



Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

Baseline patient reported outcomes are more consistent predictors of long-term functional disability than laboratory, imaging or joint count data in patients with early inflammatory arthritis: A systematic review



James M. Gwinnutt, PhD^a, Charlotte A. Sharp, MBChB, MRCP^b,
Deborah P.M. Symmons, MD, FFPH, FRCP^a, Mark Lunt, PhD^a,
Suzanne M.M. Verstappen, PhD^{a,c,*}

^a Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

^b National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care (NIHR CLAHRC) Greater Manchester, Greater Manchester, Alliance Manchester Business School, The University of Manchester, UK

^c NIHR Manchester Biomedical Research Centre, Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, UK

ARTICLE INFO

Keywords:
Early rheumatoid arthritis
Long-term outcome
Functional disability
Systematic review

ABSTRACT

Objective: To assess baseline predictors of long-term functional disability in patients with inflammatory arthritis (IA).

Methods: We conducted a systematic review of the literature from 1990 to 2017 using MEDLINE and EMBASE. Studies were included if (i) they were prospective observational studies, (ii) all patients had IA with symptom duration ≤ 2 years at baseline, (iii) follow-up was at least 5 years, and (iv) baseline predictors of HAQ score at long-term follow-up (i.e., ≥ 5 years following baseline) were assessed. Information on the included studies and estimates of the association between baseline variables and long-term HAQ scores were extracted from the full manuscripts.

Results: Of 1037 abstracts identified by the search strategy, 37 met the inclusion/exclusion criteria and were included in the review. Older age at baseline and female gender were reported to be associated with higher long-term HAQ scores in the majority of studies assessing these relationships, as were higher baseline HAQ and greater pain scores (total patients included in analyses reporting significant associations/total number of patients analysed: age 9.8k/10.7k (91.6%); gender 9.9k/11.3k (87.4%); HAQ 4.0k/4.0k (99.0%); pain 2.8k/2.9k (93.6%)). Tender joint count, erythrocyte sedimentation rate (ESR) and DAS28 were also reported to predict long-term HAQ score; other disease activity measures were less consistent (tender joints 2.1k/2.5k (84.5%); erythrocyte sedimentation rate 1.6k/2.2k (72.3%); DAS28 888/1.1k (79.2%); swollen joints 684/2.6k (26.6%); C-reactive protein 279/510 (54.7%)). Rheumatoid factor (RF) and erosions were not useful predictors (RF 546/4.6k (11.9%); erosions 191/2.7k (7.0%)), whereas the results for anti-citrullinated protein antibody positivity were equivocal (ACPA 2.0k/3.8k (52.9%)).

Conclusions: Baseline age, gender, HAQ and pain scores are associated with long-term disability and knowledge of these may aid the assessment of prognosis.

© 2018 The Authors. Published by Elsevier HS Journals, Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Funding: James Gwinnutt is supported by the Arthritis Research UK Centre for Epidemiology grant [20380] and [21229].

Charlotte A Sharp is supported by the National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care (NIHR CLAHRC) Greater Manchester. The views expressed are those of the authors and not necessarily those of the NIHR, the NHS or the Department of Health.

This research was supported by researchers at the NIHR Manchester Biomedical Research Centre.

Competing interests: none.

*Corresponding author at: Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. Tel.: +44 01612755663.

E-mail address: suzanne.verstappen@manchester.ac.uk (S.M.M. Verstappen).

<https://doi.org/10.1016/j.semarthrit.2018.03.004>

0049-0172/© 2018 The Authors. Published by Elsevier HS Journals, Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Inflammatory arthritis (IA), and its subset rheumatoid arthritis (RA), are chronic conditions characterised by synovial joint inflammation [1]. Negative outcomes associated with these conditions include premature mortality [2,3], joint destruction [4,5], and functional disability [6–8]. The term functional disability refers to the difficulties patients with IA have in performing everyday tasks. Preventing or minimising functional disability is a key goal in IA management.

In the past, functional disability was assessed using the Steinbrocker Functional Class system, in which the physician scored the

patient from class 1 (indicating little or no disability), to class 4 (indicating patients were bed-ridden or confined to a wheel chair) [9]. Whilst this system was quick and reflected clinicians' judgement, only having four levels of disability meant the measure was insensitive to change [10]. Later the Health Assessment Questionnaire – Disability Index (HAQ) was developed [11]. The HAQ comprises 20 questions in eight subsections assessing different aspects of everyday life, yielding a score of 0–3, with 0 indicating no disability and 3 representing substantial levels of disability. The HAQ has become the gold standard for measuring disability in patients with IA and has been shown to be a valid measure of disability [12,13]. A minimum clinically important difference was estimated to be between 0.20 and 0.22 [14], although later estimates have put the value as low as 0.09 within an observational cohort setting [15].

Longitudinally, functional disability measured using the HAQ has been shown to follow a J-shaped trajectory, with initial improvements in disability one to two years following symptom onset, followed by increasing HAQ scores over the subsequent 5–10 years [6]. Being able to predict which patients are likely to develop major problems in performing daily tasks is useful for patients and clinicians. Clinicians can target patients susceptible to high levels of long-term disability to receive additional interventions alongside their pharmacological therapy. Patients too may be able to modify their lifestyle to reduce future disability. A systematic review of predictors of HAQ score in patients with RA was published in 2003 [16]. A further literature review was published in 2010 including studies with patients with a range of disease durations at baseline (< 1 to 12 years) and follow-up lengths (1–15 years) [17]. However, the latter was not a systematic review and since 2003 a number of additional manuscripts investigating predictors of functional disability have been published. Furthermore, due to the J-shaped trajectory of functional disability, baseline predictors of short term (i.e., between 0 and 5 years) HAQ score may not be the same as predictors of long-term (i.e., ≥5 years) HAQ. Therefore, it is important to consider predictors of long-term functional disability separately from predictors of short-term functional disability, as measured by the HAQ.

The aim of this systematic review was to critically evaluate the available literature on baseline predictors of long-term (i.e., ≥5 years) functional disability in patients with early IA.

Methods

To address these aims, we performed a systematic review using the MEDLINE and EMBASE databases, including studies published between 01/01/1990 and 05/10/2017. The inclusion criteria were (i) all patients had IA (≥2 swollen joints lasting for ≥4 weeks), RA (defined as meeting any of the published criteria sets [18–20]), or undifferentiated arthritis; (ii) all patients had less than or equal to two years symptom duration at baseline; (iii) analysis had to assess baseline predictors of long-term functional disability measured using the HAQ at ≥5 years following baseline; (iv) studies had to be observational; (v) studies published in English (or a translation available). Exclusion criteria were (i) randomised controlled trials, clinical trials, cross-sectional studies or case-series; (ii) studies including children; (iii) studies including non-human animals; (iv) conference abstracts. The study was designed and reported according to PRISMA guidelines [21].

A search strategy was devised which included both text words and MESH terms (Supplementary file 1). This search strategy yielded 1037 titles and abstracts, 532 from MEDLINE and 505 from EMBASE. Of these 263 were identified as duplicates by reference managing software (Endnote) and were removed.

Each of the remaining titles and abstracts was independently screened based on the inclusion and exclusion criteria by two reviewers using a standardized form (JG and CS). In case of any discrepancies in agreement between the two reviewers ($n = 53$) a third reviewer was consulted (SV). Of 774 titles and abstracts screened, 73 met the inclusion criteria and the full manuscript was read by the same reviewers. Of these, 33 papers were included in the review. The reference lists of these manuscripts were screened. Four additional studies were added to the review, meaning a total of 37 studies were included (Fig. 1).

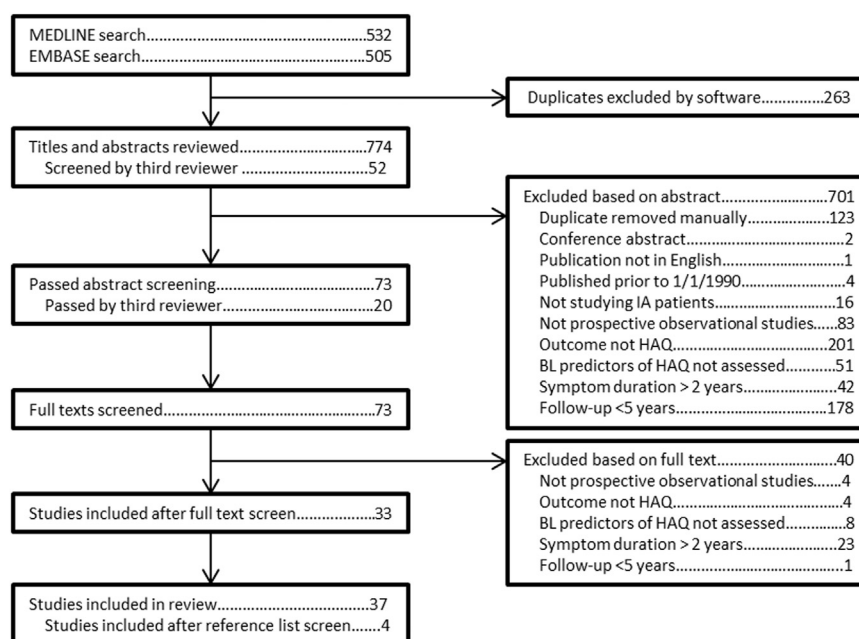


Fig. 1. A flow-diagram of the screening strategy. BL = baseline, HAQ = Health Assessment Questionnaire, IA = inflammatory arthritis, N = number.

Table 1
Descriptive statistics of included studies

Study	Country	N	Age (years), mean (SD) unless otherwise stated	Women (%)	Follow-up (years)	HAQ – baseline, mean (SD) unless otherwise stated	HAQ – final follow-up, mean (SD) unless otherwise stated
Ahlmen [4]	SW	549	Women 54 (16) Men 61 (13)	63	5	Women 1.1 (0.6) Men 0.8 (0.6)	Women 0.7 (0.7) Men 0.5 (0.6)
Ajeganova [26]	SW	1596	55.6 (14.6)	68	15	1.0 (0.6)	0.6 (0.6)
Andersson [43]	SW	1430	Immigrants (I) 55 (13) Non-immigrants (non-I) 55 (14)	I 76 non-I 70	15	I 1.2 (0.7) Non-I 1.0 (0.6)	I 0.7 Non-I 0.6 (no SDs reported)
Bansback [44]	UK	985	Median 55	66	5	Median 1.0	–
Benton [27]	NZ	42	Median (range) 48.5 (27–75)	62	6	Median (range) 0.6 (0–1.8)	Median (range) 0.3 (0–1.6)
Bjork [49]	SW	189	Women 53 (15) Men 58 (13)	69	5	Women 0.9 (0.6) Men 0.8 (0.5)	–
Burr [50]	UK	463	Median (IQR) 55.4 (45.8, 65.4)	66	5	Median (IQR) 0.9 (0.4, 1.5)	Median (IQR) 1.1 (0.4, 1.8)
Camacho [24]	UK	3666	< 55 (%) / 55–74 (%) / ≥ 75 (%) Women 53.6 / 37.6 / 8.8 Men 41.3 / 46.0 / 12.7	66	15	Median (IQR) Women 0.9 (0.4, 1.6) Men 0.6 (0.1, 1.3)	–
Camacho [29]	UK	1872	Median (IQR) Parous 54.3 (44.6, 65.2) Nulliparous 54.1 (36.4, 70.0)	100	15	Median (IQR) Parous 0.9 (0.4, 1.6) Nulliparous 0.9 (0.4, 1.6)	–
Combe [34]	FR	191	50.5 (14.7)	73	5	Median (range) 1.3 (0, 2.8)	Median (range) 0.6 (0, 3.0)
Combe [51]	FR	813	48.1 (12.6)	77	5	1.0 (0.7)	0.5 (0.6)
Contreras-Yanez [25]	MX	107	39.1 (13.3)	89	5	1.5 (0.9, 2.1)	–
Eberhardt [28]	SW	63	52.4 (13.7)	62	5	–	Women 0.9 (0.6) Men 0.4 (0.4)
Eberhardt [52]	SW	99	52.1 (12.8)	67	5	–	Medians ranged from 1.0 to 1.4 depending on immunogenetic group
Forslind [46]	SW	92	53 (15)	66	5	Median (IQR) 1.0 (0.6, 1.4)	0.4 (0.0, 1.1)
Genevay [23]	FR	25	50.6 (15.5)	72	avg 8.5	0.8 (0.8)	–
Hallert [35]	SW	251	Women 55 (16) Men 60 (14)	65	8	Women 0.9 (0.6) Men 0.8 (0.6)	Women 0.9 (0.6) Men 0.4 (0.4)
Humphreys [40]	UK	1995	55 (43, 66)	66	20	Median (IQR) 0.9 (0.4, 1.5)	–
Jäntti [33]	FIN	121	–	–	20	–	–
Kapetanovic [53]	SW	183	52 (12)	63	20	0.9 (0.6)	1.1 (0.7)
Koevoets [39]	NL	508	54 (13)–55 (14)	86	5	1.4 (0.7)	0.6 (0.6)
Kroot [38]	NL	273	ACPA+ 51.1 (15.1) ACPA– 52.4 (14.8)	+ve 62 –ve 73	6	ACPA+ 0.7 (0.4) ACPA– 0.7 (0.4)	–
Kuiper [45]	NL	332	Postmenopausal women 66 (43–83) Premenopausal women 36 (17–53) Older men 63 (43–88) Younger men 41 (23–53)	63	6	–	–
Kuuliala [48]	SW	85	52.4 (range 18–78)	64	5	0.8 (0.5, 1.2)	0.9 (0.4, 1.3)
Lindqvist [54]	SW	183	51.4 (12.4)	63	10	Median 0.8 (unreported IQR)	Median 1.1 (unreported IQR)
Ling [41]	UK	2158	Median 55 (43–67)	65	5	Median (IQR) 0.9 (0.3, 1.5)	Median (IQR) 0.9 (0.1, 1.6)
Malm [32]	SW	1387	55 (14.1)	70	15	1.0 (0.6)	–
Manivel [47]	SW	773	–	–	5	Median Anti-CII+ 1.0 Anti-CII– 1.3	Median Anti-CII+ 0.5 Anti-CII– 0.6
Nair [30]	NL	1034	55 (14.5)	66	avg 6.5	1.2 (0.7)	–
Naseem [42]	UK	843	Median (IQR) 53 (40, 65)	64	5	Median (IQR) 0.8 (0.3, 1.4)	Median (IQR) 0.3 (0.3, 1.5)
Thyberg [55]	SW	251	55 (14)	68	8	0.9 (taken from figure)	0.8 (taken from figure)
Verstappen [36]	NL	112	49 (12.4)	76	avg 7	1.1 (0.7)	0.7 (0.6)
Welsing [56]	NL	378	54.8 (14.8)	64	9	Median (IQR) 0.5 (0.2, 1.1)	0.6
Welsing [31]	NL	185	55	64	9	Median 0.5	Median 0.6
Wiles [37]	UK	684	Median (IQR) 55 (42, 68)	67	5	Median (IQR) 0.8 (0.3, 1.5)	Median (IQR) 0.9 (0.3, 1.6)
Wiles [57]	UK	528	53 (41–66)	67	5	–	Median (IQR) 0.9 (0.3, 1.6)
Woolf [58]	UK	88	45	72	5	–	–

ACPA = anti-citrullinated protein antibodies, Anti-CII = anticollagen type II antibodies, FIN = Finland, FR = France, HAQ = Health Assessment Questionnaire, IQR = Interquartile range, MX = Mexico, NL = The Netherlands, NZ = New Zealand, RA = rheumatoid arthritis, SD = standard deviation, SW = Sweden, UK = United Kingdom.

Quality assessment

Two reviewers assessed the conduct and reporting of each study using a system adapted from Pasma et al. [22]. Details on the methods and results of the quality assessment can be found in [Supplementary file 2](#).

Data abstraction

A data abstraction form was created to extract and summarise information from each included study (see data abstraction form in [Supplementary file 3](#)), including: number of patients in each study, the length of follow-up, age, gender, baseline and follow-up HAQ scores and information on analyses carried out assessing the association between baseline predictors and follow-up HAQ score. The predictors of long-term HAQ score were grouped into five categories and presented in tables: demographics, patient reported outcomes, disease activity, autoantibody status and miscellaneous. Each of these tables (i.e., other than [Tables 1 and 2](#)) displays results from studies that performed multivariable analyses first, followed by studies that only performed univariable analyses. Within these subsections the studies were sorted by sample size. The statistical method of each analysis is reported, followed by effect sizes with 95% confidence intervals.

Results

A summary of the included studies ($N = 37$), including demographics, follow-up lengths and baseline and final follow-up HAQ scores is presented in [Table 1](#). The studies are presented in

alphabetical order of first author to aid cross-reference between tables. Sample sizes ranged from $n = 25$ [23] to $n = 3666$ [24], and follow-up duration from 5 to 20 years (median (IQR) = 6 (5, 10) years). The median age of the patients ranged from 39.1 [25] to 55.6 years [26] (median = 53 years; 27/37 studies reported median age for the entire cohort). The proportions of women ranged from 62% [27,28] to 100% [29] (median = 66%; 33/37 studies reported the proportion of women). [Table 2](#) summarises the results for each of the predictors assessed in the review.

Assessment of baseline predictors

Demographics

The majority of studies assessing the association between age and long-term HAQ score reported that older age at symptom onset was associated with higher HAQ scores at long-term follow-up (18 studies total, 13 (72%) reported a significant association including 11 multivariable analyses) ([Table 3](#)); 10.7k patients were included, of which 9.8k were included in analyses that reported a significant association (91.6%). The largest study ($N = 3666$) assessed the association between age and higher HAQ scores over 15 years. The HAQ scores of men aged between 55 and 74 years were, on average, 0.19 (95% CI: -0.01, 0.39) higher and those of men aged ≥ 75 years were, on average, 1.81 (95% CI: 1.25, 2.36) higher than those men < 55 years of age. Older women also had higher HAQ scores compared to younger women, but to a lesser degree (mean difference (95% CI): < 55 years = ref, 55–74 = 0.26 (0.12 to 0.40), $\geq 75 = 0.51$ (0.05, 0.98)) [24].

A majority of studies investigating the association between gender and later HAQ scores reported that women had significantly higher long-term HAQ scores than men. In total, 21 studies

Table 2
Summary of predictors

Category	Baseline predictor	Maximum follow-up length	Studies assessing the association		Studies reporting a significant association		% of total sample in significant studies	Level of evidence [§]
			Number of studies	Number of patients	Number of studies	Number of patients		
Demographics	Age at baseline	20	18	10.7k	13	9.8k	91.6%	✓✓
	Female gender	15	21	11.3k	13	9.9k	87.4%	✓✓
Patient reported outcomes	HAQ	20	10	4.0k	9	4.0k	99.0%	✓✓
	Pain VAS	15	6	2.9k	5	2.8k	93.6%	✓✓
Disease activity	DAS28	6	5	1.1k	4	888	79.2%	–
	Swollen joint count	20	5	2.6k	1	684	26.6%	✗
	Tender joint count	15	4	2.5k	2	2.1k	84.5%	✓
	Ritchie Index	6	2	233	2	233	100%	–
	CRP	6	4	510	2	279	54.7%	–
	ESR	20	6	2.2k	2	1.6k	72.3%	✓
Serology	RF	20	11	4.6k	3	546	11.9%	✗✗
	ACPA	20	6	3.8k	1	2.0k	52.9%	–
Other	Erosions	20	5	2.7k	1	191	7.0%	✗✗
	Genetics	20	8	4.0k	1	2.2k	54.6%	–
	Morning stiffness	20	3	424	3	424	100%	–
	BMI	15	1	1.6k	1	1.6k	100%	–
	Immigrant status	15	1	1.4k	1	1.4k	100%	–
	SES	5	2	1.0k	2	1.0k	100%	–
	Reproductive factors	15	2 [†]	– [†]	2	– [†]	–	–
Other biomarkers	20	5	– [†]	2	– [†]	–	–	

[†]Reproductive factors/other biomarkers included are heterogeneous and therefore analysis populations were not summed.

[§]Key.

✓✓ = $\geq 85\%$ total participants in studies reporting significant association & ≥ 2000 total participants studied;

✓ = $\geq 60\%$ & $< 85\%$ total participants in studies reporting significant association & ≥ 2000 total participants studied;

– = $\geq 40\%$ & $< 60\%$ or < 2000 total participants;

✗ = $\geq 15\%$ & $< 40\%$ total participants in studies reporting significant association & ≥ 2000 total participants studied;

✗✗ = $< 15\%$ total participants in studies reporting significant association & ≥ 2000 total participants studied.

ACPA = anti-citrullinated protein antibody, BMI = body mass index, CRP = C-reactive protein, DAS28 = Disease activity Score (28), ESR = erythrocyte sedimentation rate, HAQ = Health Assessment Questionnaire, N = number, RF = rheumatoid factor, SES = socioeconomic status, VAS = visual analogue scale.

Table 3
Baseline demographic predictors of follow-up HAQ score

Study details			Predictor: Age		Predictor: Gender		Adjusted for
Authors	N	Analysis method	Associated with HAQ	Effect size ^a	Associated with HAQ	Effect size ^a	
<i>Multivariable analyses</i>							
Camacho [24]	3666	Multivariable linear random effects model (< 55 years used as the reference category)	✓	Men: 55–74 <i>b</i> 0.19 (−0.01, 0.39) ≥75 <i>b</i> 1.81 (1.25, 2.36) Women: 55–74 <i>b</i> 0.26 (0.12, 0.40) ≥75 <i>b</i> 0.51 (0.05, 0.98)	✓	Women vs. men: <i>b</i> 0.24 (0.20, 0.29) Further adjustment*: <i>b</i> 0.29 (0.25, 0.34)	Age at final follow-up, year recruited to the study *Further adjustment: baseline disease duration and DMARDs within 6 months of symptom onset.
Malm [32]	1387	Logistic regression (HAQ cut-off = 0.75)	✓	Age at onset (years): OR 1.03 (1.02, 1.04)	✓	Women vs. men: OR 2.53 (1.85, 3.46)	Disease duration
Nair [30]	1034	Linear mixed model	✓	Age at onset (years): Cohort 1 <i>b</i> 0.00, <i>p</i> < 0.01 Cohort 2 <i>b</i> 0.00, <i>p</i> = 0.01	✓	“sex”: Cohort 1 <i>b</i> 0.08, <i>p</i> < 0.01 Cohort 2 <i>b</i> 0.19, <i>p</i> < 0.01	Treatment, SHS, HAQ (t-1), DAS28, BMI, RF
Combe [51]	813	Logistic regression (HAQ cut-off = 0.3)	✓	“older age”: OR 1.91 (1.32, 1.77)	✓	Women vs. men: OR 1.60 (1.02, 2.50)	Baseline: HAQ, pain
Wiles [37]	684	Generalised estimating equations analysis (HAQ cut-off = 1)	✓	< 47 years at onset (ref cat): 47–63 OR 1.45 (1.06, 2.00) ≥64 OR 3.21 (2.33, 4.42)	✓	Women vs. men: OR 1.70 (1.29, 2.24)	Duration from symptom onset to baseline, Time-varying: morning stiffness, RF, rheumatoid nodules, number of deformed joints
Ahlmen [4]	549	ANCOVA	–	–	✓	Mean (SD) HAQ: Men 0.51 (0.56) Women 0.73 (0.68) <i>p</i> < 0.01	Age
Wiles [57]	528	Logistic regression Model 1 HAQ cut-off = 1, Model 2 HAQ cut-off = 1.5	✓	< 47 years (ref cat): 47–63: Model 1 OR 2.06 (1.11, 3.83) Model 2 OR 1.62 (0.79, 3.32) ≥ 64: Model 1 OR 3.46 (1.77, 6.76) Model 2 OR 2.70 (1.29, 5.67)	–	–	Year 1 HAQ, Nodules, Knee involvement factor, Tenderness factor (factors created using principal component analysis)
Welsing [56]	378	General linear mixed model	✓	Age at onset (years): <i>b</i> 0.01 (0.01, 0.20)	✓	Women vs. men: <i>b</i> 0.22 (0.08, 0.36)	Baseline: RF; time-varying: SHS, squared SHS, DAS28
Kroot [38]	273	Multiple regression	✓	Age at entry (years): <i>b</i> 0.01, <i>p</i> < 0.01	✓	Female gender: <i>b</i> −0.128, <i>p</i> < 0.05	RF, DAS28, HLA-DR4 gene, ACPA
Hallert [35]	251	Generalised estimating equations analysis	✗	NS – coefficients and confidence interval not reported	✗	NS – coefficients and confidence interval not reported	DMARD use, biologic use, grip force, SOFI-hand, SOFI-upper extremity, SOFI-lower extremity, GAT, pain, walking time
Bjork [49]	189	Projections to latent structure discriminant analysis (HAQ cut-off = 0.08)	✗	Baseline age: VIP 0.22 (“not important”)	✓	“sex”: VIP 1.39 (“important”)	Baseline: HAQ, grip force, SOFI-lower limb, gender, walking speed, GAT, wellbeing, CRP, SOFI-hand, ESR, tender joints, PGA, pain, SOFI-upper limb, swollen joints

Table 3 (continued)

Study details			Predictor: Age		Predictor: Gender		Adjusted for
Authors	N	Analysis method	Associated with HAQ	Effect size ^a	Associated with HAQ	Effect size ^a	
Welsing [31]	185	Mixed model (HAQ was log transformed)	✓	Age at onset per year: <i>b</i> 0.02, <i>p</i> < 0.01	✓	Women vs. men: <i>b</i> 0.38 <i>p</i> = 0.02	DAS28, Modified SHS, Modified SHS squared, age*modified SHS
Lindqvist [54]	183	Stepwise logistic regression (HAQ cut-off = 1.0)	✗	NS – coefficients and confidence interval not reported	✗	NS – coefficients and confidence interval not reported	Genotype, RF, HAQ, ESR, active joint count
Kapetanovic [53]	183	Hierarchical linear regression	–	–	✗	HAQ at final follow-up: “sex” <i>b</i> –0.095, <i>p</i> = NS HAQ over time (AUC): “sex” <i>b</i> –0.20, <i>p</i> < 0.01	CCI, DAS, joint damage
Verstappen [36]	112	Logistic regression (HAQ cut-off = 1)	✓	Age at onset: OR 1.05 (1.01, 1.09)	✗	Women vs. men: OR 0.90 (0.37, 2.17)	Disease duration (natural log transformed)
Contreras-Yanez [25]	107	Multivariable linear regression	✓	Age at baseline (years): <i>b</i> 0.10, <i>p</i> = 0.001	✗	NS – coefficients and confidence intervals not reported	Variables tested in univariable analysis: age, gender, disease duration, DAS28, persistence of DMARDs, comorbidity
Kuuliala [48]	85	Logistic regression (HAQ cut-off = 0.9)	✗	Age at entry (years): OR 1.02 (0.97, 1.07)	✓	Women vs. men: OR 5.51 (1.81, 16.8)	RF, Shared epitope, tertiles of soluble E-selectin
Eberhardt [28]	63	Logistic regression (HAQ cut-off = 1)	–	–	✓	Women vs. men: OR 1.02 <i>p</i> < 0.01	“[demographic,] clinical, radiographic and laboratory data”
<i>Univariable analyses</i>							
Koevoets [39]	508	Generalised estimating equations analysis	–	–	✓	Women vs. men: <i>b</i> 0.14 (0.05, 0.24)	–
Kuiper [45]	332	Student's <i>t</i> test	✓	Older men had higher HAQ scores than younger men (<i>p</i> < 0.01)	✓	Women had higher HAQ scores than men (<i>p</i> < 0.05)	–
Combe [34]	191	Spearman's test	✗	NS – coefficients and confidence interval not reported	✗	NS – coefficients and confidence interval not reported	–
Jääntti [33]	121	Somers' <i>d</i>	✓	Age at entry: Somers' <i>d</i> 0.30 (0.16, 0.45)	✗	“sex”: Somers' <i>d</i> 0.01 (–0.31, 0.33)	–

See Table 2 for acronym definitions: ACPA, BMI, DAS28, ESR, HAQ, N, RF.

^a Brackets indicate 95% confidence interval unless otherwise stated; ANCOVA = analysis of covariance, AUC = area under the curve, *b* = regression coefficient, BL = baseline, CCI = Charlson Comorbidity Index, DMARDs = disease modifying anti-rheumatic drugs, FU = follow-up, GAT = Grip Ability Test, NS = non-significant, OR = odds ratio, PGA = patient global assessment, RA = rheumatoid arthritis, SD = standard deviation, SE = standard error, SHS = Sharp score, SOFI = Signals of Functional Impairment, VIP = variable influence on projection.

assessed the association between gender and subsequent HAQ scores. Of these, 11 analyses (9 multivariable) reported that women had significantly higher HAQ scores at long-term follow-up than men; one multivariable analysis reported that men had significantly higher HAQ scores than women; three multivariable analyses reported a significant association but the direction of the association was unclear (i.e., the coefficient was labelled “gender” and the reference category (men/women) was not clearly reported); six analyses (4 multivariable) reported no significant association between gender and future HAQ score (Table 3). In total, 11.3k patients were included, of which 9.9k were included in analyses that reported a significant association (87.4%). The average difference between the HAQ scores of women and men ranged

from 0.08 [30] to 0.38 [31], based on studies reporting a significant association between female gender and higher HAQ score from linear regression analysis. A study by Malm et al. including 1.4k patients followed for 15 years reported that women had a two and a half times increased odds of having a HAQ score over 0.75 at the 15th year assessment compared to men (OR 2.53, 95% CI: 1.85, 3.46) [32].

Patient reported outcomes

Nine (7 multivariable) of the 10 studies that investigated the relationship reported a positive association between higher disability at baseline and higher disability at long-term follow-up, whilst the remaining multivariable analysis approached significance

(Table 4). Nine of these reported a positive association between baseline HAQ and follow-up HAQ, whilst one used an alternative measure of baseline functional disability [33]. In total, 4.0k patients were included, of which 3.97k were included in analyses that reported a significant association (99.0%). One study ($N = 191$) reported that each unit increase in HAQ score at baseline was associated with a 0.39 ($p = 0.0001$) increase in HAQ score at five

years [34]. Another study ($N = 1.4k$) reported that each unit increase in baseline HAQ was associated with a 3.57 (95% CI: 2.84, 4.49) times increased odds of having HAQ > 0.75 at 15th year assessment [32].

Four out of six analyses (all multivariable) assessing the relationship reported a positive association between baseline **pain** visual analogue scale (VAS) scores and follow-up HAQ scores;

Table 4
Baseline patient reported outcomes as predictors of follow-up HAQ score

Study details			Predictor: Baseline HAQ		Predictor: Baseline pain		Adjusted for
Study	N	Analysis method	Associated with HAQ	Effect size ^a	Associated with HAQ	Effect size ^a	
<i>Multivariable analyses</i>							
Malm [32]	1387	Logistic regression (HAQ cut-off = 0.75)	✓	Per unit baseline HAQ: OR 3.57 (2.84, 4.49)	✓ (VAS)	Per unit baseline pain VAS: OR 1.02 (1.02, 1.03)	Age, gender, disease duration
Bansback [44]	985	Logistic regression (HAQ cut-off = 1.5)	✓	Per unit baseline HAQ: OR 1.70, $p < 0.01$	–	–	Baseline: Carstairs deprivation index, functional grade, haemoglobin level, Larsen score; year 1: functional grade, HAQ, DAS28
Combe [51]	813	Logistic regression (HAQ cut-off = 0.3)	✓	Baseline HAQ cut-off = 0.88: OR 2.90 (2.00, 4.19)	✓ (VAS)	Baseline pain VAS cut-off = 34: OR 1.69 (1.17, 2.44)	Age, gender
Hallert [35]	251	Generalised estimating equations analysis	–	–	✓ (VAS)	Per unit baseline pain VAS: b 0.01, $p < 0.01$	DMARD use, biologic use, grip force, SOFI-hand, SOFI-upper extremity, SOFI-lower extremity, GAT, pain, walking time
Combe [34]	191	Spearman correlation & linear regression	✓	ρ 0.47, $p < 0.01$ b 0.39, $p < 0.01$	✓ (VAS)	ρ 0.32, $p < 0.01$ b not reported	ESR, CRP, Ritchie index
Bjork [49]	189	Projections to latent structure discriminant analysis (HAQ cut-off = 0.08)	✓	Baseline HAQ cut-off=0.08: VIP = 1.97 (“important”)	× (VAS)	VIP = 0.30 (“not important”)	Baseline: age, gender, grip force, SOFI-lower limb, walking speed, GAT, wellbeing, CRP, SOFI-hand, ESR, tender joints, PGA, SOFI-upper limb, swollen joints
Verstappen [36]	112	Logistic regression (HAQ cut-off = 1)	✓	Per unit baseline HAQ: OR 2.63 (1.30, 5.32)	✓ (VAS)	Per unit baseline pain VAS: OR 1.02 (0.99, 1.03)	Disease duration (natural log transformed)
Eberhardt [28]	63	Logistic regression (HAQ cut-off = 1.0)	✓	Baseline HAQ cut-off = 1.00: OR 2.08, $p < 0.01$	–	–	“[demographic,] clinical, radiographic and laboratory data”
Benton [27]	42	Logistic regression (HAQ cut-off = 0.25) OR are the odds of being in the low HAQ group	×	Per unit baseline HAQ: OR 0.16 (0.02, 1.01)	–	–	Baseline: DAS, Ritchie index, CRP, Sharp score; one year: DAS, Ritchie, CRP, HAQ
<i>Univariable analyses</i>							
Jäntti [33]	121	Somers' d	✓	Somers' $d = 0.28$ (0.11, 0.45)	–	–	–
Contreras-Yanez [25]	107	Student's t test Comparison groups: HAQ ≤ 0.2 at 5 years, yes/no.	✓	Median (IQR) baseline HAQ: HAQ ≤ 0.2 1.4 (0.8–2) HAQ > 0.2 2.1 (1.6–3) $p < 0.01$	–	–	–

See Table 2 for acronym definitions: DAS28, ESR, HAQ, N.

See Table 3 for acronym definitions: b, DMARD, GAT, OR, PGA, SOFI, VIP.

^a Brackets indicate 95% confidence interval unless otherwise stated; CRP = C-reactive protein, IQR = interquartile range, VAS = visual analogue scale, ρ = Spearman's rho.

Table 5
Baseline disease activity measures as predictors of follow-up HAQ score

Study Details			Predictor: Baseline DAS28		Predictor: Baseline joint counts		Adjusted for
Study	N	Analysis method	Associated with HAQ	Effect size ^a	Associated with HAQ	Effect size ^a	
<i>Multivariable analyses</i>							
Malm [32]	1387	Logistic regression (HAQ cut-off = 0.75)	–	–	✗ (SJC28)	Per baseline swollen joint: OR 1.02 (0.99, 1.04)	Age, gender, disease duration
					✓ (TJC28)	Per baseline tender joint: OR 1.05 (1.03, 1.07)	
Wiles [37]	684	Generalised estimating equations analysis (HAQ cut-off = 1) Joint areas clustered using principal component factor analysis	–	–	✓ (swelling on different joint sites)	MCP OR 1.57 (1.38, 1.77) Wrist OR 1.49 (1.34, 1.67) Elbow OR 1.20 (1.09, 1.33) Shoulder OR 1.00 (0.88, 1.14)	Age at symptom onset, gender, delay to presentation, morning stiffness, RF, number of deformed joints, nodules
					✓ (tenderness factor)	Knee OR 1.46 (1.33, 1.61) Ankle OR 1.38 (1.24, 1.53) MTP OR 1.15 (1.00, 1.31) Tenderness OR 1.53 (1.37, 1.70)	
Kroot [38]	273	Multiple linear regression	✓	Per unit baseline DAS28: <i>b</i> 0.10, <i>p</i> < 0.01	–	–	RF, HLA-DR4 gene, ACPA, age, gender
Combe [34]	191	Spearman correlation & linear regression	✓ (univariable) ✗ (multivariable)	ρ 0.263 <i>b</i> not reported, not associated	✗ (SJC, univariable) ✓ (TJC, univariable) ✗ (TJC, multivariable) ✓ (Ritchie, multivariable)	Baseline swollen joints: ρ 0.00, <i>p</i> = 0.45 Baseline tender joints: ρ 0.27, <i>p</i> < 0.01 Baseline Ritchie Index: ρ 0.29, <i>p</i> < 0.01 <i>b</i> 0.02, <i>p</i> = 0.05	Baseline: ESR, CRP, HAQ, pain VAS
Bjork [49]	189	Projections to latent structure discriminant analysis (HAQ cut-off = 0.08)	–	–	✗ (SJC) ✗ (TJC)	Baseline SJC: VIP 0.09 (“not important”) Baseline TJC: VIP 0.42 (“not important”)	Baseline: age, gender, HAQ, grip force, SOFI-lower limb, walking speed, GAT, wellbeing, CRP, SOFI-hand, ESR, PGA, pain, SOFI-upper limb
Lindqvist [54]	183	Stepwise Logistic regression (HAQ cut-off = 1.0)	–	–	✗ (active joint count)	NS – coefficients and confidence interval not reported	Age, gender, genotype, RF, HAQ, ESR
Verstappen [36]	112	Logistic regression (HAQ cut-off = 1.0)	–	–	✗ (Thomson joint score)	Per unit baseline Thomson score: OR 1.003	Disease duration (natural log transformed)
Benton [27]	42	Logistic regression (HAQ cut-off = 0.25) OR are the odds of being in the low HAQ group	✗	Per unit baseline DAS28: OR 0.72 (0.35, 1.49)	✓ (Ritchie index)	Per unit baseline Ritchie index: OR 0.86 (0.74, 1.00)	Baseline: HAQ, CRP, Sharp score; one year: DAS, Ritchie, CRP, HAQ

Table 5 (continued)

Study Details			Predictor: Baseline DAS28		Predictor: Baseline joint counts		Adjusted for
Study	N	Analysis method	Associated with HAQ	Effect size ^a	Associated with HAQ	Effect size ^a	
<i>Univariable analyses</i>							
Koevoet [39]	508	Generalised estimating equations analysis	✓	Per unit baseline DAS28: <i>b</i> 0.13 (0.08, 0.18)	–	–	–
Jäntti [33]	121	Somers' d	–	–	x (SJC)	<i>d</i> = –0.03 (–0.19, 0.14)	–
Contreras-Yanez [25]	107	Student's T test Comparison groups: HAQ ≤ 0.2 at 5 years, yes/no	✓	Median (IQR) baseline DAS28: HAQ ≤ 0.2: 6.0 (4.9–6.9) HAQ > 0.2: 6.8 (6.0–7.7). <i>p</i> = 0.02	–	–	–

See Table 2 for acronym definitions: ACPA, CRP, DAS28, ESR, HAQ, N, RF, VAS.

See Table 3 for acronym definitions: b, GAT, NS, OR, PGA, SOFI, VIP.

See Table 4 for acronym definitions: IQR, ρ .

^a Brackets indicate 95% confidence interval unless otherwise stated; MCP = metacarpophalangeal joint, MTP = metatarsophalangeal joint SJC = swollen joint count, TJC = tender joint count.

another study approached significance (Table 4). In total, 2.9k patients were included, of which 2.8k were included in analyses that reported a significant association (93.6%). One study using generalised estimating equations analysis over eight years of follow-up, reported that each centimetre increase in pain VAS at baseline was associated with an average increase of 0.06 HAQ score over follow-up [35]. Two studies reported a 2% increased odds in being in a higher HAQ category at follow-up per millimetre increase in pain VAS at baseline, one after seven years of follow-up [36], the other after 15 [32].

Disease activity

One multivariable analysis reported a significant positive association between baseline **swollen joint count** and follow-up HAQ scores [37], whilst four other analyses (two multivariable) reported no association. In total, 2.6k patients were included, of which 684 were included in analyses that reported significant results (26.6%).

Two multivariable analyses reported a significant positive association between baseline **tender joint count** and follow-up HAQ score, whilst two other multivariable analyses did not report a significant association (Table 5). In total, 2.5k patients were included, of which 2.0k were included in analyses that reported significant results (84.5%).

Furthermore, two small studies ($N = 191$ and 42) reported a positive association between baseline **Ritchie Index** (which includes a measure of joint tenderness) and subsequent HAQ scores [27,34]. Thus, the evidence regarding the predictive ability of baseline tender joint counts suggests it may be a useful predictor of long-term HAQ scores, whereas baseline swollen joint count is unlikely to be a predictor of long-term disability.

Two studies (one multivariable) reported that higher **C-reactive protein** (CRP) level was associated with higher long-term functional disability, whilst two multivariable analyses reported no significant association. In total, 510 patients were included, of which 279 were included in analyses that reported significant results (54.7%).

Two multivariable analyses reported a significant association between higher baseline **erythrocyte sedimentation rate** (ESR) and higher follow-up HAQ score, although with small effect sizes (HAQ at 15 years > 0.75: OR 1.01 per unit increase in ESR at baseline, 95% CI: 1.002, 1.012 [32]; mean increase in HAQ score at 5 years: 0.008 per unit increase baseline ESR, $p = 0.006$ [34]). Four smaller analyses (three multivariable) reported no significant

association (Table 6). In total, 2.2k patients were included, of which 1.6k were included in analyses that reported significant results (72.3%). Thus, there is inconsistent evidence about the relationship between higher CRP and long-term functional disability but ESR is likely to be a weak predictor of HAQ score.

Of the five studies which assessed the association, three univariable analyses and one multivariable analysis reported a positive association between baseline **Disease Activity Score (28)** (DAS28) and follow-up HAQ scores, whilst one multivariable analysis did not report a significant association (Table 5). In total, 1.1k patients were included, of which 888 were included in analyses that reported a significant association (79.2%). The average increase in HAQ score at follow-up per unit increase in baseline DAS28 ranged from 0.100 [38] to 0.130 [39], based on analyses reporting significant associations from linear regressions.

Serology

Rheumatoid factor (RF) positivity did not predict higher HAQ scores in the majority of included studies. Eight analyses (four multivariable) reported no association between RF positivity and later HAQ scores, whilst three analyses (two multivariable) did report an association (Table 6). The two largest multivariable analyses (mean difference in HAQ between RF+ and RF– = –0.03 (95% CI: –0.12, 0.05) [$N = 1995$] [40], 0.00014 ($p = 0.9959$) [$N = 1034$] [30]) and the largest univariable analysis (mean difference in HAQ between RF+ and RF– = –0.027 (95% CI: –0.130, 0.076) [$N = 508$] [39]) found no association. In total, 4.6k patients were included, of which 546 were included in analyses that reported significant results (11.9%).

The largest analysis assessing the association between **anti-citrullinated protein antibodies** (ACPA) positivity and subsequent HAQ scores reported a significant association ($N = 1995$; mean difference in HAQ between ACPA+ and ACPA– = 0.12 (95% CI: 0.02, 0.21)) [40], but five other analyses (two multivariable) found no association (Table 6). In total, 3.8k patients were included, of which 2.0k were included in analyses that reported significant results (52.9%). Thus at present the literature is equivocal as to whether ACPA positivity is a useful predictor of increased long-term functional disability.

Erosions

Four out of five studies (three multivariable) reported no significant association between **erosion score** at baseline and

Table 6
Baseline blood analyses as predictors of follow-up HAQ score

Study details			Predictor: Baseline ESR/CRP		Predictor: Baseline RF/ACPA		Adjusted for
Study	N	Analysis method	Associated with HAQ	Effect size ^a	Associated with HAQ	Effect size ^a	
<i>Multivariable analyses</i>							
Humphreys [40]	1995	Generalised estimating equations analysis	–	–	✗ (RF) ✓ (ACPA)	RF + vs. RF–: <i>b</i> –0.03 (–0.12, 0.05) ACPA+ vs. ACPA–: <i>b</i> 0.12 (0.02, 0.21)	Age, gender, smoking status, polynomials of disease duration, year of recruitment
Malm [32]	1387	Logistic regression (HAQ cut-off = 0.75)	✓ (ESR)	Per unit baseline ESR: OR 1.01 (1.00, 1.01)	–	–	Age, gender, disease duration
Nair [30]	1034	Linear mixed model	–	–	✗ (RF)	RF+ vs. RF–: Cohort 1 <i>b</i> 0.00, <i>p</i> = 0.99 Cohort 2 <i>b</i> 0.00, <i>p</i> = 0.90	Age, gender, treatment, Sharp Score (van der Heijde modification), HAQ (t-1), DAS28, BMI
Burr [50]	640	Logistic regression (HAQ cut-off = 1)	–	–	✗ (ACPA)	Per unit baseline ACPA titre: OR 1.00 (0.99, 1.01)	Baseline: age, gender, symptom duration, CRP, RF, HAQ, swollen joint count, tender joint count
Kroot [38]	273	Multivariable linear regression	–	–	✓ (RF) ✗ (ACPA)	RF+ vs. RF–: <i>b</i> 0.15, <i>p</i> < 0.05 ACPA+ vs. ACPA–: <i>b</i> 0.00, <i>p</i> = NS	Age, gender, DAS28, HLA-DR4 gene,
Combe [34]	191	Multivariable linear regression	✓ (CRP) ✓ (ESR)	Per unit baseline CRP: <i>b</i> 0.01 (<i>p</i> < 0.01) Per unit baseline ESR: <i>b</i> 0.01 (<i>p</i> < 0.01)	–	–	Baseline: DAS, swollen joint count, tender joint count, HAQ, pain VAS
Bjork [49]	189	Projections to latent structure discriminant analysis (HAQ cut-off = 0.08)	✗ (CRP) ✗ (ESR)	Per unit baseline CRP: VIP 0.63 (“not important”) Per unit baseline ESR: VIP 0.49 (“not important”)	–	–	Baseline: Age, gender, HAQ, grip force, SOFI-lower limb, walking speed, GAT, wellbeing, swollen joint count, SOFI-hand, tender joint count, PGA, pain, SOFI-upper limb
Welsing [31]	185	General linear mixed model	–	–	✓ (RF)	RF+ vs. RF–: <i>b</i> 0.19 (0.03, 0.35)	Baseline: age, sex; time-varying: Sharp score, squared Sharp score, DAS28
Lindqvist [54]	183	Stepwise logistic regression (HAQ cut-off = 1.0)	✗ (ESR)	NS – coefficients and confidence intervals not reported	✗ (RF)	NS – coefficients and confidence intervals not reported	Age, gender, genotype, HAQ, active joint count
Verstappen [36]	112	Logistic regression (HAQ cut-off = 1)	✗ (ESR)	Per unit baseline ESR: OR 1.00 (0.99, 1.02)	–	–	Disease duration (natural log transformed)
Kuuliala [48]	85	Logistic regression (HAQ cut-off = 0.9)	–	–	✗ (RF)	RF+ vs. RF–: OR 1.09 (0.33, 3.57)	Age, gender, shared epitope, tertiles of soluble E-selectin
Benton [27]	42	Logistic regression (HAQ cut-off = 0.25) OR are the odds of being in the low HAQ group	✗ (CRP)	Per unit baseline CRP: OR 0.99 (0.96, 1.03)	–	–	Baseline: DAS, Ritchie index, HAQ, Sharp score; one year: DAS, Ritchie index, CRP, HAQ
<i>Univariable analyses</i>							
Koevoets [39]	508	Generalised estimating equations analysis	–	–	✗ (RF) ✗ (ACPA)	RF+ vs. RF–: <i>b</i> –0.03 (–0.13, 0.08) ACPA+ vs. ACPA–: <i>b</i> –0.03 (–0.13, 0.07)	–
Thyberg [55]	251	Chi-Square (HAQ cut-off=1)	–	–	✗ (ACPA)	NS difference between proportion of ACPA+ patients between HAQ subgroups	–

Table 6 (continued)

Study details			Predictor: Baseline ESR/CRP		Predictor: Baseline RF/ACPA		Adjusted for
Study	N	Analysis method	Associated with HAQ	Effect size ^a	Associated with HAQ	Effect size ^a	
Jäntti [33]	121	Somers' d	x (ESR)	d = 0.12 (-0.04, 0.28)	x (RF)	d = -0.18 (-0.62, 0.26)	-
Contreras-Yanez [25]	107	Student's t test Comparison groups: HAQ ≤ 0.2 at 5 years, yes/no	-	-	x (RF) x (ACPA)	Proportion baseline RF+: HAQ ≤ 0.2 82.1% HAQ > 0.2 82.6% p = 1.00 Proportion baseline ACPA+: HAQ ≤ 0.2 85.7% HAQ > 0.2 87.0% p = 1.00	-
Woolf [58]	88	Calculated sensitivity and specificity of having HAQ > 0 at 5 years	✓ ("raised CRP")	Specificity/ sensitivity: 93/74	✓ (RF)	Specificity/sensitivity: 64/37	-
Genevay [23]	25	Mann-Whitney	-	-	x (RF)	Mean HAQ: RF+ 0.78 RF- 0.80 p = NS	-

See Table 2 for acronym definitions: ACPA, BMI, CRP, DAS, ESR, HAQ, N, RF, VAS.

See Table 3 for acronym definitions: b,GAT,NS,OR,PGA, SOFI, VIP.

^a Brackets indicate 95% confidence interval unless otherwise stated.

subsequent higher HAQ scores (Table 7). One univariable analysis reported a significant correlation [34], but with a low Spearman's rho ($\rho = 0.167$) indicating a weak relationship. Of the 2.7k patients included in analyses assessing the association, only 191 were included in the analysis that reported a significant association (7.0%). Therefore, based on current evidence, baseline erosions are not a predictor of long-term functional disability in patients with early inflammatory arthritis.

Morning stiffness

All three analyses (one multivariable; total patients included = 424) that assessed the relationship between **morning stiffness** and long-term HAQ score reported a positive association (Table 7). One study reported a low Spearman's rho ($\rho = 0.211$) indicating a weak relationship [34], another study reported a 1% increased odds of having a HAQ > 1 after seven years per minute increase in morning stiffness at baseline compared to no morning stiffness (max = 180; OR 1.008, 95% CI: 1.001, 1.016) [36]. This suggests that baseline morning stiffness may be weakly associated with long-term functional disability, but all the studies reporting on this relationship were relatively small ($N < 200$).

Genetic factors

Eight analyses (four multivariable) assessed the association between **RA susceptibility genes** (HLA and PTPN22 variants) and long-term functional disability (Table 7). One large study reported significant associations between different amino acids at positions 11, 71 and 74 of HLA-DRB1 and small increases or decreases in disability over five years [41]. Of the other studies, seven studies examined different HLA regions as the independent variable and one study examined PTPN22 variants [42], all reporting no significant associations. Therefore, the published literature suggests that specific amino acids at different positions of the HLA-DRB1 gene are weakly associated with long-term disability. Other genes within the HLA region do not predict long-term functional disability, with little research on other genetic regions.

Other factors

The one multivariable analysis ($N = 1.6k$) which assessed **body mass index** (BMI) as a predictor of long-term HAQ reported that each unit increase in BMI at baseline was associated with a 0.2 increase in HAQ at 15 years follow-up (Table 7) [26].

The one univariable analysis ($N = 1.4k$) which assessed whether **immigrant status** predicted long-term HAQ score reported that immigrants to Sweden had significantly higher HAQ scores after 15 years compared to non-immigrants (Table 7) [43]. Bansback et al. reported that those in the highest category (i.e., most deprived) of the Carstairs Deprivation Index, a measure of **socioeconomic status**, at baseline had an almost two-fold increased adjusted odds of having a HAQ > 1.5 after five years, compared to those in the lowest category (OR 1.984, $p = 0.044$, $N = 985$) [44]. Eberhardt et al. ($N = 63$) reported that compared to those with 0–9 years of education, patients who had 10–11 years had a 13% lower odds of having HAQ score > 1.0 at five years and those with ≥12 years of education had 26% lower odds, after adjusting for confounders [28].

Two studies reported on **reproductive factors**. One reported a significant association between being parous at baseline vs. nulliparous and subsequent lower HAQ score over 15 years of follow-up ($N = 1.9k$) [29], and the other reported that women who were postmenopausal at baseline had significantly higher HAQ scores six years later than women who were premenopausal ($N = 332$) [45]. However, the latter study did not control for age.

Four studies examined the association between **other biomarkers** and subsequent HAQ scores (Table 7). No association was found between antifilaggrin antibody, antiperinuclear factor, antikeratin or anticollagen type II antibody status and subsequent HAQ scores [23,46,47]. However, anti-carbamylated protein antibody positivity and being in the highest tertile of sE-selectin level were associated with higher long-term HAQ score [40,48].

Discussion

This systematic review identified 37 studies that assessed the association between a total of 20 baseline variables and

Table 7
Miscellaneous predictors of long-term functional disability

Predictor	Study	N	Analysis method	Associated with HAQ	Effect size ^a	Adjusted for
BMI	Ajeganova [26]	1596	Multivariable linear regression	✓ (BMI)	Per unit baseline BMI: <i>b</i> 0.02 (0.01, 0.03)	Age, duration of follow-up, gender, ever glucocorticoid use, ever biologic use
Erosions	Multivariable analyses					
	Malm [32]	1387	Logistic regression (HAQ cut-off = 0.75)	✗ (x-ray erosions)	Baseline erosions yes vs. no: OR 1.24 (0.93, 1.66)	Age, gender, disease duration
	Bansback [44]	985	Logistic regression (HAQ cut-off = 1.5)	✗ (Larsen score)	Per unit baseline Larsen Score: OR 1.01 (<i>p</i> = 0.20)	Carstairs deprivation index; baseline: functional grade, HAQ, haemoglobin level; Year 1: functional grade, HAQ, DAS28
	Benton [27]	42	Logistic regression (HAQ cut-off = 0.25) OR are the odds of being in the low HAQ group	✗ (Sharp score)	Per unit baseline Sharp Score: OR 0.96 (0.84, 1.08)	Baseline: DAS, Ritchie index, CRP, HAQ; one year: DAS, Ritchie, CRP, HAQ
	Univariable analyses					
	Combe [34]	191	Spearman correlation	✓ (Sharp score)	ρ 0.17, <i>p</i> = 0.04	–
	Jäntti [33]	121	Somers' <i>d</i>	✗ (Larsen score)	<i>d</i> = 0.01 (–0.62, 0.26)	–
Genetic factors	Multivariable analyses					
	Ling [41]	2158	Generalised Linear Latent and Mixed Models Comparison between amino acid at a particular position vs. all other amino acids at that position	✓ (Amino acids at HLA-DR4)	Valine 11 <i>b</i> 0.02 (0.00, 0.04) Proline 11 <i>b</i> –0.03 (–0.07, 0.00) Serine 11 <i>b</i> –0.02 (–0.04, 0.00) Arginine 71 <i>b</i> 0.02 (0.00, 0.04) Alanine 71 <i>b</i> –0.06 (–0.09, –0.02) Glutamic acid 71 <i>b</i> –0.06 (–0.09, –0.03) Other positions not significant	Age at symptom onset and disease duration at follow-up
	Kroot [38]	273	Multiple regression	✗ (HLA-DR4)	HLA-DR4+ vs. HLA-DR4–: <i>b</i> < 0.001, <i>p</i> = NS	Age, gender, RF, DAS, ACPA
	Lindqvist [54]	183	Stepwise logistic regression (HAQ cut-off = 1.0)	✗ (HLA-DRB alleles)	NS – coefficients and confidence interval not reported	Age, gender, RF, HAQ, ESR, active joint count
	Kuuliala [48]	85	Logistic regression (HAQ cut-off = 0.9)	✗ (shared epitope)	Shared epitope: None OR 1 (ref) single copy OR 0.41 (0.07, 2.35) double copy OR 1.14 (0.21, 6.23)	Age, gender, sE-selectin, RF
	Univariable analyses					
	Naseem [42]	843	Mann-Whitney	✗ (PTPN22)	No association between PTPN22 SNPs and HAQ at 5 years	–
	Combe [34]	191	Kruskal-Wallis test	✗ (HLA-DRB1)	NS – coefficients and confidence interval not reported	–
	Jäntti [33]	121	Somers' <i>d</i>	✗ (HLA-B27)	HLA-B27+ vs HLA-B27–: <i>d</i> = –0.01 (–0.33, 0.31)	–
	Eberhardt [52]	99	Wilcoxon/Mann-Whitney <i>U</i> test	✗ (HLA-DRB1/DQB antigens)	NS difference in HAQ	–
Immigrant status	Andersson [43]	1430	Mann-Whitney <i>U</i>	✓ (Immigrant)	Mean HAQ at 5 years: Immigrants 0.69 Non-immigrants 0.56 <i>p</i> = 0.04	–
Morning stiffness	Verstappen [36]	112	Logistic regression (HAQ cut-off = 1)	✓ (morning stiffness)	Per minute baseline morning stiffness: OR 1.01 (1.00, 1.02)	Disease duration (natural log transformed)
	Combe [34]	191	Spearman correlation	✓ (morning stiffness)	ρ 0.21, <i>p</i> = 0.05	–
	Jäntti [33]	121	Somers' <i>D</i>	✓ (morning stiffness)	<i>d</i> = 0.28 (0.00, 0.55)	–

Table 7 (continued)

Predictor	Study	N	Analysis method	Associated with HAQ	Effect size ^a	Adjusted for	
Other Biomarkers	Multivariable analyses						
	Humphreys [40]	1995	Generalised estimating equations analysis	✓ (anti-CarP)	Anti-CarP+ vs. anti-CarP-: <i>b</i> 0.12 (0.02, 0.21)	Age, gender, smoking status, polynomials of disease duration, year of recruitment	
	Kuuliala [48]	85	Logistic regression (HAQ cut-off = 0.9)	✓ (sE-selectin)	Level of sE-selectin: 1st tertile OR 1 (ref) 2nd tertile OR 2.45 (0.70, 8.59) 3rd tertile OR 4.18 (1.15, 15.22)	Age, gender, shared epitope, RF	
	Univariable analyses						
	Manivel [47]	773	ANOVA	✗ (Anti-CII)	NS difference between Anti-CII+ and Anti-CII- patients	–	
	Forslind [46]	92	Mann-Whitney U	✗ (AFA)	NS difference between AFA+ and AFA- patients	–	
	Genevay [23]	25	Mann-Whitney U	✗ (APF) ✗ (AKA)	Mean HAQ at follow-up: APF+ 0.94 APF- 0.75 <i>p</i> = NS AKA+ 0.82 AKA- 0.78 <i>p</i> = NS	–	
	Reproductive factors	Camacho [29]	1872	Linear random effects model	✓ (parous)	Parous vs. nulliparous at baseline: <i>b</i> -0.19 (-0.34, -0.05)	Age, disease duration, SES, smoking RF, ACPA, comorbidities, ACR RA criteria
		Kuiper [45]	332	Student's t test	✓ (menopause)	Postmenopausal women had higher HAQ than premenopausal women (<i>p</i> < 0.01)	–
	Socioeconomic status	Bansback [44]	985	Logistic regression (HAQ cut-off = 1.5)	✓ (Carstairs deprivation index)	Carstairs index: 1 OR 1.0 (ref) 2 OR 0.78 (<i>p</i> = 0.44) 3 OR 1.44 (<i>p</i> = 0.24) 4 OR 1.73 (<i>p</i> = 0.08) 5 OR 1.98 (<i>p</i> = 0.04)	Baseline: HAQ, functional grade, haemoglobin level, Larsen score; year 1: functional grade, HAQ, DAS28
Eberhardt [28]		63	Logistic regression (HAQ cut-off = 1.0); education groups (Years): 0–9, 10–11, ≥12	✓ (Years of education)	Per education group change: OR 0.87 (<i>p</i> = 0.05)	“[demographic,] clinical, radiographic and laboratory data”	

See Table 2 for acronym definitions: ACPA, BMI, CRP, DAS28, ESR, HAQ, RF, SES.

See Table 3 for acronym definitions: *b*, NS, OR, RA.

See Table 4 for acronym definitions: *ρ*.

^a Brackets indicate 95% confidence interval unless otherwise stated; ACR = American College of Rheumatology, AFA = antiferlaggrin antibodies, AKA = antikeratin antibody, ANOVA = analysis of variance, Anti-CarP = anticarbamylated protein antibodies, Anti-CII = anticollagen type II antibodies, APF = antiperinuclear factor, sE-selectin = soluble E-selectin, SNP = single nucleotide polymorphism.

subsequent long-term functional disability, as measured by the HAQ, in patients with inflammatory arthritis. There was highly consistent evidence of an association between female gender, higher baseline age and higher baseline HAQ score, with subsequent higher HAQ scores. There was moderately consistent evidence of an association between higher baseline pain, DAS28 and morning stiffness and subsequent increased HAQ score. However in general, studies reported weak or no association between higher baseline swollen joint count, erosions, HLA genetic variations or RF positivity with later HAQ scores. The literature is equivocal regarding the relationship between ACPA positivity and subsequent HAQ scores.

The findings of this review are in agreement with a review carried out by Scott et al. in 2003, which reported that women and those of older age at baseline were more likely to have high disability in the future [16]. Scott et al. also found that higher pain at baseline was associated with higher subsequent disability. However, Scott et al. reported that RF positivity and a high number of erosions were associated with increased disability at follow-up. The association between more erosions and subsequent higher HAQ score was also reported in a review by Bombardier et al. [59]. This is likely to be because both of these previous reviews included patients with any disease duration, whilst the current review only included studies confined to early arthritis patients (symptom

duration ≤2 years at baseline) who may not yet have developed erosions.

Baseline HAQ score was the only variable that was shown to be associated with higher HAQ score at follow-up consistently across all studies assessing the relationship (with nine studies reporting a significant association and one study trending towards significance). Higher levels of pain and morning stiffness at baseline may also be useful predictors of subsequent higher HAQ score, although the evidence for this is weaker. Furthermore, four out of five studies assessing the relationship reported a significant relationship between baseline DAS28 and later functional disability. However, the longest follow-up of these studies was six years.

Also of clinical interest are the results of studies assessing the association between RF and ACPA positivity and later HAQ scores. None of the three large cohort studies with over 1000 patients at baseline reported a significant association between RF positivity and later higher HAQ scores and only one study out of six reported a significant association between ACPA positivity at baseline and later higher HAQ scores. However, this was by far the largest study to assess the association, including almost 2000 patients in the analysis [40].

This review has a number of strengths. Limiting the review to studies of patients with early arthritis allows us to examine which factors early in the disease process predict later functional

disability. Furthermore, we have stratified the presentation of results into multivariable and univariable analyses, and then sorted within these sets based on the sample size of the studies. Therefore analyses with high power which control for confounding are presented first, allowing the reader to easily assess the quality of the studies presented.

A drawback to this review is that a meta-analysis could not be performed due to the heterogeneity between the studies. Almost every study assessed the association between baseline variables and subsequent HAQ scores in a different way, using different analysis techniques and controlling for different combinations of covariates. Any meta-analysis combining these studies would be uninterpretable. Furthermore, we have included all studies published since 1990 that met the inclusion criteria. Thus secular trends in disease severity could be influencing the results of the review [60,61] or differences in the available treatments and treatment strategies over time may mean that studies published over this period are not comparable.

The majority of studies included within the review were judged to be of moderate quality (Supplementary File 2). Studies often did not report on the amount of missing data. Other studies used complete case analyses, which could mean that the results of the studies are biased. Furthermore, studies often included covariates in analyses but only reported on the primary predictor defined in the research question. Therefore, these covariates could not be included in the review, despite contributing to the analyses.

In conclusion, this review has demonstrated that female gender and higher baseline age, HAQ score, pain score and duration of morning stiffness have been consistently reported to predict long-term increased functional disability. Furthermore, most studies assessing the association reported no association between RF and erosion and early IA patients' long-term disability. This study indicates the relative importance of patient reported outcomes over blood test results in predicting the long-term prognosis (in terms of physical disability) of patients with IA.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.semarthrit.2018.03.004>.

References

- [1] Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. *Lancet* 2010;376:1094–108.
- [2] Humphreys JH, Warner A, Chipping J, et al. Mortality trends in patients with early rheumatoid arthritis over 20 years: results from the Norfolk Arthritis Register. *Arthritis Care Res (Hoboken)* 2014;66:1296–301.
- [3] Gonzalez A, Maradit KH, Crowson CS, et al. The widening mortality gap between rheumatoid arthritis patients and the general population. *Arthritis Rheum* 2007;56:3583–7.
- [4] Ahlmen M, Svensson B, Albertsson K, Forslind K, Hafstrom I. Influence of gender on assessments of disease activity and function in early rheumatoid arthritis in relation to radiographic joint damage. *Ann Rheum Dis* 2010;69:230–3.
- [5] van Tuyl LH, Boers M, Lems WF, et al. Survival, comorbidities and joint damage 11 years after the COBRA combination therapy trial in early rheumatoid arthritis. *Ann Rheum Dis* 2010;69:807–12.
- [6] Norton S, Fu B, Scott DL, et al. Health Assessment Questionnaire disability progression in early rheumatoid arthritis: systematic review and analysis of two inception cohorts. *Semin Arthritis Rheum* 2014;44:131–44.
- [7] Wiles NJ, Scott DG, Barrett EM, et al. Benchmarking: the five year outcome of rheumatoid arthritis assessed using a pain score, the Health Assessment Questionnaire, and the Short Form-36 (SF-36) in a community and a clinic based sample. *Ann Rheum Dis* 2001;60:956–61.
- [8] Courvoisier N, Dougados M, Cantagrel A, et al. Prognostic factors of 10-year radiographic outcome in early rheumatoid arthritis: a prospective study. *Arthritis Res Ther* 2008;10:R106.
- [9] Steinbrocker O, Traeger C, Batterman RC. Therapeutic criteria in rheumatoid arthritis. *J Am Med Assoc* 1949;140:659–62.
- [10] Liang MH, Jette AM. Measuring functional ability in chronic arthritis: a critical review. *Arthritis Rheum* 1981;24:80–6.
- [11] Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137–45.
- [12] Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005;23:S14–8.
- [13] Daltroy LH, Larson MG, Eaton HM, Phillips CB, Liang MH. Discrepancies between self-reported and observed physical function in the elderly: the influence of response shift and other factors. *Soc Sci Med* 1999;48:1549–61.
- [14] Redelmeier DA, Lorig K. Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology. *Arch Intern Med* 1993;153:1337–42.
- [15] Pope JE, Khanna D, Norrie D, Ouimet JM. The minimally important difference for the health assessment questionnaire in rheumatoid arthritis clinical practice is smaller than in randomized controlled trials. *J Rheumatol* 2009;36:254–9.
- [16] Scott DL, Smith C, Kingsley G. Joint damage and disability in rheumatoid arthritis: an updated systematic review. *Clin Exp Rheumatol* 2003;21:S20–7.
- [17] Toussiroit E. Predictive factors for disability as evaluated by the health assessment questionnaire in rheumatoid arthritis: a literature review. *Inflamm Allergy Drug Targets* 2010;9:51–9.
- [18] Ropes MW, Bennett G, Cobb S, Jacox R, Jessar R. 1958 revision of diagnostic criteria for rheumatoid arthritis. *Arthritis Rheum* 1958;2:16–20.
- [19] Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
- [20] 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;62:2569–81.
- [21] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097.
- [22] Pasma A, van't Spijker A, Hazes JM, Busschbach JJ, Luime JJ. Factors associated with adherence to pharmaceutical treatment for rheumatoid arthritis patients: a systematic review. *Semin Arthritis Rheum* 2013;43:18–28.
- [23] Genevay S, Hayem G, Verpillat P, Meyer O. An eight year prospective study of outcome prediction by antiperinuclear factor and antikeratin antibodies at onset of rheumatoid arthritis. *Ann Rheum Dis* 2002;61:734–6.
- [24] Camacho EM, Verstappen SM, Lunt M, Bunn DK, Symmons DP. Influence of age and sex on functional outcome over time in a cohort of patients with recent-onset inflammatory polyarthritis: results from the Norfolk Arthritis Register. *Arthritis Care Res (Hoboken)* 2011;63:1745–52.
- [25] Contreras-Yanez I, Pascual-Ramos V. Window of opportunity to achieve major outcomes in early rheumatoid arthritis patients: how persistence with therapy matters. *Arthritis Res Ther* 2015;17:177.
- [26] Ajeganova S, Andersson ML, Hafstrom I. Association of obesity with worse disease severity in rheumatoid arthritis as well as with comorbidities: a long-term followup from disease onset. *Arthritis Care Res (Hoboken)* 2013;65:78–87.
- [27] Benton N, Stewart N, Crabbe J, Robinson E, Yeoman S, McQueen FM. MRI of the wrist in early rheumatoid arthritis can be used to predict functional outcome at 6 years. *Ann Rheum Dis* 2004;63:555–61.
- [28] Eberhardt KB, Fex E. Functional impairment and disability in early rheumatoid arthritis—development over 5 years. *J Rheumatol* 1995;22:1037–42.
- [29] Camacho EM, Harrison M, Farragher TM, et al. Parity, time since last live birth and long-term functional outcome: a study of women participating in the Norfolk Arthritis Register. *Ann Rheum Dis* 2011;70:642–5.
- [30] Nair SC, Bijlsma JW, van der Werf JH, et al. Do radiographic joint damage and disease activity influence functional disability through different mechanisms? Direct and indirect effects of disease activity in established rheumatoid arthritis. *J Rheumatol* 2013;40:1505–12.
- [31] Welsing PM, Borm GF, van RP. Minimal clinically important difference in radiological progression of joint damage. A definition based on patient perspective. *J Rheumatol* 2006;33:501–7.
- [32] Malm K, Bergman S, Andersson M, Bremander A. Predictors of severe self-reported disability in RA in a long-term follow-up study. *Disabil Rehabil* 2015;37:686–91.
- [33] Jantti JK, Kaarela K, Luukkainen RK, Kautiainen HJ. Prediction of 20-year outcome at onset of seropositive rheumatoid arthritis. *Clin Exp Rheumatol* 2000;18:387–90.
- [34] Combe B, Cantagrel A, Goupille P, et al. Predictive factors of 5-year health assessment questionnaire disability in early rheumatoid arthritis. *J Rheumatol* 2003;30:2344–9.
- [35] Hallert E, Bjork M, Dahlstrom O, Skogh T, Thyberg I. Disease activity and disability in women and men with early rheumatoid arthritis (RA): an 8-year followup of a Swedish early RA project. *Arthritis Care Res (Hoboken)* 2012;64:1101–7.
- [36] Verstappen SM, Jacobs JW, Huisman AM, van Rijthoven AW, Sokka T, Bijlsma JW. Functional Health Assessment Questionnaire (HAQ) and Psychological HAQ are associated with and predicted by different factors in rheumatoid arthritis. *J Rheumatol* 2007;34:1837–40.
- [37] Wiles N, Dunn G, Barrett E, Silman A, Symmons D. Associations between demographic and disease-related variables and disability over the first five

- years of inflammatory polyarthritis: a longitudinal analysis using generalized estimating equations. *J Clin Epidemiol* 2000;53:988–96.
- [38] Kroot EJ, de Jong BA, van Leeuwen MA, et al. The prognostic value of anti-cyclic citrullinated peptide antibody in patients with recent-onset rheumatoid arthritis. *Arthritis Rheum* 2000;43:1831–5.
- [39] Koevoets R, Dirven L, Klarenbeek NB, et al. Insights in the relationship of joint space narrowing versus erosive joint damage and physical functioning of patients with RA. *Ann Rheum Dis* 2013;72:870–4.
- [40] Humphreys JH, Verheul MK, Barton A, et al. Anticarbamylated protein antibodies are associated with long-term disability and increased disease activity in patients with early inflammatory arthritis: results from the Norfolk Arthritis Register. *Ann Rheum Dis* 2016;75:1139–44.
- [41] Ling SF, Viatte S, Lunt M, et al. HLA-DRB1 amino acid positions 11/13, 71, and 74 are associated with inflammation level, disease activity, and the health assessment questionnaire score in patients with inflammatory polyarthritis. *Arthritis Rheumatol* 2016;68:2618–28.
- [42] Naseem H, Thomson W, Silman A, Worthington J, Symmons D, Barton A. The PTPN22_{C1858T} functional polymorphism is associated with susceptibility to inflammatory polyarthritis but neither this nor other variants spanning the gene is associated with disease outcome. *Ann Rheum Dis* 2008;67:251–5.
- [43] Andersson ML, Bergman S, Soderlin MK. The effect of socioeconomic class and immigrant status on disease activity in rheumatoid arthritis: data from BARFOT, a multi-centre study of early RA. *Open Rheumatol J* 2013;7:105–11.
- [44] Bansback N, Young A, Brennan A, Dixey J. A prognostic model for functional outcome in early rheumatoid arthritis. *J Rheumatol* 2006;33:1503–10.
- [45] Kuiper S, van Gestel AM, Swinkels HL, de Boo TM, da Silva JA, van Riel PL. Influence of sex, age, and menopausal state on the course of early rheumatoid arthritis. *J Rheumatol* 2001;28:1809–16.
- [46] Forslind K, Vincent C, Serre G, Svensson B. Antiflaggrin antibodies in early rheumatoid arthritis may predict radiological progression. *Scand J Rheumatol* 2001;30:221–4.
- [47] Manivel VA, Mullazehi M, Padyukov L, et al. Anticollagen type II antibodies are associated with an acute onset rheumatoid arthritis phenotype and prognosticate lower degree of inflammation during 5 years follow-up. *Ann Rheum Dis* 2017;76:1529–36.
- [48] Kuuliala A, Eberhardt K, Takala A, Kautiainen H, Repo H, Leirisalo-Repo M. Circulating soluble E-selectin in early rheumatoid arthritis: a prospective five year study. *Ann Rheum Dis* 2002;61:242–6.
- [49] Bjork MA, Thyberg IS, Skogh T, Gerdle BU. Hand function and activity limitation according to health assessment questionnaire in patients with rheumatoid arthritis and healthy referents: 5-year followup of predictors of activity limitation (The Swedish TIRA Project). *J Rheumatol* 2007;34:296–302.
- [50] Burr ML, Viatte S, Bukhari M, et al. Long-term stability of anti-cyclic citrullinated peptide antibody status in patients with early inflammatory polyarthritis. *Arthritis Res Ther* 2012;14:R109.
- [51] Combe B, Rincheval N, Benessiano J, et al. Five-year favorable outcome of patients with early rheumatoid arthritis in the 2000s: data from the ESPOIR cohort. *J Rheumatol* 2013;40:1650–7.
- [52] Eberhardt K, Fex E, Johnson U, Wollheim FA. Associations of HLA-DRB and -DQB genes with two and five year outcome in rheumatoid arthritis. *Ann Rheum Dis* 1996;55:34–9.
- [53] Kapetanovic MC, Lindqvist E, Nilsson JA, Geborek P, Saxne T, Eberhardt K. Development of functional impairment and disability in rheumatoid arthritis patients followed for 20 years: relation to disease activity, joint damage, and comorbidity. *Arthritis Care Res (Hoboken)* 2015;67:340–8.
- [54] Lindqvist E, Saxne T, Geborek P, Eberhardt K. Ten year outcome in a cohort of patients with early rheumatoid arthritis: health status, disease process, and damage. *Ann Rheum Dis* 2002;61:1055–9.
- [55] Thyberg I, Dahlstrom O, Bjork M, Arvidsson P, Thyberg M. Potential of the HAQ score as clinical indicator suggesting comprehensive multidisciplinary assessments: the Swedish TIRA cohort 8 years after diagnosis of RA. *Clin Rheumatol* 2012;31:775–83.
- [56] Welsing PM, van Gestel AM, Swinkels HL, Kiemeneij LA, van Riel PL. The relationship between disease activity, joint destruction, and functional capacity over the course of rheumatoid arthritis. *Arthritis Rheum* 2001;44:2009–17.
- [57] Wiles NJ, Dunn G, Barrett EM, Harrison BJ, Silman AJ, Symmons DP. One year followup variables predict disability 5 years after presentation with inflammatory polyarthritis with greater accuracy than at baseline. *J Rheumatol* 2000;27:2360–6.
- [58] Woolf AD, Hall ND, Goulding NJ, et al. Predictors of the long-term outcome of early synovitis: a 5-year follow-up study. *Br J Rheumatol* 1991;30:251–4.
- [59] Bombardier C, Barbieri M, Parthan A, et al. The relationship between joint damage and functional disability in rheumatoid arthritis: a systematic review. *Ann Rheum Dis* 2012;71:836–44.
- [60] Diffin JG, Lunt M, Marshall T, Chipping JR, Symmons DP, Verstappen SM. Has the severity of rheumatoid arthritis at presentation diminished over time? *J Rheumatol* 2014;41:1590–9.
- [61] Nikiphorou E, Norton S, Carpenter L, et al. Secular changes in clinical features at presentation of rheumatoid arthritis: increase in comorbidity but improved inflammatory states. *Arthritis Care Res (Hoboken)* 2017;69:21–7.