

Original article

Rheumatoid arthritis-specific cardiovascular risk scores are not superior to general risk scores: a validation analysis of patients from seven countries

Cynthia S. Crowson¹, Sherine E. Gabriel², Anne Grete Semb³, Piet L. C. M. van Riel⁴, George Karpouzas⁵, Patrick H. Dessein^{6,7}, Carol Hitchon⁸, Virginia Pascual-Ramos⁹ and George D. Kitas¹⁰ on behalf of A Trans-Atlantic Cardiovascular Consortium for Rheumatoid Arthritis

Abstract

Objectives. Cardiovascular disease (CVD) risk calculators developed for the general population do not accurately predict CVD events in patients with RA. We sought to externally validate risk calculators recommended for use in patients with RA including the EULAR 1.5 multiplier, the Expanded Cardiovascular Risk Prediction Score for RA (ERS-RA) and QRISK2.

Methods. Seven RA cohorts from UK, Norway, Netherlands, USA, South Africa, Canada and Mexico were combined. Data on baseline CVD risk factors, RA characteristics and CVD outcomes (including myocardial infarction, ischaemic stroke and cardiovascular death) were collected using standardized definitions. Performance of QRISK2, EULAR multiplier and ERS-RA was compared with other risk calculators [American College of Cardiology/American Heart Association (ACC/AHA), Framingham Adult Treatment Panel III Framingham risk score-Adult Treatment Panel (FRS-ATP) and Reynolds Risk Score] using c-statistics and net reclassification index.

Results. Among 1796 RA patients without prior CVD [mean (s.d.) age: 54.0 (14.0) years, 74% female], 100 developed CVD events during a mean follow-up of 6.9 years (12430 person-years). Estimated CVD risk by ERS-RA [mean (s.d.) 8.8% (9.8%)] was comparable to FRS-ATP [mean (s.d.) 9.1% (8.3%)] and Reynolds [mean (s.d.) 9.2% (12.2%)], but lower than ACC/AHA [mean (s.d.) 9.8% (12.1%)]. QRISK2 substantially overestimated risk [mean (s.d.) 15.5% (13.9%)]. Discrimination was not improved for ERS-RA (c-statistic = 0.69), QRISK2 or EULAR multiplier applied to ACC/AHA compared with ACC/AHA (c-statistic = 0.72 for all) or for FRS-ATP (c-statistic = 0.75). The net reclassification index for ERS-RA was low (−0.8% vs ACC/AHA and 2.3% vs FRS-ATP).

Conclusion. The QRISK2, EULAR multiplier and ERS-RA algorithms did not predict CVD risk more accurately in patients with RA than CVD risk calculators developed for the general population.

Key words: rheumatoid arthritis, cardiovascular disease, risk prediction, risk assessment

¹Department of Health Sciences Research and Department of Medicine, Mayo Clinic, Rochester, MN, ²Rutgers Robert Wood Johnson Medical School, New Brunswick, NJ, USA, ³Department of Rheumatology, Diakonhjemmet Hospital, Preventive Cardio-Rheuma Clinic, Oslo, Norway, ⁴Department of Rheumatic Diseases, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands, ⁵Division of Rheumatology, Los Angeles Biomedical Research Institute, Harbor UCLA Medical Center RHU, Torrance, CA, USA, ⁶Cardiovascular Pathophysiology and Genomics Research Unit, University of Witwatersrand, Johannesburg, South Africa, ⁷Rheumatology Division, Universitair Ziekenhuis and Vrije Universiteit, Brussels, Belgium,

⁸Department of Medicine, University of Manitoba, Winnipeg, Manitoba, Canada, ⁹Department of Immunology and Rheumatology, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, México and ¹⁰Clinical Research, Unit, Dudley Group NHS Foundation Trust, West Midlands, UK

Submitted 21 September 2016; revised version accepted 31 January 2017

Correspondence to: Cynthia S. Crowson, Mayo Clinic, 200 First St SW, Rochester, MN 55906, USA.
E-mail: crowson@mayo.edu

Rheumatology key messages

- RA-specific risk calculators did not predict cardiovascular disease more accurately than general population risk calculators.
- The expanded cardiovascular risk score for RA produced comparable or lower estimates for cardiovascular disease than other calculators.
- The EULAR multiplier only reclassified a few RA patients above treatment threshold for cardiovascular disease.

Introduction

Cardiovascular disease (CVD) risk among patients with RA is increased compared with the general population [1]. CVD risk algorithms developed for the general population are not accurate in RA patients [2, 3]. The QRISK2 calculator includes RA as a risk factor for CVD [4]. The EULAR recommendations for CVD risk management proposed a 1.5 multiplier for RA patients with at least two of: disease duration >10 years, RF and/or ACPA positivity, and presence of severe extra-articular manifestations. However, the multiplier is not based on direct evidence and it is likely to be conservative [1]. A CVD risk algorithm known as the Expanded cardiovascular Risk Score for RA [ERS-RA], recently developed and internally validated using the Consortium of Rheumatology Researchers of North America (CORRONA) registry, may provide advantages over established risk scores [5]. The ERS-RA reported improved CVD risk prediction after augmenting the traditional CV risk factors with information on RA-specific characteristics, namely RA disease activity, disability, corticosteroid use and disease duration. However, the internal validation compared base and expanded models built on the same dataset because lack of blood pressure and lipid values precluded comparison with general population risk scores.

External validation is necessary to confirm any purported improvement and to determine generalizability of these calculators to other populations. The objective of this study was to externally validate the QRISK2, EULAR 1.5 multiplier and ERS-RA using data from A Trans-Atlantic Cardiovascular Risk Consortium for RA.

Methods**Study populations**

This validation study combined cohorts of patients with RA from seven different countries (UK, Norway, Netherlands, USA, South Africa, Canada and Mexico). Other A Trans-Atlantic Cardiovascular Risk Consortium RA cohorts were excluded due to lack of information on disease activity or CVD death, which were required to evaluate the ERS-RA. Patients in each cohort were included based on physician diagnosis of RA and/or fulfilment of 1987 or 2010 ACR criteria for RA. Cohorts followed patients prospectively through study visits at regular intervals or retrospectively through medical record review. Data were collected using standardized definitions. Each cohort obtained ethics committee approval at their institution and the Mayo Clinic Institutional Review Board approved the cohort aggregation. Informed

consent or authorization for use of medical records for research was obtained from each participant.

Traditional CVD risk factors collected at baseline were: age, sex, smoking status (current, former, never), blood pressure, lipid levels (total cholesterol, high density lipoprotein cholesterol, low density lipoprotein cholesterol and triglycerides), BMI, family history of premature CVD, diabetes mellitus and hypertension. Hyperlipidaemia was defined as use of lipid-lowering medications and/or low density lipoprotein cholesterol ≥ 4.1 mmol/l (i.e. 160 mg/dl). Hypertension was defined based on physician diagnosis and/or use of anti-hypertensive medication. Diabetes mellitus was defined based on physician diagnosis and/or use of oral hypoglycaemic medications or insulin. Family history of premature CVD was defined as coronary artery disease in first-degree relatives prior to age 55 years in males and 65 years in females. The CVD outcome was defined as the first event of myocardial infarction, ischaemic stroke or CVD death.

Disease-specific factors collected at baseline included ESR, CRP, swollen and tender joint counts using 28 joints, patient and physician global assessment visual analogue scales, disease duration, HAQ disability index (HAQ-DI) and all anti-rheumatic medications including corticosteroids [6]. Serologic status was defined as positive if either RF or ACPA was reported positive in clinical tests (RF tested in 99% and ACPA tested in 95% of patients). The DAS (DAS28 ESR) was calculated from ESR, swollen and tender joint counts and physician global assessment [7]. The Clinical Disease Activity Index (CDAI) was calculated from swollen and tender joint counts and physician and patient global assessments [8].

The ERS-RA was calculated using baseline data on age, sex, current smoking status, diabetes mellitus, hypertension, hyperlipidaemia, CDAI > 10, HAQ-DI > 0.5, use of corticosteroids and RA disease duration >10 years [5]. For QRISK2, the Townsend deprivation score, atrial fibrillation and chronic kidney disease were not available in our data, so CVD risk was calculated using a modified QRISK2 algorithm excluding these variables [4]. The EULAR 1.5 multiplier is applicable to patients with two or more of the following: disease duration >10 years, RF and/or ACPA positivity, and presence of severe extra-articular manifestations. The multiplier was assessed for both the Framingham risk score included in the Adult Treatment Panel guidelines (FRS-ATP) and American College of Cardiology/American Heart Association (ACC/AHA) Pooled Cohort Equation [4, 9–11]. The Reynolds Risk Score, which includes CRP, was also calculated among cohorts with available data on family history of premature CVD [12, 13]. The outcomes for all the included risk calculators were similar, except stroke was excluded

TABLE 1 Baseline characteristics of 1796 patients with RA

Characteristic	n	Value
Age at baseline, mean (s.d.), years	1796	54.0 (14.0)
Female sex, n (%)	1796	1337 (74.4)
White race, n (%)	998	901 (90.3)
Calendar year of RA diagnosis, mean (s.d.)	1796	1997.7 (8.7)
Systolic blood pressure, mean (s.d.), mmHg	1691	139.7 (23.9)
Diastolic blood pressure, mean (s.d.), mmHg	1690	81.3 (11.7)
Total cholesterol, mean (s.d.), mmol/l	1677	5.1 (1.2)
Low density lipoprotein cholesterol, mean (s.d.), mmol/l	1664	3.1 (1.1)
High density lipoprotein, cholesterol, mean (s.d.), mmol/l	1675	1.4 (0.8)
Triglycerides, median (IQR), mmol/l	1676	1.3 (0.9–1.8)
Current smoker, n (%)	1666	420 (25.2)
Diabetes mellitus, n (%)	1796	96 (5.3)
Family history of premature coronary heart disease, n (%)	1303	349 (26.8)
BMI, mean (s.d.), kg/m ²	1663	26.4 (4.9)
Hypertension, n (%)	1795	870 (48.5)
Anti-hypertensive medication use, n (%)	1794	380 (21.2)
Hyperlipidaemia, n (%)	1686	589 (34.9)
Lipid lowering medication use, n (%)	1794	148 (8.2)
RF and/or ACPA seropositivity, n (%)	1790	1489 (83.2)
RA disease duration, mean (s.d.), years	1796	4.9 (8.1)
Duration <1 year, n (%)		1072 (59.7)
Duration 1–10 years, n (%)		324 (18.0)
Duration >10 years, n (%)		400 (22.3)
ESR, mean (s.d.), mm/h	1786	28.2 (23.2)
CRP, mean (s.d.), mg/l	1783	19.1 (31.0)
Synthetic DMARD use, n (%)	1796	639 (35.6)
Biologic DMARD use, n (%)	1789	110 (6.1)
Corticosteroid use, n (%)	1796	318 (17.7)
DAS28ESR, mean (s.d.)	1783	4.5 (1.7)
CDAI, mean (s.d.)	434	24.6 (16.2)
CDAI > 10 or DAS28ESR > 3, n (%)	1796	1468 (81.7)
HAQ-DI, mean (s.d.)	1796	0.9 (0.8)
HAQ-DI > 0.5, n (%)		1054 (58.7)

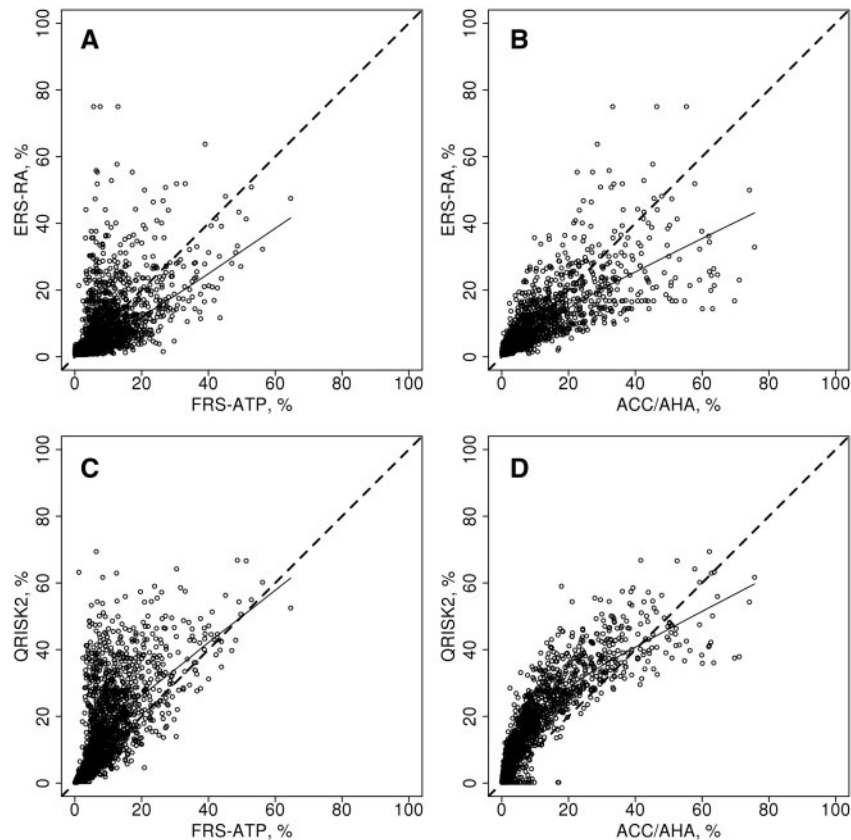
CDAI: clinical disease activity index; HAQ-DI: HAQ disability index.

from the FRS-ATP and revascularization was included in the Reynolds Risk Score. The Systematic COronary Risk Evaluation (SCORE) was not assessed because its outcome of fatal CVD events is only a subset of the outcomes used for ERS-RA, so they are not directly comparable [14].

Statistical methods

Patients with CVD prior to baseline were excluded. Multiple imputation was used to impute missing values for the CVD risk factors using 10 repetitions. Available sample size for each variable is listed in Table 1. Log-transformations were used when imputing lipid levels to avoid bias when computing lipid ratios from imputed data [15]. CDAI values were available in 434 patients. CDAI > 10 was imputed using DAS28ESR > 3.0 in 1362 patients based on a linear regression of subjects with both values ($R^2 = 0.89$, which agrees with other reports) [16]. The imputation demonstrated only a 10% misclassification rate with 90% sensitivity and 86% specificity. Sensitivity analyses on the 434 patients with available CDAI were also performed.

Evaluating performance of the risk calculators involved assessment of both calibration and discrimination. Calibration is the ability to accurately predict the absolute risk of developing CVD. Expected events were computed from predicted risk estimates and calibration was assessed by comparing observed to expected CVD events [17]. Observed risk estimates for FRS-ATP and Reynolds utilized modified CVD outcomes required for these scores. For this comparison, follow-up >10 years was truncated, and events after 10 years were excluded. For patients with <10 years of follow-up, expected events were adjusted proportionately. Standardized incidence ratios (SIR), being ratios of observed to expected events, were calculated assuming the expected rates were fixed and the observed events followed a Poisson distribution [17]. An SIR >1 indicated the observed events were higher than expected meaning the predicted risk underestimated actual risk. Conversely, SIR < 1 indicated the predicted risk overestimated actual risk. Discrimination is the ability to correctly rank risk from low to high, which was assessed using Harrell's c-statistic. Net reclassification

Fig. 1 Comparison of the Expanded cardiovascular Risk Score for RA and QRISK2 with other risk calculators

The ERS-RA is plotted against the Framingham risk score from Adult Treatment Panel III guidelines (FRS-ATP; **A**), the American College of Cardiology/American Heart Association Pooled Cohort Equation (ACC/AHA; **B**), and the QRISK2 is plotted against the FRS-ATP (**C**) and the ACC/AHA (**D**). The dashed line represents identity and the solid line is a smoother line of trend.

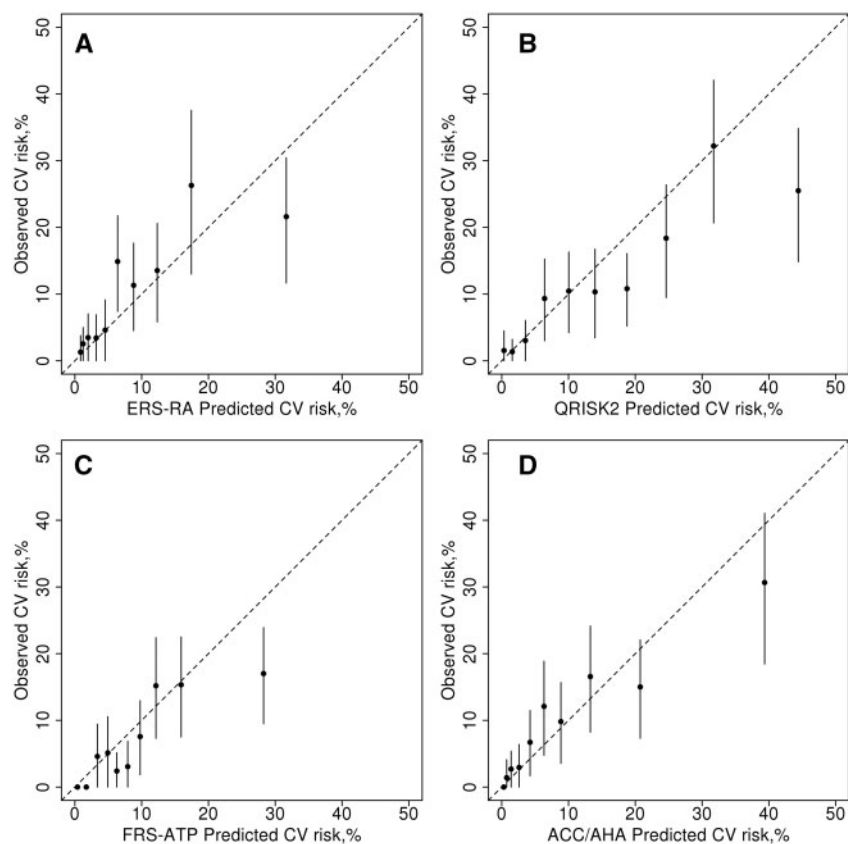
index (NRI), which compares old and new risk calculators by measuring improvement in classification of patients into risk categories, was assessed using survival methods with 95% CIs based on 1000 bootstrap samples using software obtained from <http://ncook.bwh.harvard.edu/sas-macros.html>. Risk categories representing the treatment thresholds for the FRS-ATP (i.e. <20% vs \geq 20%) and ACC/AHA (i.e. <7.5% vs \geq 7.5%) were assessed. The 10-year observed CVD event rates were calculated using Kaplan-Meier methods [18]. Analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC, USA) and R 3.1.1 (R Foundation for Statistical Computing, Vienna, Austria).

Results

In total, 1796 RA patients without prior CVD were included [mean (s.d.) age 54.0 (14.0) years, 74% female; Table 1]. During a mean follow-up of 6.9 years (12 430 person-years), 100 patients developed a CVD event. The overall incidence rate (per 1000 person-years) for first CVD event

was 8.0 (95% CI: 6.5, 9.8). By event type, the incidence rates were 6.3 (95% CI: 5.0, 7.8) for myocardial infarction, 3.1 (95% CI: 2.2, 4.2) for stroke and 1.7 (95% CI: 1.0, 2.6) for CVD death. Mean (s.d.) RA disease duration was 4.9 (8.1) years with RA disease duration >10 years in 22%. Among 434 patients with available CDAl, mean (s.d.) CDAl was 24.6 (16.2) with CDAl >10 in 80%. Mean (s.d.) DAS28ESR was 4.5 (1.7) with DAS28-ESR > 3.0 in 81% and DAS28ESR > 5.1 in 38%. Following imputation of CDAl >10 using DAS28-ESR > 3.0, 80% had imputed CDAl >10. Mean (s.d.) HAQ-DI was 0.89 (0.77) with HAQ-DI > 0.5 in 59%. As more than half of patients were enrolled at RA incidence, only 18% were using corticosteroids and 36% synthetic DMARDs at baseline.

While some patients have higher ERS-RA scores compared with other CVD risk scores, the majority of patients have lower estimated CVD risk by ERS-RA (Fig. 1A and B). Average estimated 10-year CVD risk by ERS-RA [mean (s.d.) 8.8% (9.8%)] was comparable to FRS-ATP [mean (s.d.) 9.1% (8.3%)] and Reynolds [mean (s.d.) 9.2% (12.2%)], but somewhat lower than ACC/AHA [mean

Fig. 2 Observed vs predicted 10-year risk for cardiovascular disease for each cardiovascular risk calculator

Observed vs predicted 10-year risk for cardiovascular disease (CVD) according to deciles of predicted risk obtained from the Expanded cardiovascular Risk Score for RA (ERS-RA; **A**), QRISK2 (**B**), the Framingham risk score from Adult Treatment Panel III (FRS-ATP; **C**) and the American College of Cardiology/American Heart Association Pooled Cohort Equation (ACC/AHA; **D**). The vertical lines are 95% CIs for the observed CVD risk estimates.

(s.d.) 9.8% (12.1%). QRISK2 [mean (s.d.) 15.5% (13.9%)] was higher than the other risk scores in the majority of patients (Fig. 1C and D).

Calibration was assessed by SIR. Overall the ERS-RA had a SIR of 0.84 (95% CI: 0.69, 1.02) indicating an overestimation of the observed CVD risk, which did not reach statistical significance and the QRISK2 significantly overestimated the risk of CVD (SIR: 0.45; 95% CI: 0.37, 0.55). However, the FRS-ATP (SIR: 0.63; 95% CI: 0.52, 0.79) and ACC/AHA (SIR: 0.73; 95% CI: 0.60, 0.89) both also significantly overestimated the risk of CVD in this cohort and the Reynolds risk score underestimated the CVD risk (SIR: 1.25; 95% CI: 1.05, 1.48). In fact, the ERS-RA had the SIR closest to 1, meaning it overestimated CVD risk less than the other risk calculators. When examining deciles of predicted risk, the ERS-RA significantly overestimated observed CVD risk in the highest decile, but appeared to predict risk reasonably well in the other deciles (Fig. 2A). The QRISK2 and FRS-ATP significantly overestimated risk in the highest decile and a few other deciles (Fig. 2B and C). In contrast, there were no significant

differences between predicted and observed risks by decile for ACC/AHA (Fig. 2D). Discrimination was not improved for ERS-RA (c-statistic: 0.69; 95% CI: 0.63, 0.75) and QRISK2 (c-statistic: 0.72), compared with the other risk calculators (FRS-ATP: 0.75; ACC/AHA: 0.72; Reynolds: 0.72).

The EULAR multiplier was applied to FRS-ATP and ACC/AHA, and criteria for applying the multiplier were met in 373 (21%) patients. The multiplier resulted in greater overestimation of future CVD events (SIR for FRS-ATP: 0.57, 95% CI: 0.46, 0.71 and SIR for ACC/AHA: 0.62; 95% CI: 0.51, 0.75). In addition, the c-statistics remained unchanged (0.75 and 0.72 for FRS-ATP and ACC/AHA, respectively).

Improvement of a new risk algorithm is often assessed by its ability to correctly reclassify high risk patients above the treatment threshold and low risk patients below the treatment threshold. Using the ACC/AHA 7.5% treatment threshold, the ERS-RA reclassified 118 patients without events and 6 patients with events from low to high risk and 108 patients without events and 8 patients with

TABLE 2 Reclassification of predicted 10-year risk of cardiovascular disease above and below treatment thresholds

General CVD risk scores	ERS-RA		QRISK2	
	<7.5%	≥7.5%	<7.5%	≥7.5%
ACC/AHA				
<7.5%				
N w/o events	932	118	668	382
N w/events	24	6	8	22
Observed 10-year CVD rate (95% CI)	4.1 (2.3, 5.8)	8.6 (3.7, 15.8)	2.2 (0.5, 3.8)	8.4 (4.6, 12.1)
≥7.5%				
N w/o events	108	538	12	634
N w/events	8	62	0	70
Observed 10-year CVD rate (95% CI)	9.5 (2.8, 15.8)	19.7 (14.5, 24.5)	0.0 (-, -)	17.8 (13.4, 21.9)
FRS-ATP				
<20%				
N w/o events	1440	115	1154	401
N w/events	65	10	35	40
Observed 10-year CVD rate (95% CI)	7.6 (5.6, 9.6)	18.4 (4.6, 30.2)	4.6 (2.9, 6.2)	19.0 (12.8, 25.8)
≥20%				
N w/o events	77	64	16	125
N w/events	10	15	2	23
Observed 10-year CVD rate (95% CI)	17.9 (6.2, 28.2)	28.1 (13.5, 40.2)	11.5 (0.0, 25.2)	24.0 (14.0, 32.9)

Comparing the expanded cardiovascular risk predictions score for RA (ERS-RA) and QRISK2 with American College of Cardiology/American Heart Association (ACC/AHA) Pooled Cohort Equation and Framingham risk score used in Adult Treatment Panel III guidelines (FRS-ATP) risk calculators for 1798 patients with rheumatoid arthritis. ERS-RA: expanded cardiovascular risk predictions score for RA; ACC/AHA: American College of Cardiology/American Heart Association Pooled Cohort Equation; FRS-ATP: Framingham risk score used in Adult Treatment Panel III guidelines; N: numbers; w/o: without; w/: with.

events from high to low risk. Among patients reclassified from low to high risk, the observed 10-year CVD event rate was 8.6%, which is slightly higher than 7.5% (Table 2). However, the observed 10-year CVD event rate for patients reclassified from high to low risk was 9.5%, which is not lower than 7.5%. The NRI was low (−0.8%; 95% CI: −8.2, 7.1%). Results were similar when comparing ERS-RA to FRS-ATP using the 20% treatment threshold (NRI: 2.3%, 95% CI: −8.3, 26.6%). The QRISK2 had a low NRI when compared with the ACC/AHA (−2.4%; 95% CI: −10.9, 6.5%). The NRI comparing the QRISK2 to FRS-ATP was higher than the other comparisons, but it did not reach statistical significance (NRI: 25.0%, 95% CI: −9.4, 34.7%). The EULAR multiplier only reclassified six patients above the 7.5% treatment threshold for the ACC/AHA calculator and three patients above the 20% treatment threshold for the FRS-ATP calculator, so the NRI was negligible.

Examination of the patients reclassified from low to high risk by ERS-RA revealed that this group consisted predominantly of two different types of patients. The first type included patients with few CVD risk factors and several RA-specific factors (i.e. disease duration, disease activity, disability and corticosteroid use). The second type included patients without RA-specific factors who had well-controlled hypertension and/or hyperlipidaemia, who were able to achieve lower CVD risk estimates by ACC/AHA, but not by ERS-RA. Examination of the patients reclassified from high to low risk revealed this

group consisted predominantly of patients with untreated or undiagnosed hypertension or hyperlipidaemia or those with low high-density lipoprotein cholesterol, and patients with few RA-specific factors. Some patients with one or two unfavourable RA-specific factors still had lower CVD risk estimates by ERS-RA than by ACC/AHA. Because CDAI was available in the three cohorts with the lowest CVD event rates, sensitivity analyses in this subset without imputed CDAI showed the ERS-RA substantially overestimated CVD risk (SIR: 0.32, 95% CI: 0.18, 0.61).

Each CVD risk calculator was designed to be applicable in different age ranges (ERS-RA: all adult ages; QRISK2: 35–74 years; FRS-ATP: 30–74 years; ACC/AHA: 40–79 years). In a sensitivity analysis including only patients aged 40–74 years at baseline (1388 patients with 91 first CVD events), calibration of all risk calculators improved slightly compared with the overall results, as all SIRs moved closer to 1 (ERS-RA: 0.84–0.98; QRISK2: 0.45–0.49; FRS-ATP: 0.64–0.63; and ACC/AHA: 0.73–0.93). However, discrimination worsened compared with the overall results, as the c-statistics decreased (ERS-RA: 0.69 to 0.66; QRISK2: 0.72 to 0.70; FRS-ATP: 0.75 to 0.71; and ACC/AHA: 0.72 to 0.70).

In a third sensitivity analysis, only patients with >1 year disease duration at baseline were evaluated (694 patients with 22 first CVD events and mean disease duration of 12 years). All risk calculators substantially overestimated CVD risk in this subset of patients, as all SIRs were significantly <1 [ERS-RA: 0.39 (95% CI: 0.26, 0.59); QRISK2:

0.26 (95% CI: 0.17, 0.39); FRS-ATP: 0.38 (95% CI: 0.23, 0.63); ACC/AHA: 0.41 (95% CI: 0.27, 0.62)]. Discrimination also worsened compared with the overall results, as the *c*-statistics decreased (ERS-RA: 0.69, 0.65; QRISK2: 0.72, 0.62; FRS-ATP: 0.75, 0.70; and ACC/AHA: 0.72, 0.65).

Discussion

Despite the inclusion of RA-specific risk factors, the ERS-RA, QRISK2 and EULAR 1.5 multiplier did not improve CVD risk predictions for patients with RA compared with general population CVD risk calculators. The ERS-RA algorithm produced comparable or lower risk estimates than other CVD risk calculators. While calibration was slightly improved for the ERS-RA algorithm compared with FRS-ATP or ACC/AHA, discrimination and classification were unimproved. In contrast, QRISK2 had higher risk estimates than the other risk algorithms resulting in poor calibration with no improvement in discrimination and little improvement in classification. The EULAR multiplier reduced calibration without improving discrimination or classification.

The low CVD risk in our cohort, which led to $SIR < 1$ in all assessments of calibration, was unexpected, as previous reports have found higher CVD risk in patients with RA than in the general population and have reported that risk scores for the general population underestimate risk among patients with RA [1–3]. However, our event rates are actually higher than those reported in the CORRONA cohort used to develop the ERS-RA (combined CVD event rate: 8.0 vs 6.4; myocardial infarction: 6.3 vs 2.5; stroke 3.1 vs 3.0; CVD death: 1.7 vs 1.0/1000 person-years in our cohort and CORRONA, respectively) [5]. Furthermore, all centres re-reviewed their data and found no evidence for under-ascertainment of CVD events. However, CVD event rates may differ in the general populations of the countries included, which could impact our results. Another possible explanation is that lower CVD event rates in recent years may be due to variability in aggressiveness of RA and CVD risk factor management. While this may affect our assessment of calibration, it has little impact on the relative comparisons between the risk calculators, which are the focus of our analyses.

Examination of patients reclassified by ERS-RA revealed limitations related to its reliance on physician diagnoses of hypertension and hyperlipidaemia, which are known to be under-diagnosed and under-treated in patients with RA [19, 20]. Hyperlipidaemia is a non-specific, general term which is not defined in the latest ACC/AHA, European Atherosclerosis Society or European Society of Cardiology guidelines for CVD prevention [11, 21, 22]. The lack of consensus criteria for hyperlipidaemia was worrisome; however, our cohort had higher hyperlipidaemia rates than the ERS-RA derivation cohort (35% vs 15%). Another limitation of ERS-RA was that successful treatment for hypertension and/or hyperlipidaemia did not result in reduced CVD risk estimates.

The poor calibration of QRISK2 may be related to our inclusion of cohorts from multiple countries, which may have different underlying CVD rates than the UK. Also,

the modified version of the QRISK2 without Townsend deprivation score, atrial fibrillation and chronic kidney disease may have poorer performance. Furthermore, QRISK2 is arguable as an RA-specific risk calculator since it includes an RA indicator, but not RA-specific characteristics [5]. The EULAR multiplier was only applicable in 21% of our patients and only a handful of patients were actually reclassified above the treatment thresholds for FRS-ATP and ACC/AHA, so the practical implications of the multiplier are extremely limited, which confirmed its conservativeness [1].

In contrast to general population CVD risk scores, which are developed using large population-based cohorts with long follow-up, these RA risk calculators may be less reliable due to constraints related to data availability. Large population-based cohorts of patients with RA with long follow-up are lacking, partially due to the low prevalence of RA, which is around 1% of the population. Even in the QRISK2, which is a general population CVD risk calculator developed using a large cohort of patients, the number of patients with RA upon which the RA risk estimate is based was likely small. The ERS-RA was developed using registry data and the mean follow-up was <3 years, which may lead to inaccuracies in estimation of 10-year CVD risk. The EULAR multiplier was developed based on expert opinion and evidence from the literature and its performance was not formally assessed by the developers.

Furthermore, the utility of RA-specific CVD risk calculators is debatable. Minimizing disease activity is the goal of disease management in patients with RA, and there is no evidence that even more aggressive therapy specifically to target CVD risk, instead of joint disease, would be of benefit [23]. Evidence that traditional risk factors should be managed differently in patients with RA is also lacking. Hence, while accurately predicting CVD risk in patients with RA is important, the clinical utility of an RA-specific risk calculator is questionable. Further research is needed to determine the optimal strategies to lower CVD risk in patients with high CVD risk due to RA disease activity.

Strengths of our study include the use of many well-established cohorts assembled by prominent researchers. While inclusion criteria varied between centres, all RA diagnoses were confirmed by experienced rheumatologists. Furthermore, the cohorts were either population-based or included all consecutive patients, minimizing selection bias issues. However, referral bias cannot be excluded due to the inclusion of referral centres, and these data from well-organized academic centres may not be generalizable to other practice settings. Other study limitations include possible differences in measurement of CVD risk factors and events, despite use of standardized variable definitions. Cohort surveillance methods differed as some cohorts evaluated their patients at regular intervals and others retrospectively reviewed medical records. In addition, CVD events were not adjudicated and some centres contributed few CVD events. However, inconsistencies in risk factor definitions likely mimic the real-world application of these calculators, as physicians applying them are unlikely to ensure their

measurement methods exactly agree with those used in algorithm development. Furthermore, imputation of CDAI > 10 using DAS28ESR > 3.0 could be questioned, as high disease activity is typically defined as DAS28ESR > 5.1 [24]. However, using a higher cut-point would result in lower ERS-RA estimates, which were already lower than other risk calculator estimates. So increasing the DAS28ESR cut-point would further decrease the performance of ERS-RA. While mean follow-up was only 7 years, this was similar to the QRISK2 derivation cohort [4]. Finally, the RA-specific SCORE adaptation could not be evaluated, because our cohort included the adaptation derivation cohort, and because the fatal CVD outcome used in SCORE obfuscates direct comparisons [25].

In conclusion, RA-specific risk calculators did not predict CVD risk in patients with RA more accurately than the general population risk calculators. Thus, based on our assessment, it does not appear that these RA-specific CVD risk calculators are clinically useful. However, further validation work is warranted to determine whether these RA-specific calculators may be useful in subsets of patients with RA or certain populations. Furthermore, the issue of accurate CVD risk assessment for patients with RA still needs to be addressed, so more research is needed to build an accurate CVD risk calculator for patients with RA.

Acknowledgements

Additional ATACC-RA members: Karen Douglas and Aamer Sandoo, Dudley Group NHS Foundation Trust, West Midlands, UK; Silvia Rollefstad, Eirik Ikdahl and Tore K. Kvien, Diakonhjemmet hospital, Oslo, Norway; Elke Arts and Jaap Fransen, Radboud University, Nijmegen, Netherlands; Linda Tsang, Cardiovascular Pathophysiology and Genomics Research Unit, School of Physiology, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa; Hani El-Gabalawy, University of Manitoba, Winnipeg, Manitoba, Canada; Irazú Contreras Yáñez, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubir, Mexico City, México; Eric L Matteson, Mayo Clinic, Rochester, MN, USA; Solbritt Rantapää-Dahlqvist, Solveig Wållberg-Jonsson and Lena Innala, University Hospital, Umeå, Sweden; Petros P. Sfikakis and Evi Zampeli, University of Athens, Athens, Greece; Miguel Gonzalez-Gay and Alfonso Corrales, Hospital Universitario Marques de Valdecilla, Santander (Cantabria), Spain; Mart van de Laar, Harald Vonkeman and Inger Meek, Hospital Medisch Spectrum Twente, Enschede, Netherlands; Elaine Husni and Robert Overman, Cleveland Clinic, Cleveland, Ohio, USA; Iris Colunga and Dionicio Galarza, Hospital Universitario 'Dr José E. González', Monterrey, México.

Funding: This work was supported by a collaborative agreement for independent research from Eli Lilly and a grant from the National Institutes of Health, NIAMS (R01AR046849). The content is solely the responsibility

of the authors and does not necessarily represent the official views of the National Institutes of Health.

Disclosure statement: G.K. has received honoraria from Merke, Sharp and Dohme, Union Chimique Belge, Abbvie, Pfizer, Bristol-Myers Squibb and Novartis. P.v.R. has received grants and honoraria received from BMS, Abbvie, Pfizer and Roche. A.G.S. has received honoraria as speakers' bureau Merck/Schering-Plough, UCB, Abbott, Pfizer/Wyeth, BMS and Eli-Lilly. All other authors have declared no conflicts of interest.

References

- Peters MJ, Symmons DP, McCarey D *et al.* EULAR evidence-based recommendations for cardiovascular risk management in patients with rheumatoid arthritis and other forms of inflammatory arthritis. *Ann Rheum Dis* 2010;69:325–31.
- Crowson CS, Matteson EL, Roger VL, Thorneau TM, Gabriel SE. Usefulness of risk scores to estimate the risk of cardiovascular disease in patients with rheumatoid arthritis. *Am J Cardiol* 2012;110:420–4.
- Arts EE, Popa C, Den Broeder AA *et al.* Performance of four current risk algorithms in predicting cardiovascular events in patients with early rheumatoid arthritis. *Ann Rheum Dis* 2015;74:668–74.
- Hippisley-Cox J, Coupland C, Vinogradova Y *et al.* Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* 2008;336:1475–82.
- Solomon DH, Greenberg J, Curtis JR *et al.* Derivation and internal validation of an expanded cardiovascular risk prediction score for rheumatoid arthritis: A Consortium of Rheumatology Researchers of North America Registry Study. *Arthritis Rheumatol* 2015;67:1995–2003.
- Bruce B, Fries JF. The Stanford Health Assessment Questionnaire: a review of its history, issues, progress, and documentation. *J Rheumatol* 2003;30:167–78.
- Prevoo ML, van 't Hof MA, Kuper HH *et al.* Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44–8.
- Aletaha D, Nell VP, Stamm T *et al.* Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005;7:R796–806.
- Wilson PW, D'Agostino RB, Levy D *et al.* Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97:1837–47.
- Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation* 2002;106:3143–421.
- Goff DC Jr, Lloyd-Jones DM, Bennett G *et al.* 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2014;63:2935–59.

- 12 Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. *JAMA* 2007;297:611–9.
- 13 Ridker PM, Paynter NP, Rifai N, Gaziano JM, Cook NR. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. *Circulation* 2008;118:2243–51.
- 14 Conroy RM, Pyorala K, Fitzgerald AP *et al.* Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003;24:987–1003.
- 15 Morris TP, White IR, Royston P, Seaman SR, Wood AM. Multiple imputation for an incomplete covariate that is a ratio. *Stat Med* 2014;33:88–104.
- 16 Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005;23:S100–8.
- 17 Crowson CS, Atkinson EJ, Therneau TM. Assessing calibration of prognostic risk scores. *Stat Methods Med Res* 2016;25:1692–706.
- 18 Pencina MJ, D'Agostino RB Sr, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011;30:11–21.
- 19 Panoulas VF, Douglas KM, Milionis HJ *et al.* Prevalence and associations of hypertension and its control in patients with rheumatoid arthritis. *Rheumatology* 2007;46:1477–82.
- 20 Toms TE, Panoulas VF, Douglas KM *et al.* Statin use in rheumatoid arthritis in relation to actual cardiovascular risk: evidence for substantial undertreatment of lipid-associated cardiovascular risk? *Ann Rheum Dis* 2010;69:683–8.
- 21 Reiner Z, Catapano AL, De Backer G *et al.* ESC/EAS Guidelines for the management of dyslipidaemias: the Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS). *Eur Heart J* 2011;32:1769–818.
- 22 Perk J, De Backer G, Gohlke H *et al.* European Guidelines on cardiovascular disease prevention in clinical practice (version 2012). The Fifth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of nine societies and by invited experts). *Eur Heart J* 2012;33:1635–701.
- 23 Symmons DP. Editorial: do we need a disease-specific cardiovascular risk calculator for patients with rheumatoid arthritis? *Arthritis Rheumatol* 2015;67:1990–4.
- 24 Fransen J, van Riel PL. The Disease Activity Score and the EULAR response criteria. *Clin Exp Rheumatol* 2005;23:S93–9.
- 25 Arts EE, Popa CD, Den Broeder AA *et al.* Prediction of cardiovascular risk in rheumatoid arthritis: performance of original and adapted SCORE algorithms. *Ann Rheum Dis* 2016;75:674–80.