Cohort profile: SCREE-N-RA: design, methods and perspectives of a Swiss cohort study of first-degree relatives of patients with rheumatoid arthritis

Benoit Thomas P Gilbert,1 Céline Lamacchia,3 Denis Mongin,1 Kim Lauper,1,2 Eric Trunk,1 Olivia Studer,1 Delphine S Courvoisier,1,2 Andrea Rubbert-Roth,3 Diego Kyburz,4 Burkhard Moeller,5 Axel Finckh1

ABSTRACT

Purpose: Rheumatoid arthritis (RA) is an insidious autoimmune disease, with an immunological onset years before diagnosis. Early interventions in preclinical stages could prevent or minimise the progression towards irreversible joint damage. The SCREEN-RA cohort (Evaluation of a SCREENing strategy for Rheumatoid Arthritis) aims to characterise the preclinical stages of the disease, to identify environmental risk factors, and to discover or validate novel biomarkers predictive for RA development.

Participants: SCREEN-RA includes an at-risk population for RA, namely first-degree relatives of patients with established RA.

Findings to date: The cohort started in 2009 is composed of mostly asymptomatic healthy individuals (total n=1458, 7262 person-years), with a mean age of 44 years at enrolment, 74% female and 91% Caucasian ethnicity. During the study period, 16 participants have developed RA. All participants provide baseline serum, DNA and RNA samples, and in a subset, stool samples and oral examination are performed for microbiota assessment. At enrolment, 10% of participants had asymptomatic autoimmunity associated with RA (n=147), 10% presented ‘clinically suspect arthralgias’ (n=143) and 3% reported arthralgias in conjunction with autoimmunity or high genetic risk (n=51). Studies with this cohort have uncovered risk factors for RA development, such as female hormonal factors, poor oral health or intestinal dysbiosis.

Future plans: Future directions include immunological and ‘multiomics’ approaches to discover new biological markers of progression towards RA, as well as testing preventive interventions in ‘high-risk’ population.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic autoimmune disease leading to joint destruction and extra-articular manifestations. RA has a rising prevalence1 of 0.5%–1% in the European and North American population.2,3 Important risk factors include genetics,4–6 female hormonal factors7 and environmental factors such as air pollution,8 diet and obesity9–14 or stressful events.15 The risk of RA is also strikingly associated with smoking,16–18 but only in conjunction with specific human leukocyte antigen (HLA) alleles (the so-called ‘shared epitope’), implying a strong gene–environment interaction.19–21 Recent investigations have suggested a ‘mucosal origin’ of RA autoimmunity,22 because of its remarkable association with periodontal disease22–25 and other mucosal inflammatory conditions, such as chronic intestinal conditions,26 or chronic pulmonary disorders.27–28 Underlying dysbiosis is suspected to play a key role in the development of RA,29–34 even if exact causality still remains to be determined.

The aetiology of RA is believed to result from a multistep process, where environmental factors gradually initiate a pathological activation of the immune system.35 Overall, the preclinical progression toward RA can be divided into three ‘at risk stages’:36 1. Genetic and environmental risk: First-degree relatives (FDRs) of patients with
RA have a 3–5 fold increased risk of developing the disease, which is even higher in families with multiple cases of RA. Among susceptibility genes, HLA-DRB1 variants share a common sequence in the third variable region of the MHC II binding site (referred to as the ‘shared epitope’), which is involved in the response to extracellular immune ligands. However, the risk associated with established genetic markers, even in longitudinal studies, remains modest. The latter underlines the importance of environmental factors, which are thought to act as ‘triggers’. The primary objectives are to characterise the different preclinical stages of RA, and to determine the optimal combination of biomarkers to predict the development of RA within 3–5 years. Recruitment methods include emails to patients, presentations at patient conferences, articles in general audience journals, promotion via patient associations, information to patients with RA within the Swiss Clinical Quality Management Rheumatoid Arthritis (SCQM-RA) register, advertising through radio and television and advertisement in pharmacies. Since 2018, campaigns on social networks have also been organised (Facebook, Snapchat, LinkedIn and website www.arthritis-checkup.ch).

Patient and public involvement
Patient and public organisations were involved in the project, including design and management of the study. The Swiss league against rheumatic diseases has been a long-time partner, in particular helping recruiting participants and disseminating results. Also, as future research might involve preventive interventions, a random sample of SCREEN-RA participants were asked in 2016 if they would take a hypothetical treatment or not, depending on varying levels of treatment characteristics. About one-third of the participants would be willing to take a preventive treatment if the hypothetical risk of developing RA was at least 20%. Face-to-face interviews revealed that lifestyle changes and complementary medicine were also considered. Finally, most participants would agree to enrol in a randomised controlled trial to test the efficacy of preventive interventions. We took this feedback into account for our future research, and furthermore we regularly receive input from one of the members of the rheumatology division who is also a patient with RA herself (not named in the article).

Study population
The primary study population is a genetically defined at risk population, namely FDRs of established patients with RA. The study population also comprises a minority of FDRs of patients with lupus or other connective tissue diseases, autoimmune thyroiditis or type 1 diabetes. Indeed, because of shared genetic risk factors with RA, all these conditions increase the risk of RA among FDRs in a similar magnitude. Other inclusion criteria are the absence of clinically apparent active synovitis on examination, and an age of at least 18 years. Exclusion criteria are an established diagnosis of RA, or the presence of active concomitant inflammatory arthritides (ie, patients with psoriatic arthritis, spondyloarthritis or known microcrystalline arthritis) to avoid outcome misclassification. After enrolment, all participants are followed using yearly questionnaires to detect new symptoms or signs of the disease (figure 1).

We use a combination of known risk factors for RA and clinical parameters to define groups of ‘high-risk’ participants (figure 2). These ‘high-risk’ participants satisfy at least one of the following criteria:

- Having 2 copies of the shared epitope, which doubles the risk of RA compared with having one single copy.

Cohort description
Study overview
The SCREEN-RA study is a multicentric observational cohort study across Switzerland. It enrols and follows FDRs of patients with RA and was started in 2009 with the support of the Swiss National Science Foundation. The primary objectives are to characterise the different preclinical stages of RA, and to determine the optimal preclinical phases: Asymptomatic autoimmunity can evolve over several years, towards inflammatory arthralgias, or undifferentiated arthritis, before finally leading to clinically-apparent RA. These symptomatic ‘pre-RA’ patients can be identified using specific questionnaires and/or physical examination. In particular, the European League Against Rheumatism (EULAR) has proposed clinical characteristics of arthralgias at risk for RA, namely ‘clinically suspect arthralgia’ (CSA), which increase the risk of developing RA during a 2-year follow-up.

The preclinical phases of RA represent opportunities for preventive interventions, which may allow to avert disease development or improve long-term outcomes. However, the optimal screening strategy to identify ‘at-risk’ individuals most likely to benefit from early interventions is still to be established.

To adequately define the specific preclinical phases of RA development, and to identify environmental factors driving progression from one phase to the other, longitudinal studies are required. In this article, we present a cohort study of FDRs of patients with RA, including 1458 participants.
Having serological antibodies associated with RA, which strongly increases the risk of developing RA among FDRs: ACPA seropositivity, or RF levels (either IgA or IgM isotype) three times the upper limit of the norm or anti-Ra33 antibodies three times the upper limit of the norm (IgM, IgG or IgA). The simultaneous presence of several autoantibodies above the upper limit of normal is also considered high risk for future RA development.

Having ‘CSA’ defined when satisfying four or more of the seven criteria previously validated by EULAR (ie, symptom duration <1 year, symptoms in metacarpophalangeal joints, morning stiffness duration ≥60 min, most severe symptoms in early morning, being RA-FDR, difficulty with making a fist, and positive squeeze test of metacarpophalangeal joints). ‘Undifferentiated arthritis’ was defined as one or more swollen joints on examination, in conjunction with ‘CSA’. Undifferentiated arthritis should not be classifiable as rheumatic disease, nor result from a septic or crystal aetiology. If correctly assessed, subsequent risk for RA development in the following year has been reported as high as 35%.

‘High-risk’ participants are then followed up more closely in this cohort, with a yearly in person visit and...
blood sampling, to monitor evolution of serological markers and symptoms over time.

Sample size calculation
The SCREEN-RA sample size was estimated based on the number of FDRs developing RA, to allow predictive modelling of RA in FDRs. We estimated that a minimum of 60 incident cases of RA would be needed to analyse with sufficient discriminative power a predictive model of RA in FDRs. Unaffected FDRs in multiply affected families have an incidence of RA of 8/1000 patient-years (95% CI: 4.2 to 13.6), and lower in families with only a single affected case. Patients develop autoantibodies on average 2–5 years prior to disease onset, which implies that we expect to detect autoimmunity associated with RA in up to 4% of FDRs, which is approximately what has been described in similar populations. With a minimum of 5 years of follow-up, an estimated incidence rate of RA between 0.6 and 1 case/person-year, the sample size required to characterise ~60 patients with RA was estimated to be between 1000 and 2000 individuals. To ensure the feasibility of such a long-term longitudinal follow-up while minimising costs, the study was designed with a yearly follow-up.

Study sites
Enrolment is conducted within 10 collaborative centres, within the following cities: Geneva (Hôpitaux Universitaires de Genève), Lausanne (Centre Hospitalier Universitaire Vaudois), Fribourg (Hôpital Fribourgeois), Neuchâtel (Réseau Hospitalier Neuchâtelois), Bâle (Universitätsspital Basel), Zurich (Universitätsspital Zurich), Berne (Inselspital-Hôpital universitaire de Berne), Aarau (Kantonsspital Aarau) and Saint-Gall (Kantonsspital St.Gallen). Figure 3 represents the geographical distribution of enrolled population across the involved Swiss cantons.

Questionnaires
Inclusion questionnaire
At inclusion, participants complete a questionnaire regarding demographic data and environmental factors such as alcohol consumption, nutritional habits, smoking status, infectious diseases, professional exposures, oral health, female hormonal factors and family history of autoimmune disease (table 1).

Longitudinal follow-up questionnaire
Participants receive a yearly follow-up questionnaire assessing articular symptoms, presence of immune disease, current medication and environmental factors such as smoking, nutritional or exercising activity.

Table 1 Summary of questionnaire content for participants, SCREEN-RA cohort, Switzerland, 2009–2020

| Questionnaire component | Summary of content |
|-------------------------|--------------------|
| General information     | Contact information, age, gender, ethnicity, country of origin, birth weight, years of education |
| Physical examination    | Absence of systemic inflammatory disease, height, weight, blood pressure, number of swollen joints, no of tender joints, presence of rheumatoid nodules. |
| Family information      | Family’s country of origin, number of relatives with RA or other autoimmune conditions, number of siblings, no of children, relation to the RA-diagnosed relative, age of beginning of symptoms, age of diagnosis, anti-CCP testing, medication of the RA-relative. |
| Annual follow-up questionnare | Joint pain assessment, joint swelling assessment, recent blood test for RA for Lupus, current health issues, current medication, history of infectious disease, history of female hormonal factors, vaccinal status, smoking status, consumption of tea/coffee/soft-drinks, use of vitamin supplementation, alcohol consumption, professional situation, sleeping disorders, physical activity. |
| Annual follow-up questionnare (optional) | Dust exposition at workplace, professional health, oral health, consumption of seafoods. |
| In case of stool sample (optional) | Time since last defecation, stool consistency, recent travel, use of probiotics, use of antibiotics, recent surgery, current periodontitis. A Food Frequency Questionnaire. |
| In case of oral sample (optional) | Allergies, current medication, smoking status, number of teeth, reason for tooth loss, no of implants, oral hygiene habits, chewing problems, breath problems, periodontal status. |

CCP, cyclic citrullinated peptide; RA, rheumatoid arthritis; SCREEN-RA, Evaluation of a SCREENing strategy for Rheumatoid Arthritis.
habit (table 1). Questionnaires are available in three languages (French, English and German). Questionnaires have been established in collaboration with other ongoing studies of at-risk populations, such as the American SÉRA cohort to allow replication studies in the future.

Clinical visits
At inclusion, a clinical examination is performed by a specialised nurse or a rheumatologist to assess potential tender and swollen joints and rule out the presence of RA or other autoimmune conditions. This examination is repeated yearly for the ‘high-risk’ participants, in addition to biological sampling.

Biological samples
Blood samples
Full blood samples are collected at inclusion in EDTA collection tubes for genetic testing (HLA; online supplemental file) and additional aliquots for a genomic DNA library. Full blood is further used to collect total RNA using Tempus Blood RNA Tubes (lyses whole blood cells and stabilises RNA). Genomic DNA and total RNA are isolated by standard procedures. Serum samples are collected for the assessment of autoantibodies (ACPA, RF, and anti-Ra-33 in a subset of participants) using commercially, as well as non-commercially assays (online supplemental file). Aliquots are stored at −80°C in a serum library. Participants deemed at high risk provide yearly new blood samples, while other participants provide a baseline sample.

Stool samples
A nested case–control study in 2016 was performed, with 133 stool samples. A new collection is ongoing (2019–2020; targeted n=400 stool samples), using collection-devices allowing the creation of several aliquots. Participants receive a stool collection kit and proceed to sampling at home. They temporarily freeze the fresh sample at −20°C, and bring it in a cooler box to the study centre, where the stool samples are stored at −80°C, without any additive according to published methods.

Salivary/dental plaque samples
In a subset of the cohort (n=99), gingival crevicular fluid is collected at one site in each dentition quadrant using membrane strips. The salivary microbiome is sampled collecting unstimulated saliva by spitting in a sterile plastic tube. Finally, the subgingival microbiome is sampled using sterile paper points inserted into the bottom of the pockets, at four different oral sites.

Sample storage and biobank
All biological samples are processed following standard operative procedures and stored at −80°C, in a dedicated biobank. Samples from collaborative centres are regularly shipped on dry ice to the Geneva’s main biobank. Table 2 presents the repartition of all available serum, DNA and RNA samples by baseline risk-subgroups. A total of 2301 serum samples were collected during the study period. Each serum sample is divided into 7–9 aliquots (total n=12390 aliquots). Twenty-eight per cent of participants have at least two sequential samples (mean interval between samples=2.8 years). Moreover, most serum samples are matched with RNA and DNA samples (table 2). Concentration and RNA Integrity Number of RNA samples are available, as well as concentration and 260/280 optical density ratios for DNA samples. A total of 159 participants have at least two sequential RNA samples, allowing future transcriptomic longitudinal studies. In addition, matching of RNA and DNA samples (n=1396, table 2) will be useful for future expression quantitative trait loci analysis. All matched biological samples will also allow studies of predictive associations of biomarkers, combining serological, genomic and transcriptomic information into RA risk-scores.

Data management
Data are collected through a secured online interface. Since late 2019, data are stored and monitored using Research Electronic Data Capture (REDCap) software and hosted on institutional servers, with secure backup. Previously filled-in on paper versions, questionnaires are now sent by email, and reports of physical examination or serological analysis are entered into REDCap. For external data manipulation, each patient is identified

| Baseline risk-group | Participants (n subjects) | Serum samples (n samples) | At least two sequential serum samples (n subjects) | DNA samples (n samples) | RNA samples (n samples) | Matched RNA and DNA (n samples) |
|---------------------|--------------------------|----------------------------|-----------------------------------------------|------------------------|------------------------|---------------------------------|
| 1                   | 1006                     | 1293                       | 171                                           | 1060                   | 839                    | 835                             |
| 2                   | 80                       | 242                        | 59                                            | 152                    | 140                    | 140                             |
| 3                   | 147                      | 379                        | 92                                            | 227                    | 213                    | 213                             |
| 4                   | 143                      | 222                        | 57                                            | 149                    | 125                    | 124                             |
| 5                   | 51                       | 133                        | 32                                            | 82                     | 84                     | 84                              |

Baseline risk-groups as defined in figure 2. SCREEN-RA, Evaluation of a SCREENing strategy for Rheumatoid Arthritis.
by a numerical code of 2–4 digits, which is also used to label the biological samples. The database is password protected and changes are tracked in logfiles.

**Statistical analysis**

A descriptive analysis of baseline data was performed (tables 3–5). Continuous variables are expressed as means with standard deviation (SD) whereas categorical variables are described using frequencies (percentage). \( \chi^2 \) test, or Fisher’s exact test for small size samples, were used to compare categorical variables. Continuous variables were compared between groups using Student’s t test, or Kruskal-Wallis test if not normally distributed or ANOVA if more than two groups. Two-tailed values of \( p<0.05 \) were considered significant. Missing data were imputed using value found in the nearest time point in a window of 6 months, when available. Missing data for RF, shared epitope and ACPA status were imputed as last observation carried forward. Incomplete records (ie, participant who never came to inclusion visit or finally refused blood sampling) were excluded. All analyses were conducted using R, V.3.6.2, with package tableone.

**Baseline characteristics and evolution of the SCREEN-RA population**

**Whole study population**

On 23 November 2020, SCREEN-RA cohort had enrolled 1458 individuals, 1261 of whom are still actively providing follow-up data. The total follow-up duration equals 7762 patient-years, which represents an average of 5 years of follow-up per participant. The population had a mean age at enrolment of 44 years, was 74% female and 91% from white ethnicity (table 3). The main reason for study discontinuation was loss to follow-up (65%), followed by refusal to participate further (30%). The main enrolment sites were the Geneva centre (35%) and St-Gallen centre (19%). Nineteen per cent of participants were active smokers and the mean baseline body mass index was 24 kg/m\(^2\) (tables 3 and 4).

**RA-converter subjects**

During the study period, 16 participants developed a classifiable RA, after a mean follow-up of 5.5 years. They provided a total of 48 blood samples, including postdiagnostic samples. At enrolment, compared with other FDRs taken together, RA-converters were significantly more often seropositive for ACPA (38% vs 5%; \( p<0.01 \)) and RF (63% vs 19%; \( p<0.01 \)). Interestingly, frequency of shared epitope alleles among RA-converters was not significantly different from that in the general population of SCREEN-RA.

---

**Table 3** Baseline sociodemographic characteristics of participants, SCREEN-RA cohort, Switzerland, 2009–2020

| Variable                      | No (%) (total n=1458) |
|-------------------------------|-----------------------|
| **Age group, years**         |                       |
| 18–25                         | 172 (12)              |
| 25–50                         | 779 (53)              |
| 50–75                         | 494 (34)              |
| >75                           | 13 (1)                |
| **Year of education**        |                       |
| 0–5                           | 55 (4)                |
| 5–10                          | 83 (6)                |
| >10                           | 785 (54)              |
| Not specified                 | 535 (37)              |
| **Gender**                   |                       |
| Female                        | 1086 (74)             |
| White                         | 1322 (91)             |
| **Number of RA cases in participants’ family** |        |
| 1                             | 1158 (79)             |
| 2                             | 165 (11)              |
| >3                            | 51 (4)                |
| Not specified                 | 84 (6)                |
| **Tobacco smoking**          |                       |
| Never                         | 727 (50)              |
| Previous                      | 380 (26)              |
| Current                       | 283 (19)              |
| Not specified                 | 67 (5)                |

**Table 4** Baseline biological and physical characteristics of participants, SCREEN-RA cohort, Switzerland, 2009–2020

| Variable                      | No (%) (total n=1458) |
|-------------------------------|-----------------------|
| **BMI groups**                |                       |
| <18                           | 30 (2)                |
| 18 to <25                     | 892 (61)              |
| 25 to <30                     | 385 (27)              |
| ≥30                           | 127 (9)               |
| Missing                       | 24 (2)                |
| **Biology**                   |                       |
| Provided at least two blood samples | 412 (28)             |
| Total ACPA seropositivity     | 78 (5)                |
| Total RF seropositivity       | 282 (19)              |
| IgA RF seropositivity         | 66 (5)                |
| IgM RF seropositivity         | 251 (17)              |
| Total anti-Ra33 tested        | 660 (45)              |
| Anti-Ra33 seropositivity      | 8 (0.5)               |

| **Shared epitope allele no**  |                       |
| 0 copy                        | 724 (50)              |
| 1 copy                        | 564 (39)              |
| 2 copies                      | 105 (7)               |
| Not tested                    | 65 (4)                |

RA, Rheumatoid Arthritis; SCREEN-RA, Evaluation of a SCREENing strategy for Rheumatoid Arthritis; ACPA, anticitrullinated peptide antibodies; BMI, body mass index; RF, rheumatoid factors; SCREEN-RA, Evaluation of a SCREENing strategy for Rheumatoid Arthritis; ULN, upper limit of the norm.
| Variables (at enrolment) | Low risk | High risk | P value |
|-------------------------|----------|-----------|---------|
|                         | Otherwise n (%) | Otherwise n (%) | (ANOVA or $\chi^2$) |
| **Risk groups (total n=1458):** | | | |
| NA (n=31) | 1 (n=1006) | 2 (n=80) | 3 (n=147) | 4 (n=143) | 5 (n=51) | Clinically suspect arthralgia with specific autoimmunity or high genetic risk |
| Not assigned (serological result awaited) | | | | | | |
| Asymptomatic without specific autoimmunity | | | | | | |
| High genetic risk without specific autoimmunity | | | | | | |
| Asymptomatic with specific autoimmunity | | | | | | |
| Isolated clinically suspect arthralgias | | | | | | |
| Clinically suspect arthralgia with specific autoimmunity or high genetic risk | | | | | | |
| **Demographics** | | | | | | |
| Age (years) | 42 (12) | 43 (14) | 44 (12) | 45 (15) | 50 (14) | 48 (13) | <0.001 |
| Gender (female) | 68% | 72% | 78% | 75% | 86% | 84% | 0.004 |
| White ethnicity | 84% | 92% | 88% | 90% | 91% | 82% | 0.044 |
| BMI | 27 (6) | 24 (4) | 24 (4) | 24 (4) | 25 (5) | 26 (5) | 0.02 |
| Tobacco smoking | | | | | | | 0.22 |
| current | 29% | 18% | 28% | 16% | 24% | 20% | |
| previous | 26% | 25% | 33% | 31% | 26% | 24% | |
| never | 29% | 51% | 29% | 51% | 48% | 56% | |
| **Biology** | | | | | | |
| ACPA seropositivity (commercial or non-commercial assays) | 0% | 0% | 42% | 0% | 31% | <0.01 |
| RF seropositivity (IgA or IgM) at least 1 x ULN | 12% | 9% | 66% | 17% | 53% | <0.01 |
| at least 3 x ULN | 0% | 0% | 61% | 0% | 43% | |
| Anti-RA33 antibodies (3 x ULN) | 0% | 0% | 3% | 1% | 2% | 0.02 |
| HLA-SE | | | | | | | <0.01 |
| 0 copy | 55% | 0% | 52% | 54% | 35% | |
| 1 copy | 43% | 0% | 39% | 43% | 27% | |
| 2 copies | 0% | 100% | 7% | 0% | 27% | |
| Undifferentiated arthritis | 0% | 0% | 0% | 0% | 11% | 6% | <0.01 |

'High genetic risk' defined as having two copies of the HLA-SE. 'Undifferentiated arthritis' means: presence of clinically suspect arthralgia + at least one swollen joint (patient reported or nurse examined). P values computed excluding the NA group.

ACPA, anticitrullinated peptide antibodies; ANOVA, analysis of variance; BMI, body mass index; CCP, cyclic citrullinated peptide; HLA-SE, human leucocyte antigen shared epitope allele; SCREEN-RA, Evaluation of a SCREENing strategy for Rheumatoid Arthritis; ULN, upper limit of the norm.
The proportion of highly expanded T cell clones in the peripheral blood of participants in the SCREEN-RA cohort increased the closer the participants were to the onset of RA, which is consistent with the ‘mucosal origins hypothesis’. Indeed, antigen-specific T cells are required to build antibody mediated immune responses by activating B-cells. In particular, this cytokine cross-talk takes place at the mucosal level, where B-cells will in turn generate high amounts of secreted IgA, including IgA-ACPAs in inflammatory context. An expansion of T-cell clones before RA diagnosis therefore fits in line with the current hypothesis that RA could result from the systemic-spread of an initially local mucosal immune reaction.

**Female hormonal factors**

The assessment of female hormonal factors among women in the SCREEN-RA cohort suggested that perimenopausal status was significantly associated with ACPA positivity (p<0.001), which underlines the potential role of female hormonal factors in the onset of RA. This is in line with a previous finding that the prevalence of ACPAs increases with age, peaking between 45 and 55 years old for women, but not for men.

**Periodontitis**

SCREEN-RA participants (n=99) were examined for periodontal status by a blinded periodontist, to assess the link with ACPA seropositivity. This nested case–control study revealed a higher prevalence and severity of periodontitis and poorer periodontal conditions in the ACPA positive subjects, compared with ACPA negative subjects. This finding suggests that periodontitis precedes the development of the disease and may be causally associated with the onset of RA.

**Gut microbiota**

Bacterial composition of available stool samples was determined by a blinded external research group. ‘High-risk’ samples were then compared with samples from asymptomatic participants, and revealed an expansion of *Prevotella* species, in particular *Prevotella copri*. This study was the first to confirm intestinal expansion of known RA-associated microbes in the pre-clinical phases of RA, suggesting that the association between gut microbiome and early RA might be causal. A second stool sampling campaign is currently ongoing.

**Strengths and limitations**

The main strength of the SCREEN-RA study is its longitudinal design, with physical and biological data collected in a controlled environment, using standard operating procedures. The recruitment and long-term follow-up of asymptomatic individuals allows better characterisation of the preclinical phases of RA. The variety of preclinical RA stages enrolled gives the opportunity to realise nested studies, which help to understand the link between environmental factors and specific preclinical stages of the disease, and ultimately apprehend factors driving the onset of RA. The longitudinal

**Expanded T cell clones**

The proportion of highly expanded T cell clones in the peripheral blood of participants in the SCREEN-RA cohort increased the closer the participants were to the onset of RA, which is consistent with the ‘mucosal origins hypothesis’. Indeed, antigen-specific T cells are required to build antibody mediated immune responses by activating B-cells. In particular, this cytokine cross-talk takes place at the mucosal level, where B-cells will in turn generate high amounts of secreted IgA, including IgA-ACPAs in inflammatory context. An expansion of T-cell clones before RA diagnosis therefore fits in line with the current hypothesis that RA could result from the systemic-spread of an initially local mucosal immune reaction.

**Female hormonal factors**

The assessment of female hormonal factors among women in the SCREEN-RA cohort suggested that perimenopausal status was significantly associated with ACPA positivity (p<0.001), which underlines the potential role of female hormonal factors in the onset of RA. This is in line with a previous finding that the prevalence of ACPAs increases with age, peaking between 45 and 55 years old for women, but not for men.

**Periodontitis**

SCREEN-RA participants (n=99) were examined for periodontal status by a blinded periodontist, to assess the link with ACPA seropositivity. This nested case–control study revealed a higher prevalence and severity of periodontitis and poorer periodontal conditions in the ACPA positive subjects, compared with ACPA negative subjects. This finding suggests that periodontitis precedes the development of the disease and may be causally associated with the onset of RA.

**Gut microbiota**

Bacterial composition of available stool samples was determined by a blinded external research group. ‘High-risk’ samples were then compared with samples from asymptomatic participants, and revealed an expansion of *Prevotella* species, in particular *Prevotella copri*. This study was the first to confirm intestinal expansion of known RA-associated microbes in the pre-clinical phases of RA, suggesting that the association between gut microbiome and early RA might be causal. A second stool sampling campaign is currently ongoing.

**Strengths and limitations**

The main strength of the SCREEN-RA study is its longitudinal design, with physical and biological data collected in a controlled environment, using standard operating procedures. The recruitment and long-term follow-up of asymptomatic individuals allows better characterisation of the preclinical phases of RA. The variety of preclinical RA stages enrolled gives the opportunity to realise nested studies, which help to understand the link between environmental factors and specific preclinical stages of the disease, and ultimately apprehend factors driving the onset of RA. The longitudinal
follow-up allows for more accurate causal inferences than typical case-control studies.

The principal limitation of the SCREEN-RA study is the low incidence and slow rate of RA conversion, which requires the enrolment of numerous asymptomatic participants to ensure the observation of a limited number of individuals developing a definite diagnosis of RA. Moreover, we cannot formally exclude a selection bias since symptomatic individuals could have a higher motivation to participate. However, the observed incidence rate in our cohort (~2.1 case per 1000 person-years) is still compatible with previous studies in the same population.5 7 20

An important part of the collected data is based on self-assessment; hence, we cannot completely exclude the possibility of outcome misclassification and measurement uncertainty on symptom-related items. Notably, our group 4 ‘Isolated CSA’, as presented in table 5, is likely to overestimate the proportion of individuals with true inflammatory arthralgias, because the CSA definition relies principally on self-reported symptomatology and nurse-examination. This overestimation appears on figure 4: some participants classified in group 4 often later regressed to lower-risk and asymptomatic subgroups, probably because of fluctuating aspecific symptomatology. To address this issue of misclassification, we categorised the highest-risk participants (group 5), as those presenting both CSA symptoms and biomarkers. We may further underestimate the proportion of asymptomatic ‘high-risk’ individuals (group ‘3’, ie. columns 2 and 3 in table 5), because of differential follow-up procedures. Indeed, our risk classification includes biological markers that are not immediately available, hence enrolled individuals are occasionally misclassified as low risk, due to delay in obtaining serological results (ie, non-assigned subgroup in table 5). Second, because of budget limitations, or refusal from the participants, not all individuals have been blood-sampled yearly, since, low-risk participants are not invited to provide additional blood sample unless they develop new symptoms. To address this issue, we consider increasing our blood-sampling capability to include every low-risk individual in the annual serological sampling.

**FUTURE PLANS**

Currently ongoing, a multicentre collaboration is focusing on characterising antibody production at mucosal site to identify novel biomarkers for the prediction of RA development. Analysis will include immunohistochemistry, 16s RNA sequencing, single cell cloning. A complementary project aims at pinpointing other biomarkers by large ‘multiomics’ analysis. The collected blood samples will be used to extract genomic DNA (targeted n=500) and total RNA (targeted n=700), which will be compared with DNA and RNA from patients with RA from the SCQM-RA cohort (targeted n=100).25 Finally, linking periodontitis to ACPA status previously suggested that mucosal inflammation can be an important trigger in the onset of autoimmunity associated with RA.25 One of the largest mucosal site is the gut, and our initial analysis of intestinal microbiota of SCREEN-RA participants suggested a link between gut dysbiosis and development of RA.50 Thus, we are currently resampling faecal material of participants at different preclinical stages, using more up-to-date methodology,66 as well as studying mucosal and serological immune responses against hypothesised ‘autoimmunogenic’ micro-organisms (such as *P. copri*).76 77

**CONCLUSION**

Started in 2009 in Switzerland, the SCREEN-RA cohort focuses on long-term follow-up of individuals at risk of RA. Both symptoms, signs and biological data have been collected systematically in 1458 FDR of patients with RA. Prospective cohort designs allow more reliable causal inference than case-control experiments, while providing the opportunity to realise nested studies or validation studies.

Despite slow conversion rate toward classifiable RA, the study confirmed the involvement, in early phases of RA, of previously known risk factors, such as female hormonal factors, periodontitis and autoantibodies. Future plans include validation of new RA-associated biomarkers, and assessment of host-microbial immune homeostasis in pre-clinical phases of RA. In the new era of ‘personalised medicine’, early identification and stratification of at-risk individuals will indeed be key to establish reliable diagnostic approaches. We also expect our future research to demonstrate the efficacy of targeted preventive interventions.

**COLLABORATION**

Our team welcomes collaborative projects, in particular for biomarker identification and/or replication studies. Contact senior author Pr. FINCKH (ORCID: 0000-0002-1210-4347 - Email: axel.finckh@hcuge.ch).

**Author affiliations**

1Division of Rheumatology, Geneva University Hospitals (HUG), Geneva, Switzerland
2Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, University of Manchester, Manchester, England
3Division of Rheumatology, Kantonsspital St. Gallen, St. Gallen, Switzerland
4Department of Biomedicine, University Hospital Basel, Basel, Switzerland
5Department of Rheumatology, Clinical Immunology and Allergology, Inselspital Universitätsspital Bern, Bern, Switzerland

**Twitter** Eric Trunk @erictrunk

**Acknowledgements** The authors also thank Adriana URIBE CAPARROS, Sylvette BAS, Danielle GASCON.

**Contributors** BTPG: MD-PhD candidate in the Rheumatology division. Involved in writing the manuscript, data management of the SCREEN-RA database, creation of the figures, collection of samples (serum /stool), implementation/organisation of stool collection, implementation of the Food Frequency Questionnaire, updating ethical approval and consent forms. CL: Biologist in the Rheumatology division, involved in managing the SCREEN-RA central biobank; managing the SCREEN-RA biobank-related database, classification of participants in ‘risk groups’, previously published studies (in particular ref 68), major corrections in the manuscript. DM: data manager in the rheumatology division (research group). Involved in data management of the SCREEN-RA database, creation of the figures in the manuscript (in particular, figures 2 and 3), major corrections in the manuscript (in particular, figure 4). KL: Rheumatologist and senior clinical associate. Involved in statistical analysis (R code), major corrections in the manuscript (overall structure, tables
content), previously published studies, in particular ref 72. ET: Research coordinator. Involved in updating patient information, various logistics and administrative issues for the cohort, implementation of follow-up questionnaires, coordination with other Swiss centres, communication/recruitment (social media, website, campaigning, etc.), minor corrections in the manuscript. OS: Research nurse. Involved in data acquisition (follow-up visits, sampling), expedition of stool sampling kits, optimisation of questionnaires, minor corrections in the manuscript. DSC: Professor of statistics. Involved in implementation of Redcap database, troubleshooting database, statistical analysis, implementation of the study (in particular, power and sample size calculation), major corrections in the manuscript. AR-R: Professor in a collaborative centre (St. Gallen). Involved in substantial sample collection and data acquisition in St. Gallen Canton, corrections in the manuscript. DK: professor in a collaborative centre (Basel). Involved in extension of the study in Switzerland, substantial sample collection and data acquisition in Basel Canton, corrections in the manuscript. BM: professor in a collaborative centre (Bern). Involved in extension of the study in Switzerland, substantial sample collection and data acquisition in Bern Canton, corrections in the manuscript.AF: professor in the Rheumatology division (Geneva) and Principal Investigator of the SCREEN-RA study. Involved in starting the cohort in 2009 (Geneva), development of initial questionnaires, obtaining of ethical approval and funding, overall study design, classification of participants in ‘risk groups’, supervision of most of the previous publications, major corrections in the manuscript.

**Funding**

This work has been made possible by the financial support of the Swiss National Science Foundation, the Schmidheiny Foundation, and DeReuter Foundation.

**Map disclaimer**

The inclusion of any map (including the depiction of any boundaries therein), or of any geographic or locational reference, does not imply the expression remains solely that of the relevant source and is not endorsed by BMJ. Maps are provided without any warranty of any kind, either express or implied.

**Competing interests**

None declared.

**Patient consent for publication**

Not required.

**Ethics approval**

Initial approval in 2008 by the ‘Comité départemental d’éthique de médecine interne et de médecine communautaire’, protocol 08-102, project name : ‘Évaluation d’une stratégie de dépistage de la polymyosite rhumatoïde’. Every modification on the project was then approved by relevant cantonal ethic committees (respectively for each Swiss canton for which the project was extended). The SCREEN-RA cohort has been approved by the relevant ethic Committees (project PB_2016-00889), and participants sign an informed consent before enrolment, in accordance with the Declaration of Helsinki.

**Provenance and peer review**

Not commissioned; externally peer reviewed.

**Data availability statement**

Data are available on reasonable request. Anonymised data from the SCREEN-RA cohort can be shared on request (contact senior author Pr. Finckh at axel.finckh@hcuge.ch).

**Supplemental material**

This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ DISCLAIMS all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access**

This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**

Benolt Thomas P Gilbert http://orcid.org/0000-0001-6037-6470
Kim Lauber http://orcid.org/0000-0002-4315-9009
Delphine S Courvoisier http://orcid.org/0000-0002-1956-2607

**REFERENCES**

1. Safiri S, Kolahi AA, Hoy D, et al. Global, regional and national burden of rheumatoid arthritis 1990-2017: a systematic analysis of the global burden of disease study 2017. *Ann Rheum Dis* 2019;78:1463–71.
2. Zlatkovic-Svenda MI, Saraux A, Tuncer T, et al. FRi0558 Rheumatoid Arthritis Prevalence in Europe, A EULAR-Endorsed Survey: Table 1. *Ann Rheum Dis* 2016;75:643–643.
3. Cooper GS, Stroehla BC. The epidemiology of autoimmune diseases. *Autoimmun Rev* 2003;2:119–25.
4. Viate S, Plant D, Raychaudhuri S, Genetics and epigenetics of rheumatoid arthritis. *Nat Rev Rheumatol* 2013;9:141–53.
5. Silman AJ, Hennessen E, Olier B. Incidence of rheumatoid arthritis in a genetically predisposed population. *Rheumatology* 1992;31:365–8.
6. Silman AJ, MacGregor AJ, Thomson W, et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. *Br J Rheumatol* 1995;33:903–7.
7. Alpizar-Rodríguez D, Pluchino N, Canny G, et al. The role of female hormonal factors in the development of rheumatoid arthritis. *Rheumatology* 2017;56:1254–63.
8. Hart JE, Laden F, Puett RC, et al. Exposure to traffic pollution and increased risk of rheumatoid arthritis. *Environ Health Perspect* 2009;117:1065–9.
9. Sundström B, Johansson I, Rantapää-Dahlqvist S. Interaction between dietary sodium and smoking increases the risk for rheumatoid arthritis: results from a nested case-control study. *Rheumatology* 2015;54:487–93.
10. Hu Y, Costenbader KH, Gao X, et al. Sugar-sweetened soda consumption and risk of developing rheumatoid arthritis in women. *Am J Clin Nutr* 2014;100:959–67.
11. Gan RW, Young KA, Zerbe GO, et al. Lower omega-3 fatty acids are associated with the persistence of anti-cyclic citrullinated peptide autoantibodies in a population at risk for future rheumatoid arthritis: a nested case-control study. *Rheumatology* 2016;55:376–7.
12. Di Giuseppe D, Crippa A, Orsini N, et al. Fish consumption and risk of rheumatoid arthritis: a dose-response meta-analysis. *Arthritis Res Ther* 2014;16:446.
13. Tedeschi SK, Cui J, Arkema EV, et al. Elevated BMI and antibodies to citrullinated proteins interact to increase rheumatoid arthritis risk and shorten time to diagnosis: a nested case-control study of women in the nurses’ health studies. *Arthritis Rheum* 2017;69:692–8.
14. Bemis EA, Norris JM, Seifert J, et al. Complement and its environmental determinants in the progression of human rheumatoid arthritis. *Mol Immunol* 2019;112:256–65.
15. Gross J, Oubaya N, Eymard F, et al. Stressful life events as a trigger for rheumatoid arthritis development within a year: a case-control study. *Scand J Rheumatol* 2017;46:507–8.
16. Bang S-Y, Lee K-H, Cho S-K, et al. Smoking increases rheumatoid arthritis susceptibility in individuals carrying the HLA-DRB1 shared epitope, regardless of rheumatoid factor or anti-cyclic citrullinated peptide antibody status. *Arthritis Rheum* 2010;62:760–7.
17. Criswell LA, Saag KG, Mikulsk TR, et al. Smoking interacts with genetic risk factors in the development of rheumatoid arthritis among older Caucasian women. *Ann Rheum Dis* 2006;65:1163–7.
18. Jaakkola JJK, Gissler M. Maternal smoking in pregnancy as a determinant of rheumatoid arthritis and other inflammatory polyarthritis diseases during the first 7 years of life. *Int J Epidemiol* 2005;34:664–71.
19. Keenan BT, Chibnik LB, Cui J, et al. Effect of interactions of glutathione S-transferase T1, M1, and P1 and HMox1 gene promoter polymorphisms with heavy smoking on the risk of rheumatoid arthritis. *Arthritis Rheum* 2010;62:3196–210.
20. Klareskog L, Stolt P, Lundberg K, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. *Arthritis Rheum* 2006;54:58–69.
21. Holers VM, Demoruelle MK, Kuhn KA, et al. Rheumatoid arthritis and the mucosal origins hypothesis: protection turns to destruction. *Nat Rev Rheumatol* 2018;14:542–57.
22. Al-Katma MK, Bissada NF, Bordeaux JM, et al. Control of periodontal infection reduces the severity of active rheumatoid arthritis. *J Clin Rheumatol* 2007;13:134–7.
23. Chen H-H, Huang N, Chen Y-M, et al. Association between a history of periodontitis and the risk of rheumatoid arthritis: a nationwide, population-based, case-control study. *Ann Rheum Dis* 2013;72:1206–11.
24. Demmer RT, Molitor JA, Jacobs DR, et al. Periodontal disease, tooth loss and incident rheumatoid arthritis: results from the first National health and nutrition examination survey and its epidemiological follow-up study. *J Clin Periodontol* 2011;38:1005–1006.
25. Konig MF, Abusleele L, Reinhard J, et al. Aggregatibacter actinomycetemcomitans-induced hypercitrullination links periodontal
infection to autoimmunity in rheumatoid arthritis. Sci Transl Med 2016;8:369ra176.

Nguyen Y, Mariette X, Salliot C, et al. Chronic diarrhea and risk of rheumatoid arthritis: findings from the French E3N-EPIC cohort study. Rheumatology 2020;59:376–81.

Quirke A-M, Perry E, Carwright A, et al. Bronchiectasis is a model for chronic bacterial infection inducing autoimmunity in rheumatoid arthritis. Arthritis Rheumatol 2015;67:2335–42.

Janssen KM, de Smit MJ, Brouwer E, et al. Rheumatoid arthritis-associated autoantibodies in non-rheumatoid arthritis patients with mucosal inflammation: a case-control study. Arthritis Res Ther 2015;17:174.

Kishikawa T, Maeda Y, Nii T, et al. Metagenome-wide association study of gut microbiome revealed novel aetiology of rheumatoid arthritis in the Japanese population. Ann Rheum Dis 2020;79:103–11.

Alpizar-Rodriguez D, Lesker TR, Gronow A, et al. Prevotella copri in individuals at risk for rheumatoid arthritis. Ann Rheum Dis 2019;78:590–3.

Maeda Y, Kurakawa T, Umemoto E, et al. Dysbiosis contributes to arthritis development via activation of autoimmune T cells in the intestine. Arthritis Rheum 2016;68:2646–61.

Chen J, Wright K, Davis JM, et al. An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. Genome Med 2016;8:43.

Zhang X, Zhang D, Jia H, et al. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partially normalized after treatment. Nat Med 2015;21:895–905.

Scher JU, Szczesnaik A, Longman RS, et al. Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis. Elife 2013;2:e01717.

Gerlaç DM, Raza K, van Baarsen LGM, et al. EULAR recommendations for terminology and research in individuals at risk of rheumatoid arthritis: report from the study Group for risk factors for rheumatoid arthritis. Ann Rheum Dis 2012;71:638–41.

Källberg H, Jacobsen S, Bengtsson C, et al. Alcohol consumption is associated with decreased risk of rheumatoid arthritis: results from two Scandinavian case-control studies. Ann Rheum Dis 2009;68:222–7.

Lu B, Solomon DH, Costenbader KH, et al. Alcohol consumption and markers of inflammation in women with preclinical rheumatoid arthritis. Arthritis Rheum 2010;62:3554–9.

Costenbader KH, Karlsson EW, Epstein-Barr virus and rheumatoid arthritis: is there a link? Arthritis Res Ther 2006;8:204.

Nielen MMJ, van Schaardenburg D, Reesink HW, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004;50:380–6.

Rantapää-Dahlqvist S, de Jong BAW, Berglin E, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003;48:2741–9.

Kolivula M-K, Heliovaara M, Ramberg J, et al. Autoantibodies binding to citrullinated telopeptide of type II collagen and to cyclic citrullinated peptides predict synergistically the development of seropositive rheumatoid arthritis. Ann Rheum Dis 2007;66:1450–5.

van Gaalen FA, Visser H, Huizinga TWJ. A comparison of the diagnostic accuracy and prognostic value of the first and second anti-cyclic citrullinated peptides (CCP1 and CCP2) autoantibody tests for rheumatoid arthritis. Ann Rheum Dis 2005;64:1510–2.

Saraux A, Berthelot JM, Chales G, et al. Value of laboratory tests in early prediction of rheumatoid arthritis. Arthritis Rheum 2002;47:155–62.

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.

Deane KD, Norris JM, Helmers VM. Preclinical rheumatoid arthritis: identification, evaluation, and future directions for investigation. Rheum Dis Clin North Am 2010;36:2313–41.

Verheul MK, Vierboom MPM, ’t Hart BA, et al. Anti-carbamylated protein antibodies precede disease onset in monkeys with collagen-induced arthritis. Arthritis Res Ther 2010;12:R246.

Sieghard D, Platzter A, Studnicin P, et al. Determination of autoantibody isotypes increases the sensitivity of Serodiagnostics in rheumatoid arthritis. Front Immunol 2018;9:876.

van Beers-Tas MH, Ter Wee MM, van Tuyll LH, et al. Initial validation and results of the symptoms in persons at risk of rheumatoid arthritis (SPARRA) questionnaire: a EULAR project. RMD Open 2018;4:e000641.
healthy individuals at risk for rheumatoid arthritis. Clin Rheumatol 2017;36:677–82.

73 Loutan L, Alpizar-Rodriguez D, Courvoisier DS, et al. Periodontal status correlates with anti-citrullinated protein antibodies in first-degree relatives of individuals with rheumatoid arthritis. J Clin Periodontol 2019;46:690–8.

74 Kolfenbach JR, Deane KD, Derber LA, et al. A prospective approach to investigating the natural history of preclinical rheumatoid arthritis (RA) using first-degree relatives of probands with RA. Arthritis Rheum 2009;61:1735–42.

75 Die SCQM Foundation (Swiss clinical quality management in rheumatic diseases). Available: https://www.scqm.ch/en/ [Accessed 19 Feb 2020].

76 Pianta A, Arvikar S, Strle K, et al. Evidence of the immune relevance of Prevotella copri, a gut microbe, in patients with rheumatoid arthritis. Arthritis Rheumatol 2017;69:964–75.

77 Pianta A, Arvikar SL, Strle K, et al. Two rheumatoid arthritis-specific autoantigens correlate microbial immunity with autoimmune responses in joints. J Clin Invest 2017;127:2946–56.