis of rheumatoid arthritis. *Arthritis Care Res* 2004;51:952–7.
23 Karlson EW, Mandl LA, Hankinson SE, *et al*. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis? results from the nurses' health study. *Arthritis Rheum* 2004;50:3458–67.
24 Mikuls TR, Cerhan JR, Criswell LA, *et al*. Coffee, tea, and caffeine consumption and risk of rheumatoid arthritis: results from the Iowa women's health study. *Arthritis Rheum* 2002;46:83–91.
25 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;73:219–26.
26 Clavel-Chapelon F, van Liere MJ, Giubout C, *et al*. E3N, a French cohort study on cancer risk factors. E3N group. Etude Epidémiologique auprès de femmes de l'Education Nationale. *Eur J Cancer Prev* 1997;6:473–8.
27 Guillemin F, Saraux A, Fardellone P, *et al*. Detection of cases of inflammatory rheumatic disorders: performance of a telephone questionnaire designed for use by patient interviewers. *Ann Rheum Dis* 2003;62:957–63.
28 Arnett FC, Edworthy SM, Bloch DA, *et al*. The American rheumatism association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis & Rheumatism* 1988;31:315–24.
29 Seror R, Henry J, Gusto G, *et al*. Passive smoking in childhood increases the risk of developing rheumatoid arthritis. *Rheumatology* 2019;58:1154–62.
30 Aletaha D, Neogi T, Silman AJ, *et al*. Rheumatoid arthritis classification criteria: an American College of Rheumatology/ European League against rheumatism collaborative initiative. *Ann Rheum Dis* 2010;2010:1580–8.
31 Karlson EW, Sanchez-Guerrero J, Wright EA, *et al*. A connective tissue disease screening questionnaire for population studies. *Ann Epidemiol* 1995;5:297–302.