Estimating the diagnostic accuracy of rheumatoid factor in UK primary care: a study using the Clinical Practice Research Datalink (CPRD)

Anne Miller[1], Alison L Nightingale[2] Cormac J Sammon[2] Kamal R Mahtani[3], Tim A Holt[3], Neil J McHugh[2,4], Raashid A Luqmani[5]

Institutions:

1. Department of Rheumatology, Nuffield Orthopaedic Centre, Oxford University Hospitals NHS Trust.
2. Department of Pharmacy & Pharmacology, University of Bath, Claverton Down, Bath, BA2 7AY
3. Nuffield Department of Primary Health Care Sciences, University of Oxford, Radcliffe Observatory Quarter, Woodstock Rd, Oxford OX2 6GG
4. Department of Rheumatology, Royal National Hospital for Rheumatic disease, Bath
5. Nuffield Department of Orthopaedics, Rheumatology & Musculoskeletal Science, University of Oxford

Short Title: Diagnostic accuracy of RF testing in primary care

Corresponding author:

Dr Anne Miller, Department of Rheumatology, Nuffield Orthopaedic Centre, Oxford University Hospitals NHS Trust, Telephone 01865 741155, Fax 01865 738056 email anne.miller@ouh.nhs.uk

Funding; Medical Research Council's licence agreement with MHRA, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences

Abstract

Objective

To investigate the diagnostic accuracy of rheumatoid factor (RF) as a test for rheumatoid arthritis (RA) in primary care and impact on referral times using the Clinical Practice Research Datalink.

Methods

We identified all patients with a first RF test recorded in the CPRD between 1/1/2000 and 31/12/2008 and those diagnosed with RA within two years of testing. We calculated likelihood ratios (LRs), sensitivity, specificity and predictive values (PVs) of RF for a diagnosis of RA. We compared time to hospital referral in those testing positive and negative using Kaplan Meier failure curves and Log Rank tests.

Results

Of 62436 first RF tests, 4679 (7.5%) were positive. There were 1753 incident cases of RA of which 57.8% were sero-positive. The LR+ for RF was 9.5 (CI_{95} 9.0 to 10.0) and LR- was 0.5 (CI_{95} 0.4 to 0.5). Sensitivity and specificity were 57.8% (CI_{95} 55.4% to 60.1%) and 93.9% (CI_{95} 93.7% to 94.1%) and PPV and NPV were 21.4% (CI_{95} 20.3% to 22.6%) and 98.7% (CI_{95} 98.6% to 98.8%) respectively. Median time to first hospital contact after first RF test in those with sero-positive vs sero-negative results was 54 days (CI_{95} 49 to 58) vs 150 (CI_{95} 147 to 152).

Conclusions

Only 2.8% of patients undergoing RF testing were diagnosed with RA, suggesting RF is used to screen patients with musculoskeletal symptoms rather than those with more specific features of RA. A positive RF may be helpful in diagnosing RA in primary care but performs badly in excluding RA and may delay referral.

Key words: sensitivity specificity, rheumatoid factor, rheumatoid arthritis, primary care.

Introduction

Rheumatoid arthritis (RA) is a destructive inflammatory joint disease. Bone oedema leading to erosions has been demonstrated as early as 4 weeks after symptom onset[1] and early referral of patients with suspected RA for treatment with disease modifying anti-rheumatoid drugs (DMARDS) is advocated in order to prevent long term joint damage and disability.[2-5] Many GPs undertake investigations to support early diagnosis and guide referrals and this may include testing for rheumatoid factor (RF).[6, 7] RF has a pooled sensitivity of 69% (65% to 73%) and specificity of 85% (82% to 88%)[8] and has an established role in the classification of RA.[9] NICE guidelines specifically advise against a delay in urgent referral of patients with persistent synovitis who have a negative RF.[2] However, RF results have been shown to influence referral decisions and GPs may use a negative test to exclude RA despite the presence of clinical features suggesting the disease.[6, 10] This may lead to a delay in diagnosis for patients with RF negative disease, and referral of patients with a false positive test and poor clinical evidence of disease.

Most studies on the diagnostic utility of RF have taken place in early arthritis clinics in secondary care where the pre-test probability of RA is relatively high. Few studies have been conducted in general practice where the pre-test probability of RA is low.[6, 10, 11] However, most RF testing takes place in primary care.[11] It is important for GPs and patients to have information on the usefulness of RF in excluding RA or supporting referral. Our primary aim was to determine the diagnostic value of RF for RA in primary care by estimating likelihood ratios, sensitivity, specificity and predictive values using data from the UK Clinical Practice Research Datalink (CPRD). Our secondary aim was to determine the prognostic impact of the RF result on time to referral to secondary care.

Methods

Data Source

The UK Clinical Practice Research Datalink (CPRD) is the world's largest longitudinal primary care database containing data for approximately 13.3 million individuals registered with 613 UK primary care

practices. It has been found to be generally representative of the UK population.[12] GPs use a coded thesaurus of clinical terms, called Read Codes, to record clinical features, laboratory tests and results, diagnoses, and other information. We had access to Read codes but not to supplemental free text.

Ethical Approval

The study protocol was approved by the Independent Scientific Advisory Committee (ISAC) for the CPRD, protocol number 10_105R.

Study Population

The study period ran from 1st January 2000 to 31st December 2010. Patients had contributed at least 12 months of research standard data to the CPRD before the date of their first RF test, and at least two years of follow up data subsequently, unless they died during the two years' follow-up in which case they were included to date of death.

We identified all RF tests recorded as a RF titre or a Read code for a positive or negative RF test in the CPRD records of permanently registered patients aged 16 and over between 1st January 2000 and 31st December 2008. The first test in each patient's record was identified and patients with a record of a RF test before the study period were excluded. During the course of the study we observed a bias in the way RF was recorded in the CPRD with tests being more likely to be recorded if positive(13) and negative tests either being recorded without a RF titre or not recorded at all, possibly because positive tests are more clinically meaningful to GPs. Where there was evidence of preferential data recording, missing data were considered to be Missing Not at Random (MNAR)(14, 15). In order to minimise the impact of selection bias thus introduced, we stratified the data by general practice and calendar year ('practice-years') and excluded practice-years where the proportion of positive tests was >20%. This cut-off was based on a study by Miller et al(11) that found that the mean percentage of positive RF tests in primary care was 6.3%. Based on a

comparison of the distribution of the percentage of positive RF tests against that for the study by Miller, we considered that it was likely that practice-years with >20% of positive tests were likely to be preferentially recording positive tests and were thereby MNAR; the full methods of this investigation are described elsewhere(13). The remaining tests were from practice-years where missing data were likely to be missing at random, allowing for a complete case analysis to be conducted with a lower risk of selection bias(15). From this dataset we then selected tests where a test result or RF titre was available.

We excluded patients with a diagnosis of RA more than three months before their first RF test, those with pre-existing connective tissue diseases (systemic lupus erythematosus, poly/dermatomyositis, systemic sclerosis and Sjogren's syndome), or with psoriasis or inflammatory bowel disease at any time.

Classification of RF test results

RF results were classified as positive if the recorded titre was above the upper value of the normal range (UNR) in the electronic test record or there was a Read code for a positive RF in the medical records. Where normal ranges were not recorded we assumed that a practice had sent all RF tests to the same laboratory during that calendar year and imputed the UNR value as the mode value stratified by year and general practice. Due to the variation in the UNR cut-off we calculated the difference between the RF titre and the UNR cut-off value in order to investigate the relationship between RF titre and RA diagnosis. We denoted this as the 'off-set' RF value.

Identifying cases of rheumatoid arthritis (RA)

We searched the medical records of the study population for patients with a diagnosis of RA up to two years following the first RF test. We based our case identification on the algorithm validated by Thomas et al.[16] Those with a Read code for sero-positive or erosive disease and those with at least one prescription for a DMARD within two years of diagnosis were classified as cases of RA; those with more than one record of RA diagnosis but without any prescriptions for DMARDs were classified as 'probable' RA. Identification of

cases of RA was undertaken without knowledge of the RF result. The date of RA diagnosis was taken as the first record of RA or date of first prescription for a DMARD, whichever was earlier.

Covariate identification

For each patient in the study population we determined smoking status and body mass index records nearest to their first RF test. The presence of pre-existing diabetes mellitus was determined using a previously described algorithm.[17] Pre-existing thyroid disease and cancer in the year before diagnosis was determined. The individual socioeconomic status was not available for every patient in the study population; therefore, we used the general practice Index of Multiple Deprivation (IMD) score quintile where a score of five indicates the highest level of deprivation.

Time from RF test to hospital referral or attendance

We identified all records of referral, outpatient clinic attendance and inpatient discharge summaries that were coded as having a rheumatology specialty and determined the earliest date as first contact with rheumatology. Specialty-specific data are inconsistently recorded on the CPRD; therefore we identified all records of hospital referral, outpatient clinic attendance, letters and inpatient discharge summaries, and determined the earliest date of contact with any hospital department.

Statistical analyses

All analyses were undertaken using STATA Statistical Software version 12 (StataCorp. 2011. Stata: Release 12. Statistical Software. College Station, TX: StataCorp LP).

We calculated the positive and negative likelihood ratios (LRs), sensitivity, specificity and predictive values (PVs) of RF for a diagnosis of RA in patients with positive or negative RF tests. We performed sensitivity analyses excluding patients with "probable RA", and to determine the effect of classifying a patient as having RA if the RF value was equal to or above the upper normal range value. We performed a sensitivity analysis of all test data in the original study population, including data that were considered to be MNAR to quantify the impact of this on the measures of diagnostic accuracy. Modified Poisson regression models

with robust variance estimates [18] were used to calculate the relative risks (RRs) for a diagnosis of RA within two years associated with increasing offset RF titre values.

The distribution of baseline characteristics of patients with positive and negative RF test results was compared using chi-square tests. Modified Poisson regression models [18] were used to calculate adjusted RRs characteristics found to be associated with a positive RF test including co-morbidity and records of musculoskeletal symptoms in the six months before the test. Variables were selected for inclusion into the model using forwards stepwise Poisson regression models set to accept variables at a significance of $p \leq 0.2$.[19] Cox proportional hazards models were used to calculate the hazard ratio of death within two years of first RF test. We constructed Kaplan-Meier failure curves to investigate the effect of a positive and negative RF result on referral times to a rheumatologist and to any hospital contact following first RF test. Log rank tests were used to determine the equality of the failure curves in those with positive and negative RF tests.

Results

There were 91 293 eligible first RF tests with a test result recorded on the CPRD during the study period. We excluded 28 857 tests from practice-years where there was evidence that missing data were MNAR or there was an exclusion diagnosis present, resulting in a final study population of 62 436 RF tests. The number of eligible tests that were included increased from 2234 in 2000 to 9167 in 2008. This is likely to reflect a change from manual to automatic recording of test results. There were 4697 (7.5%) positive and 57 739 (92.5%) negative tests. 42 545 (68.1%) were undertaken in females. Forty-four patients (0.08%) with a negative first RF test had a subsequent positive test within two years of their first test.

Table 1 shows the baseline characteristics of the study population stratified by RF test result. The mean age of patients who were RF(+) was significantly older than those who were RF(-); mean difference 6.0 years (95% CI 5.6 to 6.5). Current smokers and those with unknown smoking status were more likely to have a positive test result relative to non-smokers. Those with a BMI< 20 had an increased risk relative to those with a BMI between 20 and 24.

1091 patients died within two years of their first RF test of whom 186 (17.1%) had a positive result. The adjusted HR for death within two years of first RF result was 1.5 (95% CI 1.3 to 1.8) after adjustment for age, sex and all variables found to be associated with a positive RF test.

Diagnostic accuracy of RF testing

There were 1740 incident cases of RA; 1285 of these had a record of a DMARD prescription after their date of RA diagnosis. 455 had no DMARD prescriptions within two years and were classified as 'probable' RA. 1005 (57.8%) of the RA cases had a positive and 735 (42.2%) a negative first RF result. Of those without a diagnosis of RA within two years, 3692 (6.1%) had a positive and 57 004 (93.9%) a negative RF test. Table 2 shows the results of the diagnostic accuracy analysis and Table 3 demonstrates that the relative risk for a diagnosis of RA increased with increasing RF titre.

Referral to secondary care

Of the 62 436 patients with a first RF test during the study period, 1695 had a referral to rheumatology prior to the date of their first RF and were excluded from the analysis of referral times leaving 60 741 eligible patients. Figure 1 shows the Kaplan-Meier failure curves for contact with a rheumatology department in the two years following the first RF test. The Kaplan-Meier curve illustrates that when contact with rheumatology services occurred, it tended to be a short time after the first RF test. A higher proportion of patients with a RF(+) test were referred to rheumatology within two years of testing ($p<0.0001$). 3924 (86.9%) of the patients with a RF(+) test and 44 335 (78.9%) of those with a RF(-) test had a record of any hospital contact within two years of their first RF test. Figure 2 shows the Kaplan-Meier curve for contact with any hospital department in the two years following first RF test. A higher proportion of patients with a RF(+) test were seen in hospital after their first RF test ($p<0.0001$).

For patients with a record of contact with a rheumatologist after their first RF test, the median referral time was 21 days (IQR 11, 60) in patients with a RF(+) test and 42 days (IQR 14, 193) in those with a (RF-) result. In those with a record of contact with any hospital department after their first RF test the median time

to contact was 33 days (IQR 11, 128 days); in those with a RF(+) test and 79 days (IQR 20, 247 days) with a RF(-) test.

Discussion

We report the diagnostic utility of RF in a large electronic dataset of primary care patients representing approximately 8.4% of the UK population. RF was found to be highly specific with a moderately good positive likelihood ratio of 9.5 (95% CI 9.0 to 10.0). However, its sensitivity was low and the negative likelihood ratio was poor at 0.5 (95% 0.4 to 0.5). The first RF test was positive for 57.8% incident cases of RA and higher titres were associated with a diagnosis of RA. We demonstrated a rise in the relative risk of a positive test with increasing age and a higher risk of death with a positive RF (RR 1.5). We demonstrated a delay in referral of patients with a negative RF result to rheumatology and to hospital in general, although 78.9 to 86.9% patients undergoing testing were referred for hospital review in the two years following the test.

The main strength of the study is that it is based in a primary care population, where most RF requests take place and is the largest study to date of the utility of RF as a test for RA in this setting. The CPRD is an observational dataset consisting of records kept by GPs in day to day patient management. It is a valuable resource for large epidemiological studies but has limitations, for example, the loss of patients to follow-up when they change practice and the lack of valid data prior to their entry onto the database. Diagnoses are recorded using Read codes and consequently definition of cases may be hampered by variation in recording methods. We identified cases of RA using an algorithm based on that validated by Thomas et al[16] with a sensitivity of 84% [95% CI 73 to 94%] and specificity 86% [95% CI 72 to 92%]. These figures suggest that a number of cases with Read codes for RA might be misclassified, introducing potential bias in our results.

The sensitivity analyses showed that the likelihood ratios from the primary analysis were robust in excluding ‘probable’ cases of RA but that changes to the definition of a positive RF test (sensitivity analysis (b)) and the inclusion of data that were considered to be MNAR (sensitivity analysis (d)) resulted in a substantial decrease in the LR+. When data that were considered to be MNAR were included, the proportion of positive

tests in the study population increased from 7.5% to 19.5% and the proportion of negative tests decreased from 92.5% to 80.5% resulting in a reduction in the LR+ but a stable LR-. Our analysis demonstrates the preferential recording of positive tests on the CPRD. The change from manual to automatic coding of test results on the CPRD was reflected in the increase in eligible tests from the beginning to the end of the study period. However, we do not have the practice-specific data regarding the dates of this change. Since we would expect 6.3% of RF tests in primary care to be positive, we have shown that the inclusion of data that are MNAR gives rise to the analysis of a biased population.

.

We acknowledge that people diagnosed with RA on clinical grounds and referred directly without RF testing in primary care were not included, and that not all study participants were assessed by a rheumatologist for the presence or absence of RA. This may have introduced detection bias because RF(+) patients were more likely to be referred to hospital, and RF(-) patients might have not been referred during the two year follow-up period, resulting in a higher specificity than expected.

Forty two per cent of patients diagnosed with RA had a negative first RF test in primary care which is comparable with a cohort of 1892 patients with new RA where 37% RF tests were negative.[20] In Table 4 we have compared our results with those of a meta-analysis[8] and pooled values from a systematic review.[21] The sensitivity of RF was lower in early arthritis studies suggesting poorer utility earlier in the disease and consistent with our results. A poor negative likelihood ratio is reported throughout the literature. Our finding that a positive RF is associated with an increased risk of death is in keeping with other population studies[22] although the risk estimate might be confounded by the development of comorbidity after diagnosis. In general a positive likelihood ratio greater than 10 and a negative likelihood ratio less than <0.1 are considered to provide strong evidence to rule in or rule out diagnoses respectively.[22, 23] This study, and previous analyses,[8, 21] have demonstrated the positive likelihood ratio for RF to be less than 10 and the negative likelihood ratio to be greater than 0.1 indicating moderate performance ruling in and poor

performance ruling out RA. Whilst higher titres are associated with a greater likelihood of the patient having RA, a negative test does not exclude it.

Only 2.8% of patients undergoing RF testing had a final diagnosis of RA suggesting that RF is used to screen patients presenting with musculoskeletal symptoms, rather than to support a diagnosis in patients with specific features of RA. This is in keeping with a recent UK study of RF requests in primary care where 5.8% of tests were positive and 4% had an eventual diagnosis of RA.[11]According to 2010 ACR/EULAR criteria,[9] RF has an important role in the diagnosis of RA. The low diagnostic yield demonstrated in primary care is largely due to incorrect use or interpretation. Up to 87% patients undergoing RF testing were referred to hospital and a positive RF was associated with a much reduced time to first hospital contact in comparison to those with a negative RF test. This may be either because patients with a RF(+) are more unwell and referred on clinical grounds, or that a RF result influences referral decisions. The main question raised in our study is whether RF should be requested and how it should be interpreted in primary care. These are also relevant questions for other immunological tests such as antinuclear antibodies and anti-citrullinated protein antibody (ACPA). It has been suggested that ACPA might be useful as a screening test for RA because of its high specificity and similar sensitivity to RF[24] and used to guide referrals in patients with suspected inflammatory joint disease.[25] However, if used indiscriminately in patients with a low pre-test probability of RA, neither a positive RF nor a positive ACPA test result affect clinical probabilities sufficiently to secure a diagnosis[26] and its utility as a test would be hampered by the same problems of over-use and over-interpretation.

Key messages

- Only 2.8% of patients undergoing RF testing in primary care are subsequently diagnosed with RA
- A negative RF test should not be used as evidence for the absence of RA
- Patients with symptoms consistent with RA should be referred to rheumatology for assessment without delay

Acknowledgements

We would like to thank Corinne de Vries, Professor of Pharmacoepidemiology and David Mant, Emeritus Professor of General Practice for their invaluable discussion. This work was supported by the Oxford NIHR Musculoskeletal Biomedical Research Unit, Nuffield Orthopaedic Centre, University of Oxford.

Conflict of interest statement

AN and CS received funding from the University of Oxford to conduct this project. RL reports non-financial support and other from GSK, personal fees from Roche, personal fees from Janssen, other from Nordic, other from Chemocentryx, personal fees from UCB, outside the submitted work. AM, KM, TH and NM have no disclosures.

Funding

This work was supported by the Medical Research Council's licence agreement with MHRA providing access to the Clinical Practice Research Database. This study is based in part on data from the CPRD (previously termed General Practice Research Database) obtained under licence from the UK Medicines and Healthcare Products Regulatory Agency. However, the interpretation and conclusions contained in this study are those of the authors alone. Further funding was provided by the Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford and the MSK Biomedical Research Unit, Oxford.

References

1. McGonagle D, Conaghan PG, O'Connor P, et al. The relationship between synovitis and bone changes in early untreated rheumatoid arthritis: a controlled magnetic resonance imaging study. Arthritis Rheum 1999;**42**(8):1706-11 doi: 10.1002/1529-0131(199908)42:8<1706::AID-ANR20>3.0.CO;2-Z [Published Online First: 2001/04/26].
2. Rheumatoid arthritis: National clinical guideline for management and treatment in adults. In: (NICE) NIfHaCE, ed. London: NICE, 2009.

3. Luqmani R, Hennell S, Estrach C, et al. British Society for Rheumatology and British Health Professionals in Rheumatology guideline for the management of rheumatoid arthritis (after the first 2 years). Rheumatology (Oxford) 2009;**48**(4):436-9 doi: 10.1093/rheumatology/ken450a [Published Online First: 2009/01/29].

4. van der Linden MP, le Cessie S, Raza K, et al. Long-term impact of delay in assessment of patients with early arthritis. Arthritis Rheum 2010;**62**(12):3537-46 doi: 10.1002/art.27692 [Published Online First: 2010/08/20].

5. Farragher TM, Lunt M, Fu B, Bunn D, Symmons DP. Early treatment with, and time receiving, first disease-modifying antirheumatic drug predicts long-term function in patients with inflammatory polyarthritis. Ann Rheum Dis 2010;**69**(4):689-95 doi: 10.1136/ard.2009.108639 [Published Online First: 2009/10/28].

6. Sinclair D, Hull RG. Why do general practitioners request rheumatoid factor? A study of symptoms, requesting patterns and patient outcome. Ann Clin Biochem 2003;**40**:131-37 doi: 10.1258/000456303763046049 [Published Online First: 2003/03/01].

7. Suter LG, Fraenkel L, Holmboe ES. What factors account for referral delays for patients with suspected rheumatoid arthritis? Arthritis Rheum 2006;**55**(2):300-5 doi: 10.1002/art.21855 [Published Online First: 2006/04/04].

8. Nishimura K, Sugiyama D, Kogata Y, et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. Ann Intern Med 2007;**146**(11):797-808 doi: 10.7326/0003-4819-146-11-200706050-00008 [Published Online First: 2007/06/06].

9. 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):2569-81 doi: 10.1002/art.27584 [Published Online First: 2010/09/28].

10. Thomas SL, Adebajo A, Chapel H, Webley M. The use of rheumatoid factors in clinical practice. Postgrad Med 1995;**71**:674-77 doi: 10.1136/pgmj.71.841.674 [Published Online First: 1995/11/01].

11. Miller A, Mahtani KR, Waterfield MA, Timms A, Misbah SA, Luqmani RA. Is rheumatoid factor useful in primary care? A retrospective cross-sectional study. Clin Rheumatol 2013;**32**(7):1089-93 doi: 10.1007/s10067-013-2236-0 [Published Online First: 2013/03/22].

12. Campbell J, Dedman DJ, Eaton SC, Gallagher AM, Williams TJ. Is the CPRD GOLD population comparable to the UK population? Pharmacoepidemiol Drug Saf 2013;**21**(Suppl):280 doi: 10.1002/pds.3512 [Published Online First: 2013/10/07].

13. Sammon CJ, Miller A, Mahtani KR, Holt TA, McHugh N, Luqmani R, et al. Laboratory test data in electronic general practice records are 'missing not at random': is it possible to identify when they aren't? Pharmacoepidemiol Drug Saf. 2015;In Press.14. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;**338**:b2393 doi: 10.1136/bmj.b2393 [Published Online First: 2009/07/01].

15. Marston L, Carpenter JR, Walters KR, Morris RW, Nazareth I, Petersen I. Issues in multiple imputation of missing data for large general practice clinical databases. Pharmacoepidemiol Drug Saf 2010;**19**(6):618-26 doi: 10.1002/pds.1934 [Published Online First: 2010/03/23].

16. Thomas SL, Edwards CJ, Smeeth L, Cooper C, Hall AJ. How accurate are diagnoses for rheumatoid arthritis and juvenile idiopathic arthritis in the general practice research database? Arthritis Rheum 2008;**59**(9):1314-21 doi: 10.1002/art.24015 [Published Online First: 2008/09/02].

17. Mulnier HE, Seaman HE, Raleigh VS, Soedamah-Muthu SS, Colhoun HM, Lawrenson RA. Mortality in people with type 2 diabetes in the UK. Diabet Med 2006;**23**(5):516-21 doi: 10.1111/j.1464-5491.2006.01838.x [Published Online First: 2006/05/10].

18. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;**159**(7):702-6 doi: 10.1093/aje/kwh090 [Published Online First: 2004/03/23].

19. Hosmer DW, Lemeshow S. *Applied logistic regression*. New York: John Wiley & Sons, 1989.

20. Sokka T, Pincus T. Erythrocyte sedimentation rate, C-reactive protein, or rheumatoid factor are normal at presentation in 35%-45% of patients with rheumatoid arthritis seen between 1980 and 2004:

analyses from Finland and the United States. J Rheumatol 2009;**36**(7):1387-90 doi: 10.3899/jrheum.080770 [Published Online First: 2009/05/05].

21. Whiting PF, Smidt N, Sterne JA, et al. Systematic review: accuracy of anti-citrullinated Peptide antibodies for diagnosing rheumatoid arthritis. Ann Intern Med 2010;**152**(7):456-64; W155-66 doi: 10.7326/0003-4819-152-7-201004060-00010 [Published Online First: 2010/04/07].
22. Liang KP, Kremers HM, Crowson CS, et al. Autoantibodies and the risk of cardiovascular events. J Rheumatol 2009;**36**(11):2462-9 doi: 10.3899/jrheum.090188 [Published Online First: 2009/10/17].
23. Jaeschke R, Guyatt C, Lijmer JG. Diagnostic tests. In: Guyatt G, Rennie D, eds. User's guide to the medical literature. Chicago: AMA Press, 2002:121-40.
24. Avouac J, Gossec L, Dougados M. Diagnostic and predictive value of anti-cyclic citrullinated protein antibodies in rheumatoid arthritis: a systematic literature review. Ann Rheum Dis 2006;**65**(7):845-51 doi: 10.1136/ard.2006.051391 [Published Online First: 2006/04/12].
25. Steuer A, Watkins J, Smith F, Day L, Demetriadi F, Chapel H. RF latex and anti-CCP antibodies: a combined strategy for diagnosing RA in primary care? Rheumatology (Oxford) 2008;**47**(3):375-6 doi: 10.1093/rheumatology/kem367 [Published Online First: 2008/02/01].
26. Chatfield SM, Wicks IP, Sturgess AD, Roberts LJ. Anti-citrullinated peptide antibody: death of the rheumatoid factor? Med J Aust 2009;**190**(12):693-5 doi: cha10993_fm [pii] [Published Online First: 2009/06/17].

Table 1. Characteristics of patients with positive and negative RF tests

Characteristic	**Negative RF test n (%)**	**Positive RF test n (%)**	**RR (CI 95%) or p-value**	**$_{adj}$RR***
Males	18 318 (92.1)	1573 (7.9)	1.0	1.0
Females	39 421 (92.7)	3124 (7.3)	**0.9 (0.9,1.0)**	1.0 (1.0, 1.1)
Mean age (SD)	51.4 (15.6)	57.5 (15.4)	**1.0 (1.0, 1.0)**	1.0 (1.0, 1.0)
Pre-existing thyroid disease				
No	52 923 (91.6)	4275 (92.1)	1.0	1.0
Yes	4819 (8.4)	419 (8.9)	1.1 (1.0,1.2)	0.9 (0.8, 1.0)
Cancer within 1 year before RF test				
No	57 227 (99.1)	4640 (98.9)	1.0	1.0
Yes	515 (0.9)	54 (1.2)	1.3 (1.0, 1.6)	0.7 (1.2)
Pre-existing diabetes mellitus				
None	54 629 (94.6)	4390 (93.5)	1.0	1.0
Yes (Type 1 or Type 2)	3113 (5.4)	304 (6.5)	**1.2 (1.1, 1.3)**	1.0 (0.9, 1.1)
Smoking status				
Non-smoker	35 199 (61.0)	2733 (58.2)	1.0	1.0
Current smoker	10 095 (17.5)	913 (19.5)	**1.2 (1.1, 1.2)**	**1.3 (1.2, 1.4)**
Smoking status unknown	12 448 (21.6)	1048 (22.3)	**1.1 (1.0, 1.2)**	**1.3 (1.1, 1.5)**
Body Mass Index				
<20	2321 (4.0)	228 (4.9)	**1.3 (1.1, 1.4)**	**1.4 (1.2, 1.6)**
20 - 24.9	14 166 (24.5)	1081 (23.0)	1.0	1.0
25 - 29.9	15 052 (26.1)	1260 (26.8)	**1.1 (1.0, 1.2)**	1.0 (1.0, 1.1)
30 - 34.9	7345 (12.7)	595 (12.7)	1.1 (1.0, 1.2)	1.0 (0.9, 1.1)
35 – 39.9	2593 (4.5)	190 (4.1)	1.0 (0.8, 1.1)	1.0 (0.8, 1.1)
40 and over	1109 (1.9)	94 (2.0)	1.1 (0.9, 1.4)	1.2 (1.0, 1.4)
BMI unknown	15 153 (26.2)	1249 (26.6)	1.2 (0.9, 1.4)	1.1 (0.9, 1.4)
General Practice IMD quintile				

1 (least deprived)	11 034 (19.1)	840 (17.9)	1.0	1.0
2	8957 (15.5)	673 (14.3)	1.0 (0.9, 1.1)	0.9 (0.8, 1.1)
3	11 169 (19.3)	943 (20.1)	**1.1 (1.0, 1.2)**	**1.1 (1.0, 1.2)**
4	12 137 (21.0)	1009 (21.5)	1.1 (1.0, 1.2)	1.1 (1.0, 1.2)
5 (most deprived)	14 445 (25.0)	1229 (26.2)	**1.1 (1.0, 1.2)**	**1.2 (1.1, 1.3)**

* Adjusted for age and all other variables in the table. Results shown in **bold** are statistically significant. IMD, Index of Multiple Deprivation.

Table 2. Diagnostic accuracy of RF test (a) including all RA cases and (b) excluding 'probable' cases of RA (c) using all RA cases but classifying tests where the RF titre was equal to the upper normal range for the RF test as positive rather than negative and (d) including all original data (including data considered to be MNAR).

	Primary analysis (a)		**Sensitivity analysis (b)**		**Sensitivity analysis (c)**		**Sensitivity analysis (d)**	
	RA diagnosis	**No RA diagnosis**	**RA diagnosis**	**No RA diagnosis**	**RA diagnosis**	**No RA diagnosis**	**RA diagnosis**	**No RA diagnosis**
RF test positive	1005(1.6%)	3692(5.9%)	745(1.2%)	3949(6.3%)	1096(1.8%)	12 533(20.1%)	2184(2.4%)	15 663(17.2%)
RF test negative	735(1.2%)	57 004(91.3%)	540(0.9%)	57 202(91.6%)	644(1.0%)	48 163(77.1%)	1047(1.1%)	72 399(79.3%)
Diagnostic Accuracy	**Estimate**	**95% CI**	**Estimate**	**95% CI**	**Estimate**	**95% CI**	**Estimate**	**95% CI**
LR+	9.5	9.0 to 10.0	9.0	8.5 to 9.5	3.1	2.9 to 3.2	3.8	3.7 to 3.9
LR-	0.5	0.4 to 0.5	0.5	0.4 to 0.5	0.5	0.4 to 0.5	0.4	0.4 to 0.4
Sensitivity	57.8%	55.4% to 60.0%	58.0%	55.2% to 60.7%	63.0%	60.7% to 79.7%	67.6%	66.0% to 69.2%
Specificity	93.9%	93.7% to 94.1%	93.5%	93.3% to 93.7%	79.4%	79.0% to 79.7%	98.6%	98.5% to 98.6%
PPV	21.4%	20.3% to 22.6%	15.9%	14.8% to 17.0%	8.0%	7.6% to 8.5%	12.2%	11.9% to 12.5%
NPV	98.7%	98.6% to 98.8%	99.1%	99.0% to 99.1%	98.7%	98.6% to 98.8%	98.6%	98.5% to 98.6%

RA, rheumatoid arthritis; LR+, positive likelihood ratio; LR- negative likelihood ratio; PPV, positive predictive value; NPV, negative predictive value.

Table 3. Risk of RA associated with the difference between the observed RF value and the upper normal range value provided by the laboratory

Difference between RF titre and UNR value	Rheumatoid Arthritis within 24 months		$_{adj}$RR (CI 95%)
	No n (%)	Yes n (%)	
RF value less than UNR	48 163 (79.4)	644 (37.0)	1.0
RF value equal to UNR	8844 (14.6)	91 (5.2)	0.8 (0.6 to 1.0)
1-99 IU/L above UNR	3245 (5.4)	555 (31.9)	10.0 (9.0 to 11.2)
100-499 IU/L above UNR	382 (0.6)	354 (20.3)	28.5 (25.3 to 32.0)
≥500 IU/L above UNR	62 (0.1)	96 (5.5)	33.2 (28.0 to 39.3)

RF, rheumatoid factor; RA rheumatoid arthritis; UNR, laboratory upper normal range; $_{adj}$RR, relative risk adjusted for adjusted for age, sex, smoking status, BMI, symptoms in the six months before RF test(hand symptoms, shoulder symptoms, knee symptoms, tenosynovitis, pre-existing thyroid disease, back symptoms, neck symptoms), and general practice IMD score.

Table 4. Diagnostic accuracy of RF compared with other studies

Estimate of diagnostic accuracy	Current study	Nishimura et al meta-analysis (50 studies)[8]	Whiting et al Summary results (67 studies)[21]	Whiting et al Summary results from only early arthritis studies [21]
	Estimate (95% confidence interval)			
LR+	9.5 (CI9.0 to 10.0)	4.86 (3.95 to 5.97)	3.3 (2.7 to 3.9)	4 (2.5 to 6.5)
LR-	0.5 (0.4 to 0.5)	0.38 (0.33 to 0.44)	0.4 (0.35 to 0.42)	0.5(0.45 to 0.59)
Sensitivity	57.8% (55.4% to 60.0%)	69% (65% to 73%)	70% (66% to 73%),	56% (50% to 62%)
Specificity	93.9% (93.7% to 94.1%)	85% (82% to 88%)	79% (74% to 83%),	86% (78% to 92%)
PPV	21.4% (20.3% to 22.6%)			
NPV	98.7% (98.6% to 98.8%)			

Figure 1. Kaplan-Meier failure curve for contact with rheumatology in the two years following first RF test, stratified by RF test result

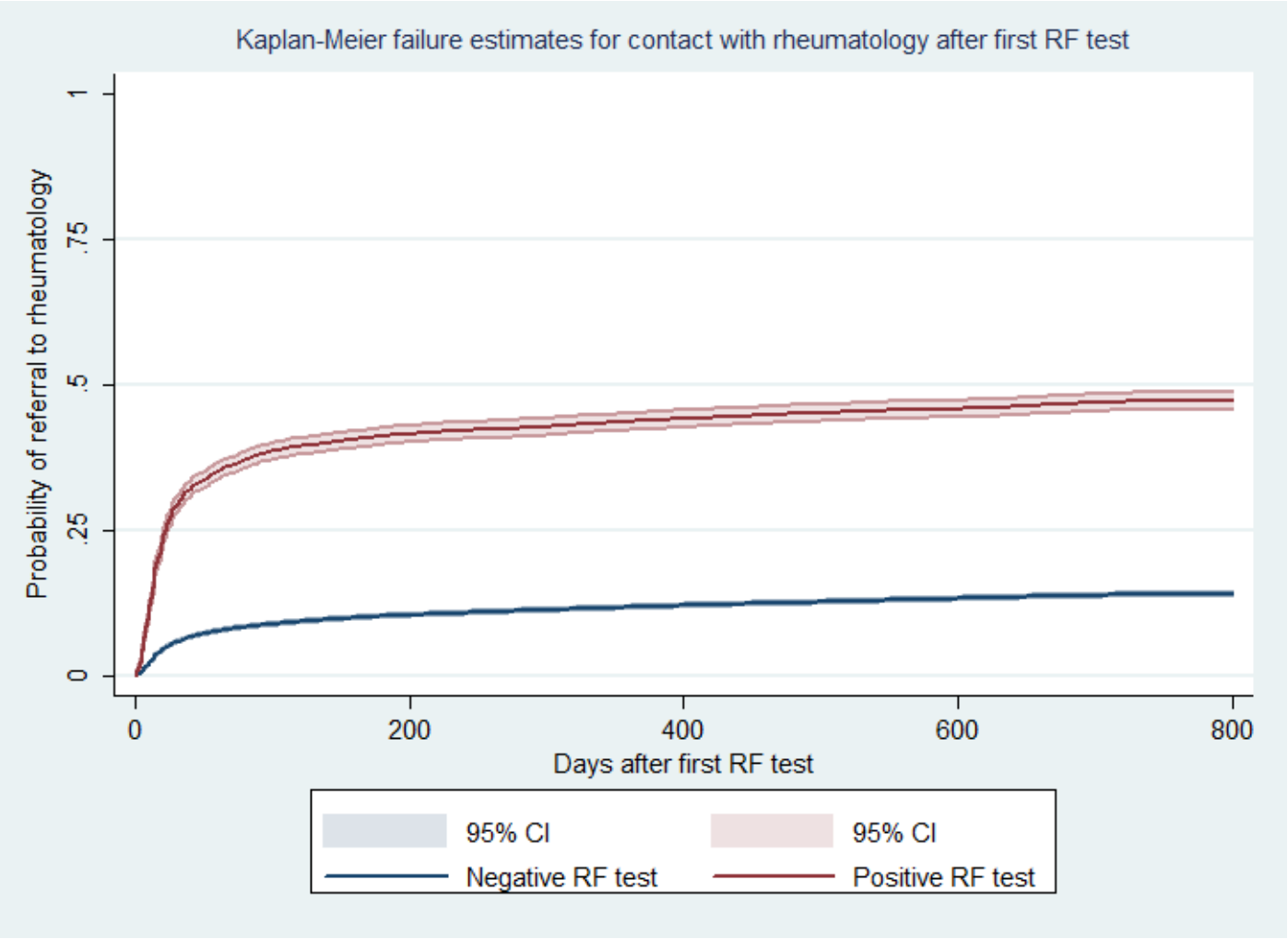

Figure 2. Kaplan-Meier failure curve for contact with any hospital department in the two years following first RF test, stratified by RF test result