OSTEOARTHRITIS

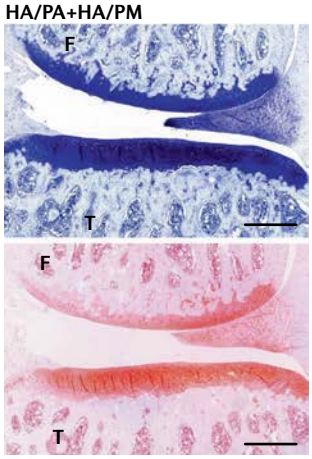
Restoring cartilage lubrication to heal post-traumatic OA

The extracellular matrix at the surface of healthy articular cartilage and its interaction with synovial fluid molecules ensure a low degree of friction within the joint. Lubricin and lipid components of the synovial fluid form a nanofibre with hyaluronic acid (HA) that binds to type II collagen and fibronectin in cartilage, effectively lubricating the joint. This lubrication is lost when cartilage is damaged, as occurs during osteoarthritis (OA).

"During the early stage of OA, cartilage friction is increased, resulting in the formation of chondral debris and subsequent inflammation," explains Li Ren of South China University of Technology, corresponding author of a new study on the restoration of cartilage lubrication as a treatment for OA. "The chondral debris and inflammation inside the joint lead to a feedback loop that accelerates the damage. HA can be injected into the joints of patients every month as a lubricant to treat OA. However, HA is quickly degraded by the enzymes present in the joint due to inflammation."

Ren and colleagues studied the natural interactions between cartilage and synovial fluid in order to create a lubricant that would be resistant to degradation by enzymes associated with inflammation in the joint. "Our strategy is biomimicry," states Ren. "Many researchers have focused their work on increasing the degradation-resisting properties of lubricants by methods such as cross-linking or increasing the molecular weight. However, in our opinion, the lubricant should interact with the cartilage surface in a more natural way."

Using their biomimicry strategy, the research team developed two synthetic nanofibres with an HA backbone: one that mimics lubricin (called HA/PM) and one that mimics lipids (called HA/PA). These nanofibers were designed to be biocompatible and to bind to fibronectin and type II collagen, properties that were confirmed both in vitro using human cartilage explants treated with trypsin to mimic osteoarthritic cartilage, and in vivo in the joints of healthy rats.



Toluidine blue and safranin O staining of the knee joints of rats with osteoarthritis 8 weeks after intra-articular injection of HA/PA and HA/PM. Scale bar, 500 μm. Adapted from Xie, R. et al. *Nat. Biomed. Eng.* https://doi.org/10.1038/s41551-021-00785-y (2021), Springer Nature Limited.

Using surgical transection of the anterior cruciate ligament to mimic post-traumatic OA in rats, Ren and colleagues examined the effects of their biomimetic lubricants on diseased cartilage. The combination of HA/PM and HA/PA performed better than either lubricant alone and was able to restore lubrication of the joint to near normal levels, thereby aiding cartilage healing. The combination treatment also performed better than either HA alone or an injectable HA formulation currently used in the clinic. Cartilage in affected joints 8-weeks post-treatment with HA/PM and HA/PA looked very similar to cartilage from healthy joints, with a uniform distribution of aggrecan and type II collagen and a smooth surface, suggesting that cartilage regeneration had taken place.

The researchers hope to perform follow-on studies in large animals with joint anatomy that is more similar to humans, with a view to moving their lubricant towards the clinic.

Joanna Clarke

ORIGINAL ARTICLE Xie, R. et al. Biomimetic cartilage-lubricating polymers regenerate cartilage in rats with early osteoarthritis. *Nat. Biomed. Eng.* https://doi.org/10.1038/s41551-021-00785-y (2021)

IN BRIEF

SPONDYLOARTHRITIS

Anti-TNF dose can be reduced in PsA and axSpA

Results from a retrospective cohort study suggest that use of a disease activity-guided dose optimization (DAGDO) strategy to taper TNF inhibitor therapy is safe and effective in patients with psoriatic arthritis (PsA) or axial spondyloarthritis (axSpA) and low disease activity. The mean percentage of daily defined dose of TNF inhibitor was 108% in the full-dose period and 78% after DAGDO in those with PsA ($n = 153$; median follow-up period 46 months), and 108% in the full-dose period and 72% after DAGDO in those with axSpA ($n = 171$; median follow-up period 44 months), with no difference in disease activity scores.

ORIGINAL ARTICLE Michielsens, C. A. J. et al. Tumour necrosis factor inhibitor dose adaptation in psoriatic arthritis and axial spondyloarthritis (TAPAS): a retrospective cohort study. *Rheumatology* https://doi.org/10.1093/rheumatology/keab741 (2021)

RHEUMATOID ARTHRITIS

Cancer recurrence risk not increased by DMARDs

A meta-analysis of 12 retrospective studies determined that the overall risk of any new or recurrent cancer in patients with rheumatoid arthritis (RA) and a history of cancer is not increased in those who initiate any biologic DMARD (bDMARD) compared with those who do not start bDMARDs (RR 1.09; 95% CI 0.92–1.32, $P = 0.31$); the total number of patients in the studies could not be calculated owing to missing data but was at least 7,560. On the basis of data from four studies, bDMARD treatment was associated with an increased risk of new or recurrent skin cancer (RR 1.32; 95% CI 1.02–1.72); however, skin cancer risk was not increased when melanomas were excluded.

ORIGINAL ARTICLE Wetzman, A. et al. Risk of cancer after initiation of targeted therapies in patients with rheumatoid arthritis and a prior cancer: systematic review with meta-analysis. *Arthritis Care Res.* https://doi.org/10.1002/acr.24784 (2021)

OSTEOARTHRITIS

Which NSAIDs are best for OA treatment?

In a network meta-analysis of 192 randomized trials (102,829 participants) that evaluated different preparations and doses of NSAIDs, opioids or paracetamol to treat osteoarthritis (OA), the oral NSAIDs etoricoxib 60 mg per day and diclofenac 150 mg per day seemed to be the most effective interventions for improving pain and function but were associated with an increased risk of adverse events. Topical diclofenac 70–81 mg per day also seemed effective and was generally safer than oral diclofenac. The clinical benefit of opioids for OA did not outweigh the risk of harm, regardless of dose.

ORIGINAL ARTICLE da Costa, B. R. et al. Effectiveness and safety of non-steroidal anti-inflammatory drugs and opioid treatment for knee and hip osteoarthritis: network meta-analysis. *BMJ* **375**, n2321 (2021)

RHEUMATOID ARTHRITIS

Treatment withdrawal is feasible in RA remission

Many patients with rheumatoid arthritis (RA) can remain in remission after tapering or withdrawal of DMARD therapy, according to findings from the phase III RETRO trial. Among 282 patients with sustained (≥12 months) remission, 43.3% of those who then discontinued DMARDs at 6 months maintained relapse-free remission at 12 months, compared with 58.6% of those whose dose of DMARDs was halved and 81.2% who continued to receive a full dose. Most patients whose disease relapsed regained remission after restarting treatment.

ORIGINAL ARTICLE Tascilar, K. et al. Treatment tapering and stopping in patients with rheumatoid arthritis in stable remission (RETRO): a multicentre, randomised, controlled, open-label, phase 3 trial. *Lancet Rheumatol.* https://doi.org/10.1016/S2665-9913(21)00220-4 (2021)

PSORIATIC ARTHRITIS

Does psoriasis treatment affect PsA development?

Joseph F. Merola and Alexis Ogdie

The results of several retrospective studies have reported that systemic treatment in patients with psoriasis reduces the risk of incident psoriatic arthritis (PsA). Although encouraging from a prevention perspective, such studies are limited in their ability to provide a conclusive understanding of PsA risk, preventing a clear picture from emerging.

Refers to Acosta Felquer M. L. et al. Treating the skin with biologics in patients with psoriasis decreases the incidence of psoriatic arthritis. *Ann. Rheum. Dis.* https://doi.org/10.1136/annrheumdis-2021-220865 (2021).

Individuals with psoriasis who are at increased risk of developing psoriatic arthritis (PsA) provide a unique opportunity to study disease progression and potentially intervene to prevent PsA from developing[1]. This disease prevention concept has been addressed in a study by Acosta Felquer et al.[2], as well as in two additional studies[3,4], all of which used retrospective cohort study designs to assess whether systemic treatment of psoriasis reduces the risk of incident PsA. Althought the studies present somewhat similar results, we urge caution when interpreting the conclusions as retrospective study designs have inherent limitations.

In their publication, Acosta Felquer et al. present a retrospective cohort study from a single health maintenance organization (HMO) in Argentina[2]. In this study, all patients with psoriasis not diagnosed with PsA at baseline who were seen in the HMO between 2000 and 2018 were monitored in follow-up for the development of PsA. Follow-up time started at initiation of a topical therapy, biologic DMARD (bDMARD) or conventional synthetic DMARD (csDMARD) and ended at time of PsA diagnosis or the end of follow-up in the electronic medical record. In total, 103 patients receiving bDMARD therapy were observed for an average of 4.4 years, 6 of whom developed PsA (0.43 per 100 person years), and 229 patients receiving csDMARDs were observed for an average of 1 year, 2 of whom developed PsA (1.20 per 100 person years). By contrast, 231 cases of PsA were diagnosed among individuals receiving topicals treatments ($n = 1{,}387$) over an average of 10 years of follow-up (1.67 per 100 person years). Thus, the incidence of PsA in those receiving topical therapy was much higher than in those receiving bDMARDs or csDMARDs. The difference between those receiving bDMARDs and the remainder of the cohort remained similarly 'protective' after propensity score matching. One interesting aspect of this study was that the researchers found a relatively high incidence of PsA[2]; this result could further support the dermatology–rheumatology collaborative care model[5] that this centre utilizes.

In another single, combined dermatology–rheumatology centre study, Gisondi et al. utilized a similar design to assess incident PsA risk[3]. Patients were included if they had completed at least three phototherapy treatment courses or had received bDMARD therapy for at least 5 years. Using traditional multivariable models, the hazard ratio (HR) for bDMARD therapy compared with phototherapy was 0.31 (95% CI 0.13–0.74) after adjusting for age, sex, psoriasis severity and nail psoriasis. However, after applying propensity score matching, the HR was 2.07 (95% CI 0.87–4.93)[3]. These results were not only attentuated from those presented by Acosta Felquer et al.[2], but were in fact 'flipped' to suggest a possible association between bDMARD use and increased risk of incident PsA after applying propensity score matching (albiet with wide confidence intervals).

In a third retrospective cohort study, Shalev Rosenthal et al.[4] compared patients with psoriasis who had used bDMARDs with those who had not used bDMARDs (but had been prescribed either two systemic therapies or one systemic therapy and phototherapy) within an HMO. The primary outcome was an International Classification of Diseases (ICD9) code for PsA, ankylosing spondylitis, enthesitis, spondylosis or undifferentiated arthritis in the 10 years following therapy initiation. The two groups were quite different at baseline, becoming more similar after propensity score matching but differing with regards to male:female ratio and year of diagnosis. Over a mean of 8 years, 109 bDMARD-exposed (16%) and 75 bDMARD-naive (11%) patients developed PsA and/or possible PsA (for example, ICD9 codes for other types of inflammatory arthritis) with an HR of 1.39 (95% CI 1.03–1.87).

> "retrospective designs … can only provide information about the beginning of the story"

Although the results of these studies[2–4] are largely congruent with our hypothesis that treatment of psoriasis improves the risk for PsA[1], some caveats are needed when interpreting their conclusions. First, as is often the case with retrospective cohort studies that compare two or more therapies, the groups of patients are clearly not comparable. The potential for confounding by indication is substantial; for example, reasons behind the choice of treatment might be related to the severity of psoriasis (which is known to be associated with development of PsA[6]) or to more complex and often undocumented factors (such as whether a patient qualifies for or even wants bDMARD therapy). Additionally, differential follow-up times (bDMARDs tend to be started later in the disease course) might allow for more time to develop and identify PsA in one group than in another. All three studies[2–4] employed propensity score matching, in which a score that reflects the probability of receiving treatment is created from measured covariates and patients with similar scores from different groups are matched, as a way to balance the treatment groups. However, a core principle of interpreting the results of propensity score matching causally is that there cannot be substantial unmeasured

confounding. In this scenario, there is clear unmeasured confounding as only limited ways exist to measure or identify why one patient is prescribed a bDMARD when an otherwise similar patient is not. Thus, the results of these studies[2–4] cannot be interpreted causally but only as observed assocations within a given patient population. Because of the types of imbalances in treatment groups in both measured and unmeasured characteristics, retrospective designs such as those used in these three studies can only provide information about the beginning of the story. The development of PsA risk scores could improve our understanding of the results of retrospective studies; for example, patients could be stratified on the basis of risk of PsA to examine differences by risk level.

Next, the lopsided incidence of PsA between groups in these studies[2–4] could be related more to study design than to biology. Patients at high risk of PsA or who are experiencing joint symptoms might have been excluded or have already been diagnosed with PsA at the time of therapy initiation (particularly given the access to rheumatologists in the dermatology–rheumatology collaboration centres). This factor could be particularly problematic in designs in which inclusion requires a specific duration of therapy. In addition, dermatologists, particularly those trained to detect PsA in a collaborative centre, might be selecting systemic agents over topical therapy or phototherapy for patients with musculoskeletal symptoms and other risk factors for PsA development. Furthermore, by only studying a population that is seen by dermatologists (as was done in these three studies) the true risk in the full population of individuals with psoriasis could be underestimated or overestimated. Dermatologists in the USA, for example, tend to undercapture musculoskeletal complaints compared with general practitioners[7]. The dermatology-based population of patients with psoriasis might also be different for other reasons, opening up the opportunity for collider stratification bias, a form of selection bias in which the results in the selected population can differ from the results in the full population, which has been observed in other aspects of psoriatic disease[8].

Lastly, relatively few outcomes were observed in all three studies[2–4]. This low number of outcomes leads to wide confidence intervals, suggesting that larger studies are needed to confirm their findings. We propose large, multicentre or even multinational studies, in which a broader population can be observed.

In summary, considerable gaps remain in our knowledge of how systemic treatment of psoriasis potentially affects progression to PsA. These three retrospective studies[2–4], although representing positive steps forward, leave us without the conclusive data that only a prospective randomized controlled trial might afford. In BOX 1, we summarize additional biases that should be considered when interpreting observational studies that seek to examine the effect of bDMARD therapy on the development of PsA. It is clear that thoughtfully designed prospective studies are needed to move the field forward; ongoing work from a prevention of PsA study group (PAMPA) is seeking to answer this question by defining a group of patients with psoriasis who are at-risk for progression to PsA and assessing disease modification with an interventional study design[1,9].

Box 1 | **Considerations when interpreting observational data**

Several forms of bias can influence the results of retrospective cohort studies examining the outcomes of therapy, which are important to consider when interpreting data to infer causality[10].

Confounding by indication
Patients are prescribed a medication for reasons ('an indication') beyond a simple set of rules and often these reasons are not measured, interfering with interpretation of the results.

Confounding by prognosis
Therapy is prescribed because a patient has a poor prognosis; in this case, the groups might have differential risks for the outcome (or prognosis) at baseline that cannot be adjusted for in the analysis.

Unmeasured confounding
Important covariates that would ideally be included in the analysis are not measured or captured in the dataset and could influence both the receipt of therapy and the outcome of interest.

Survival bias
A patient must 'survive' to receive the therapy of interest, leading to differences in the groups (particularly the duration of disease); this is a form of selection bias.

Protopathic bias
A therapy is prescribed because of a symptom or an undiagnosed disease that is also the outcome of interest.

Collider stratification bias
Studying a subset of the population might result in spurious results that are not replicated when examining the full the population.

Joseph F. Merola[1] ✉ *and Alexis Ogdie*[2]

[1]*Department of Dermatology and Department of Medicine, Division of Rheumatology and Immunology, Harvard Medical School, Brigham and Women's Hospital, Boston, MA, USA.*

[2]*Departments of Medicine and Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.*

✉*e-mail: jfmerola@bwh.harvard.edu*

https://doi.org/10.1038/s41584-021-00706-y

1. Scher, J. U. et al. Preventing psoriatic arthritis: focusing on patients with psoriasis at increased risk of transition. *Nat. Rev. Rheumatol.* **15**, 153–166 (2019).
2. Acosta Felquer, M. L. et al. Treating the skin with biologics in patients with psoriasis decreases the incidence of psoriatic arthritis. *Ann. Rheum. Dis.* https://doi.org/10.1136/annrheumdis-2021-220865 (2021).
3. Gisondi, P. et al. Biological disease-modifying antirheumatic drugs may mitigate the risk of psoriatic arthritis in patients with chronic plaque psoriasis. *Ann. Rheum. Dis.* https://doi.org/10.1136/annrheumdis-2021-219961 (2021).
4. Shalev Rosenthal, Y. et al. Psoriatic arthritis incidence among patients receiving biologic medications for psoriasis: a nested case control study. *Arthritis Rheumatol.* https://doi.org/10.1002/art.41946 (2021).
5. Haberman, R. et al. Bridging the gaps in the care of psoriasis and psoriatic arthritis: the role of combined clinics. *Curr. Rheumatol. Rep.* **20**, 76 (2018).
6. Ogdie, A. et al. Body surface area affected by psoriasis and the risk for psoriatic arthritis: a prospective population-based cohort study. *Rheumatology* https://doi.org/10.1093/rheumatology/keab622 (2021).
7. Ogdie, A. et al. Longitudinal analysis of the patient pathways to diagnosis of psoriatic arthritis. *Arthritis Res. Ther.* **23**, 252 (2021).
8. Nguyen, U.-S. D. T. et al. Smoking paradox in the development of psoriatic arthritis among patients with psoriasis: a population-based study. *Ann. Rheum. Dis.* **77**, 119–123 (2018).
9. Bell, S. et al. Aiming for cure and preventive initiatives in psoriatic disease: building synergy at NPF, GRAPPA, and PPACMAN. *Curr. Rheumatol. Rep.* **22**, 78 (2020).
10. Hernan, M.A. & Robins, J. M. *Causal Inference: What If* (Boca Raton: Chapman & Hall, 2020).

Acknowledgements
The authors wish to acknowledge C. Ritchlin, J. Scher, J. Gelfand and the PAMPA study group dedicated to investigating opportunities in the prevention of psoriatic arthritis.

Competing interests
J.F.M. declares he is a consultant and/or investigator for Abbvie, Amgen, Biogen, BMS, Dermavant, Eli Lilly, Janssen, Leo Pharma, Novartis, Pfizer, Regeneron, Sanofi, Sun Pharma and UCB. A.O. declares she has consulted for Abbvie, Amgen, BMS, Celgene, CorEvitas, Gilead, Janssen, Lilly, Novartis, Pfizer and UCB and has received grants from Abbvie, Novartis and Pfizer to the University of Pennsylvania and Amgen to the Forward Databank. Her spouse has received royalties from Novartis.

An introduction to machine learning and analysis of its use in rheumatic diseases

Kathryn M. Kingsmore ✉, Christopher E. Puglisi, Amrie C. Grammer and Peter E. Lipsky

Abstract | Machine learning (ML) is a computerized analytical technique that is being increasingly employed in biomedicine. ML often provides an advantage over explicitly programmed strategies in the analysis of multidimensional information by recognizing relationships in the data that were not previously appreciated. As such, the use of ML in rheumatology is increasing, and numerous studies have employed ML to classify patients with rheumatic autoimmune inflammatory diseases (RAIDs) from medical records and imaging, biometric or gene expression data. However, these studies are limited by sample size, the accuracy of sample labelling, and absence of datasets for external validation. In addition, there is potential for ML models to overfit or underfit the data and, thereby, these models might produce results that cannot be replicated in an unrelated dataset. In this Review, we introduce the basic principles of ML and discuss its current strengths and weaknesses in the classification of patients with RAIDs. Moreover, we highlight the successful analysis of the same type of input data (for example, medical records) with different algorithms, illustrating the potential plasticity of this analytical approach. Altogether, a better understanding of ML and the future application of advanced analytical techniques based on this approach, coupled with the increasing availability of biomedical data, may facilitate the development of meaningful precision medicine for patients with RAIDs.

Machine learning (ML) is an extension of computer science and statistics and involves the use of algorithms to recognize relationships in data and/or predict outcomes[1]. ML models are not explicitly programmed[2,3], meaning that each decision that the model makes is not determined by a user-inputted 'if-then' statement. Because ML models do not have direct instructions about how to approach a problem, ML models 'learn' or rather extract knowledge from input data[4]. Like statistical modelling, ML can be used to determine associations within data[5]. Whereas a rules-based statistical model is aimed at explaining specified or hypothesis-driven relationships in data[6], ML attempts to discover underlying connections in data and make decisions based on newly discovered associations. ML often uses unknown relationships that cannot be identified with other statistical techniques to predict outcomes. Moreover, ML uncovers unanticipated connections in the data, which is favourable for hypothesis generation.

The use of ML applications has increased because of the capacity of models to analyse high-dimensional data (that is, data in which there are more data points than samples). This 'big data' from electronic health records (EHRs), imaging, genetics and transcriptomics often makes purely statistical analyses unfeasible. For example, ML is well suited to address many of the challenges that arise in analysis of EHRs of patients with rheumatic autoimmune inflammatory diseases (RAIDs), such as the low accuracy of standard disease identification codes, low prevalence of RAIDs in the general population, and a need to determine which of many clinical features are important. Consequently, numerous studies have employed ML to identify patients with RAIDs and others have similarly employed ML techniques to extract significant clinical variables from the text of EHRs and classify patients[7–13]. In addition, ML can be used in big data analyses to identify associations when there is no apparent hypothesis to be tested. ML analysis of high-dimensional data is particularly useful for research in complex chronic diseases, such as RAIDs, in which the patient population is extremely heterogeneous, the conditions evolve over time and multiple factors contribute to disease burden[14–16]. Advanced analytical strategies, such as ML, offer a means of determining disease patterns among

Machine learning
(ML). A subset of artificial intelligence that utilizes software to predict outcomes and recognize relationships in data without explicit programmes for each step.

Algorithms
Mathematical or computational methods that can be applied to data to form a model.

Model
A framework built upon input data that can classify, regress or cluster.

AMPEL BioSolutions and RILITE Research Institute, Charlottesville, VA, USA.

✉*e-mail: kathryn.allison@ampelbiosolutions.com*

https://doi.org/10.1038/s41584-021-00708-w

Key points

- Appropriate application of machine learning (ML) algorithms and model construction, including that using data from patients with rheumatic autoimmune inflammatory diseases (RAIDs), involves preprocessing, feature selection, comparisons of multiple models to determine which is most appropriate for the data, and proper validation.
- ML has been applied to various types of data from patients with RAIDs, including medical records and imaging data to classify patients, sequencing data to predict genetic risk loci, biometric data to identify disease activity, transcriptomic data to classify or cluster patient subtypes, and demographic, genetic and genomic data to predict treatment response.
- Most published studies that describe the employment of ML in RAIDs, however, only serve as proof-of-principle studies as they lack adequate sample sizes or external test datasets; consequently, clinical translation of ML in rheumatology is in a nascent stage.
- Current ML studies provide hypotheses that can be validated in large retrospective datasets or used to design prospective trials characterized by correct data collection and sample sizes that are suitable for the application of ML.

Statistical modelling
A model that relies on explicitly programmed mathematical functions to explain relationships in data.

Classification
Prediction of a categorical outcome.

Regression
Prediction of a quantitative outcome.

Clustering
Grouping of data points with similar characteristics.

Labelled
Data for which the class or outcome value is known.

Class
A group with a label that is produced from classification.

Clusters
Groups without a label that are produced from clustering.

Supervised models
Models trained on labelled data that are used to predict classes or quantitative values.

Unsupervised models
Models trained on unlabelled data that are used to find associations and patterns that result in groups of similar samples.

Imputation
A method of replacing missing values with data points.

Data scaling
The processes of transforming data into a format that a computer algorithm can use, which can also involve normalization.

patients with RAIDs and aid in prediction of numerous outcome measurements.

Although the integration of data science and ML with the clinical practice of rheumatology is at a nascent stage, ML is frequently used in other areas of biomedicine to analyse medical images[17], subtype patient cohorts (especially based on omics data)[18], predict drug response[19] and aid in guiding personalized medicine[20]. ML-driven analyses of these and other outcome measures could greatly improve both research and clinical practice in RAIDs[6,21–23]. For example, in diseases such as primary Sjögren's syndrome (pSS) and systemic lupus erythematosus (SLE), for which few approved therapies are available, ML has the potential to aid in identification of candidates for drug repurposing[24]. Similarly, in rheumatoid arthritis (RA), psoriatic arthritis (PsA) and ankylosing spondylitis (AS), there is a need to predict the most effective therapy from many approved biologics and targeted small molecules[25,26]. In addition, ML might be useful to classify patients with RAIDs or predict disease outcomes based on available genetic[27] or ancestry[28] data. Altogether, ML has the potential to elucidate unknown connections between features of patients with RAIDs and ultimately help to facilitate effective patient care strategies. In this Review, we describe the common algorithms used for ML, explain the workflow of building an ML model, discuss the strengths and weakness of ML, and illustrate these principles with examples of ML application in the field of rheumatology.

Principles of machine learning

Outcomes

The output of ML is the construction and application of a predictive model. ML algorithms are routinely employed to build models for one of three outcomes: classification, regression or clustering. Classification is the prediction of discrete categories of data samples[29] into a labelled group, known as a 'class'; regression is the prediction of a numerical outcome for each sample[30]; and clustering is the grouping of similar observations into unlabelled groups formally known as 'clusters' (FIG. 1). In theory, in SLE, ML classification models could distinguish patients with SLE from controls, whereas regression models could be used to predict the value for autoantibody titre in each patient based on their diagnosed class or other features. Patients could further be assigned to classes such as 'active disease' or 'inactive disease' based on their predicted autoantibody titre value. Clustering of the same SLE and control samples would group samples on the basis of their similar characteristics into subsets.

Types of machine learning

ML algorithms for classification, regression and clustering belong to one of four types of learning, including supervised, unsupervised, semi-supervised and reinforcement learning[31], of which supervised and unsupervised algorithms are the most common. Supervised and unsupervised algorithms can be differentiated based on whether the outcome variable is labelled[32,33] — that is, the identity of the model outcome variable is known (for example, SLE or healthy, or positive or negative for autoantibodies). Data labels are often determined by an expert in the field or based on measurement of a specific parameter (such as autoantibody titre). Supervised models are constructed to predict known groups or values (that is, labelled classes or values, respectively), whereas unsupervised models are constructed to predict groups of unknown quality (that is, unlabelled clusters)[34], such as samples grouped together based on algorithm-detected characteristics in the data. For example, a supervised ML model could be applied to predict the disease activity status of patients with SLE with specified disease activity class (active or inactive), whereas an unsupervised model could group the same patients with SLE based on shared characteristics among patients, such as similar gene expression patterns, without knowing their disease activity status or clinical presentation.

Machine learning workflow

Although the intrinsic architecture of models built for classification, regression and clustering may differ, the model construction process is similar. The general workflow for all ML models involves data preprocessing and model construction (FIG. 2). Supervised classification and regression models additionally involve model validation and assessment[34].

Data preprocessing

Data preprocessing is a fundamental part of ML[35]. Because the nature of the input data will affect the ultimate model outcome prediction, careful curation of input data before application of an ML algorithm is essential to improving the usefulness and credibility of the model. Data curation and preprocessing include managing incomplete data (imputation)[36], transforming the data into a form that a computer can utilize (data scaling)[37,38], and selecting the most appropriate variables to use as input into the model (feature selection)[39].

Biomedical datasets are often subject to missing data points. For many ML models, it is best to approximate or replace missing data points, a process known as imputation[36], for better model performance[40]. Depending on the type of data, there are many different statistical techniques for imputation (reviewed

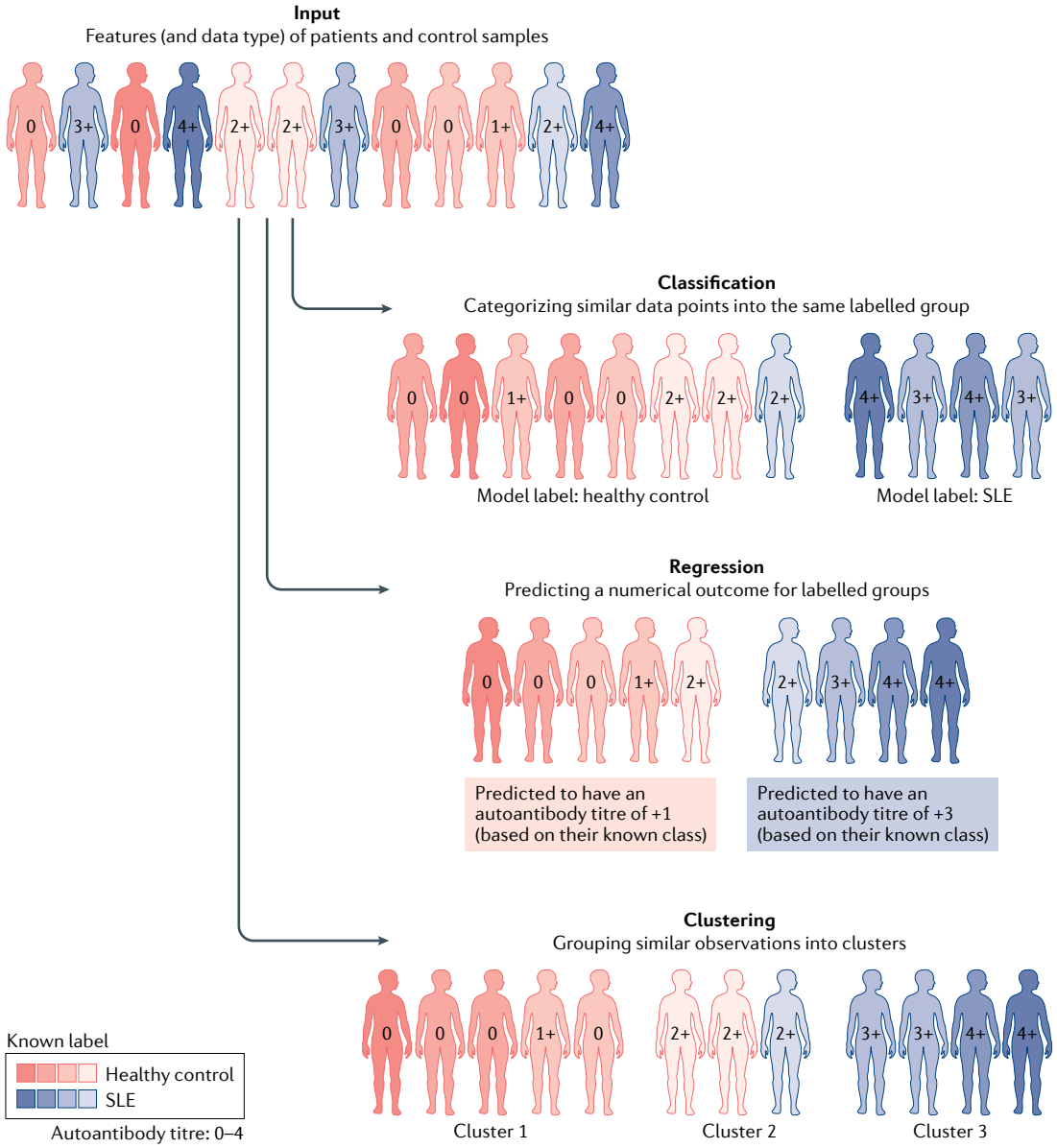


Fig. 1 | **The output of a machine learning model is classification, regression or clustering.** Autoantibody titre data were provided for seven healthy individuals and five patients with systemic lupus erythematosus (SLE). A classification model was built using the autoantibody titre to predict the label (class) of the sample, a regression model was built that used the label of the sample to predict the autoantibody titre of the patient, and a clustering model aggregated the samples into groups (clusters) according to similar values of autoantibody titre.

elsewhere[41]). In addition, ML algorithms themselves can be used for imputation[42–44].

After imputing missing values, biomedical data often need to be transformed into a numerical format that a computer can process. For example, EHR data contain notes that might be in unstructured text format, such as International Classification of Diseases (ICD) codes, prescriptions and patient demographics, or numeric format, including values derived from laboratory measurements. Natural language processing (NLP) is a specific branch of ML comprising computational methods by which computers interpret human language[45]. NLP algorithms vary in complexity. Some NLP algorithms create a histogram of word frequencies in a text (for example, bag of words (BOWs))[46], whereas others create word maps that represent conceptually similar ideas in close geometric proximity using vectors (for example, word embeddings)[47]. Thereby, similar vectors have similar associations that capture the word's meaning. NLP in biomedicine often makes use of concept unique identifiers (CUIs), which are specific identifiers given to related concepts in a particular discipline. For example, 'arthralgia' and 'joint pain' map to the CUI C0003862. The Unified Medical Language System is an inventory of CUIs specific to medicine that helps to delineate conceptual ideas for a word with multiple meanings[48].

In addition, it is often necessary to transform or normalize numerical data before model construction, a process known as data scaling[37,38]. Numerous techniques are available for data scaling, including

Feature selection
The process of selecting the best set of variables to be used as input for the model.

Natural language processing
(NLP). A data scaling process that is also a branch of ML, which allows computers to interpret human language.

Dimensionality reduction
The process of reducing the number of input variables (features).

Variance
Error as a result of the fluctuations in the observations, or how much the observations differ from the average value.

the min–max algorithm and the z-score algorithm[37]. Dimensionality reduction techniques can also be used to transform existing variables into a smaller set of variables[49]. One of the most commonly used dimensionality reduction techniques, principal component analysis (PCA), reduces high-dimensional data to a set of principal components based on the intrinsic features and variance of the data[50]. Additional dimensionality reduction techniques, such as multi-dimensional scaling (MDS), low variance and high correlation filters, forward selection and backward elimination, have been reviewed elsewhere[51–53].

ML models perform best when the number of input variables is optimized. Feature selection is a dimensionality reduction technique that is used to determine the most appropriate variables to use as inputs into an ML algorithm, as all measured variables might not provide information that is necessary for outcome prediction[54].

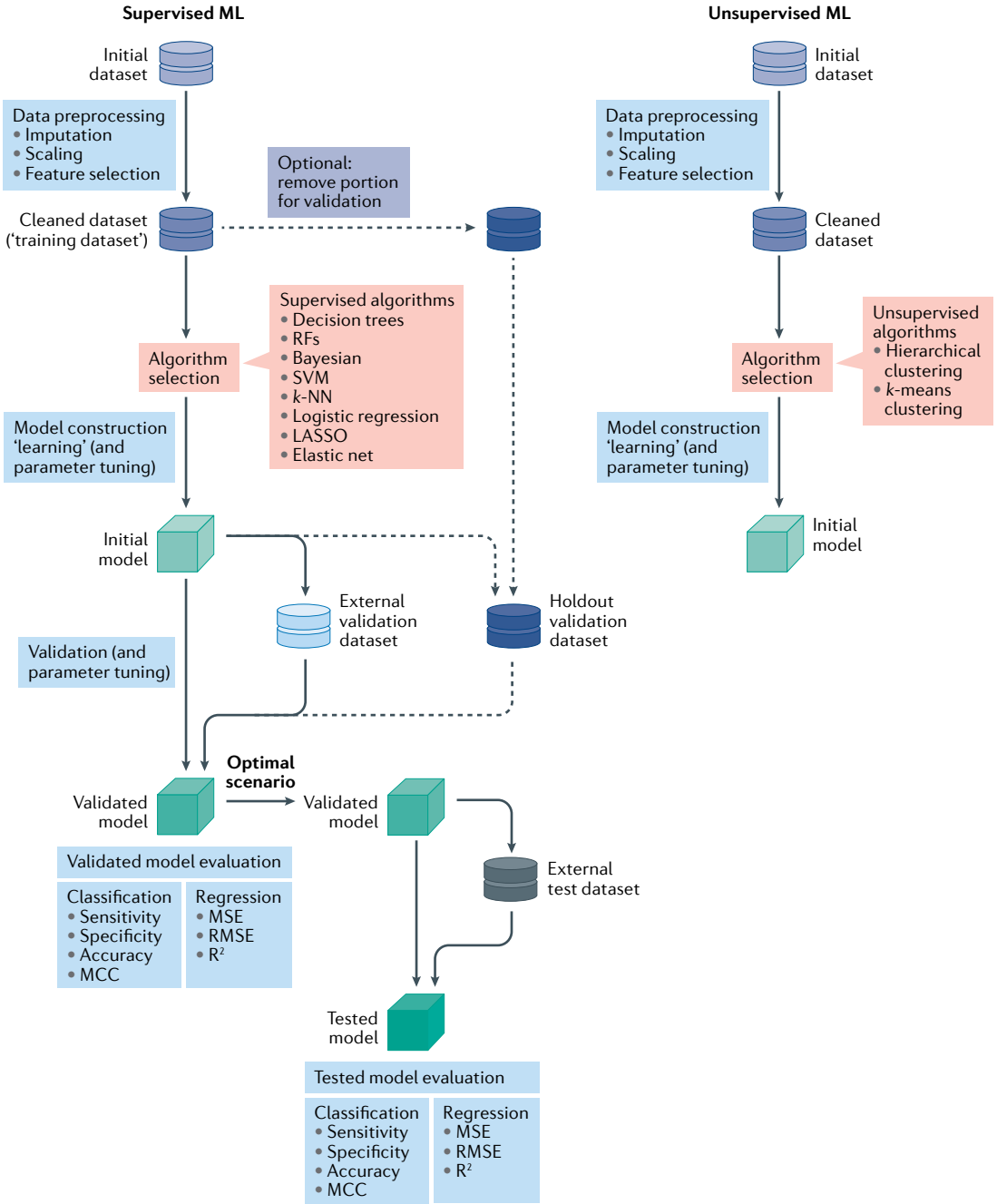


Fig. 2 | **Machine learning model workflow.** The application of machine learning (ML) models involves data processing and model construction 'learning'. For supervised models, there are additional validation and evaluation steps. Modified from "How to Build a Machine Learning Model"[98] with permission from Dr. Chanin Nantasenamat. *k*-NN, *k*-nearest neighbours; LASSO, least absolute shrinkage and selection operator; MCC, Matthews correlation coefficient; MSE, mean squared error; RF, random forest; RMSE, root mean squared error; R^2, coefficient of determination; SVM, support vector machine.

Biased
A biased model is one that fails to capture underlying patterns in data and thus there is a difference between the true values and the values predicted by the model.

Decision trees
Supervised method that asks a series of 'yes or no' questions with labelled data to classify or regress.

Clustering algorithms
Unsupervised methods that assign observations to subsets using mathematically calculated distances.

Neural networks
Supervised or unsupervised methods that build a series of networks to predict or classify. They are named because the structure of the model is aimed at mimicking the way in which a human brain operates.

Ensemble algorithms
Supervised methods that aggregate several predictors from multiple machine learning models (for example, random forest).

Bagging
Algorithm that generates training sets by sampling of the training data with replacement to generate individual models that are characteristic of the sample, which are then aggregated to build a final model.

Boosting
Algorithm that adds an additional simpler model to minimize the existing error during each iteration of a supervised model.

Bayesian algorithms
Supervised methods that solve classification problems by predicting the most probably hypothesis, given the input data (for example, naive Bayes).

Instance-based
Supervised methods that memorize instances seen in training to make predictions (for example, support vector machines and *k*-nearest neighbours).

Regression algorithms
Supervised methods that use linear or polynomial functions for or as a fundamental part of prediction (for example, linear regression and logistic regression).

Thus, it is crucial to identify the most relevant set of independent variables (features) before or during algorithm application. Employing a model with too many features is likely to increase computational expense and the likelihood of overfitting, whereas the application of a model with too few features could result in decreased model performance or underfitting[55,56]. In addition, some ML algorithms are highly sensitive to collinear features (that is, features that are correlated to one another and thus assumed to provide redundant information)[57]. Although an investigator could build a model based on the features they selected as clinically relevant to the hypothesis to be tested, it is more reliable to employ computational feature selection and subsequently compare the model's ability to select features with what is expected clinically, as a clinician could be biased by current knowledge of the disease, and ML might provide alternative relationships or results. There are many types of feature selection algorithms, including filter, wrapper and embedded methods[58]. Filter methods are employed independently of ML algorithms[59,60]. Filter methods compute the importance between features and the outcome variable using statistical tests. Conversely, wrapper methods are employed as an additional algorithm during model construction. Wrapper methods determine the usefulness of features based on multiple iterations of an ML model[59,60]. Embedded methods similarly carry out feature selection during ML model construction[39,60] but, unlike wrapper methods, they can be a part of the ML algorithm that is used for model construction. Embedded methods evaluate the importance of variables from the model-generated outcome[61].

Model construction

Choosing the most appropriate algorithm. ML algorithms used to build models vary in complexity. Common algorithms used for classification include decision trees, random forest (RF), gradient boosting machine (GBM), naive Bayes, support vector machine (SVM), *k*-nearest neighbours (*k*-NN) and logistic regression; those for regression include least absolute shrinkage and selection operator (LASSO) and ridge regression; and the two most common clustering algorithms are hierarchical clustering and *k*-means clustering[62]. Some algorithms, such as decision trees, SVM and neural networks, can be employed for both classification and regression. The principles, strengths and weaknesses of these algorithms are summarized in Supplementary Table 1.

In addition, ML algorithms can be categorized into related types based on their architecture. For example, RF and GBM are ensemble algorithms, which aggregate ML algorithms to classify or regress samples[63,64]. Thus, ensemble models are capable of making more accurate predictions than individual algorithms[63,64]. The two predominant strategies employed to create ensemble methods are bagging (also known as bootstrap aggregating)[65] and boosting. Algorithms that employ bagging sample different portions of the data with replacement during each iteration of the model[66], and, thereby, individual models from each iteration are built from separate groups of samples[67]. In ensemble models built with bagging, each sample is run through each individual model and the outcome recorded. The final model is an aggregate of the individual models, and the final classification or prediction for each sample is based on the weighted average of the individual model classifications or predictions. RF is one of the most frequently employed ensemble techniques that uses bagging[66]. Other ensemble methods employ boosting to build aggregated models. Boosting techniques add new, 'weak' models to the existing model at each iteration[68,69]. Each new model is aimed at further minimizing the error resulting from the previous aggregated model and is thereby trained on the existing error of the aggregated model[68]. The GBM algorithm is an ensemble technique that uses boosting and is based on the principles of gradient descent (that is, an optimization technique that modifies a parameter iteratively to determine the minimum of that particular function)[68].

Bayesian algorithms, including naive Bayes, use Bayes' theorem of probability to determine relationships among input data[70]. Bayesian algorithms are often used for classification, whereby the probability of each sample belonging to each possible class is calculated, and the event (class) with the highest probability is selected[71]. Instance-based learning, which includes SVM and *k*-NN algorithms, differs from other algorithms as it is not aimed at generalizing input data, but instead makes predictions based on specific examples (instances)[29] seen in training[72]. Regression algorithms employ mathematical functions with coefficients and variables to classify samples or regress numerical values[73]. For example, logistic regression uses the logistic function as the basis for building a binary ML classifier. Regularization algorithms are a specific type of regression algorithm and thereby are similarly built using mathematical functions characterized by coefficients in order to classify or predict variables. However, in contrast to logistic regression, regularization algorithms are embedded methods that employ feature selection during model construction. Regularization algorithms, such as LASSO and ridge regression, penalize the use of too many variables by shrinking model coefficients towards zero[74]. Both LASSO and ridge regression are aimed at minimizing coefficient magnitude to reduce model complexity and prevent overfitting[73,74]. In a regularization model constructed to predict disease activity status in patients with SLE, some variables such as antinuclear antibody titre or serum complement levels might be determined to contribute more or less to the accuracy. If the complement level reduced model accuracy, the coefficient for complement level in the LASSO model could be set to zero. Two additional dimensionality reduction algorithms that can also be used for classification include linear discriminant analysis and partial least squares discriminant analysis (PLS-DA)[75].

Finally, neural networks, which are algorithms that are designed to mimic the function of a simplified human brain[76], can be used for supervised or unsupervised learning. Neural networks take inputs and process them into an output through a series of layers, which mimic neurons firing to other neurons[77]. The first layer is the input layer, the last layer is the output layer, and all layers in between are known as hidden layers[33,78]. Each

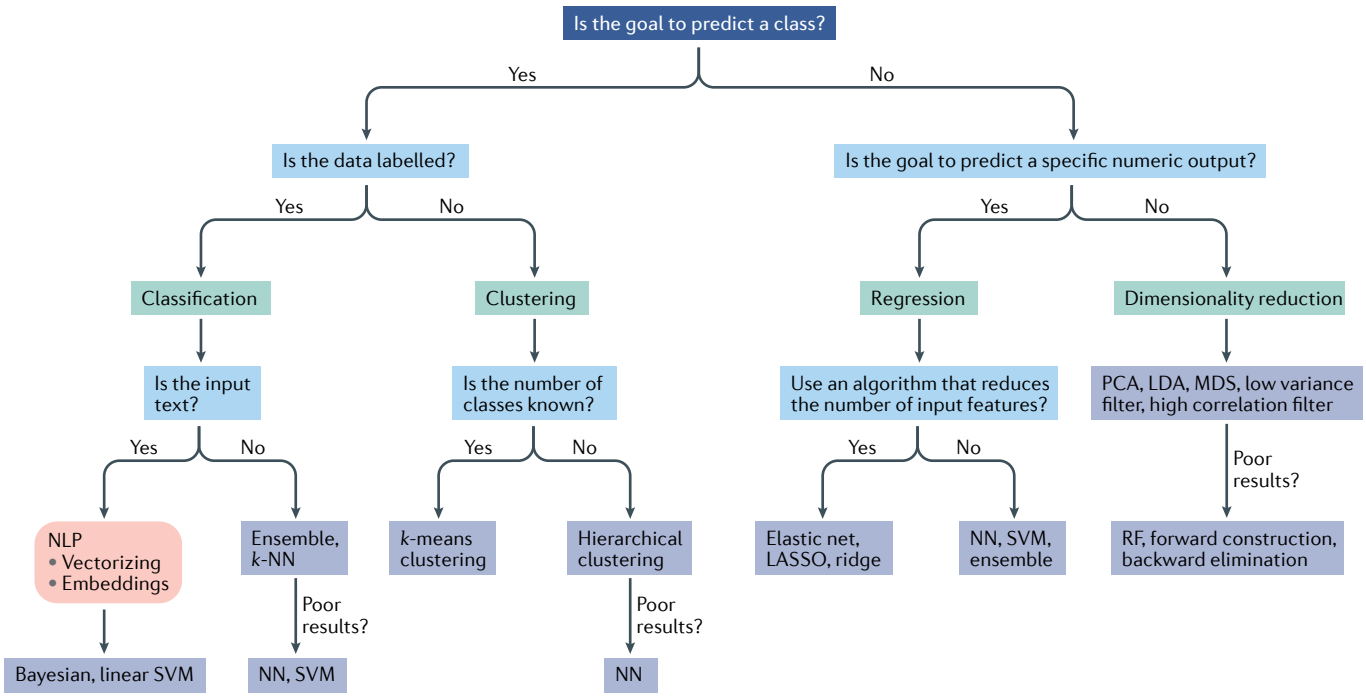


Fig. 3 | **Guidelines for selecting the most appropriate machine learning algorithm.** The goal of machine learning (ML) is classification, regression, clustering or dimensionality reduction. Based on the characteristics of the input data (for example, labelled versus unlabelled) and the desired outcome (that is, predict a group or predict a value), there are often more appropriate algorithms to apply, yet it is still recommended that multiple algorithms are compared in order to discover the model with the greatest predictive power. *k*-NN, *k*-nearest neighbours; LASSO, least absolute shrinkage and selection operator; LDA, linear discriminant analysis; MDS, multi-dimensional scaling; NLP, natural language processing; NN, nearest neighbours; PCA, principal component analysis; RF, random forest; SVM, support vector machine.

connection is assigned a weight, or importance[77]. Each neuron in the second layer can be thought of as a mathematical function whose variables are a weighted combination of each of the values of the neurons in the layer behind it. Furthermore, every neuron in the second layer is connected and determines the value of every neuron in the third layer, and so on until the final layer is reached[73]. With each iteration, the connections between each neuron are assigned a new weight to minimize error by gradient descent[79]. Neural networks are highly versatile and effective but computationally expensive. There are multiple variants of neural networks, including convolutional, recurrent and recursive neural networks[80–82]. Neural networks with roughly two or more hidden layers are categorized as deep-learning algorithms[82].

Algorithm selection is dependent on the questions being asked about the data and on qualities of the input data. A general framework for deciding on applicable algorithms based on the characteristics of data is depicted in FIG. 3. The algorithms proposed are those that to our knowledge are best at balancing interpretability and accuracy and are less prone to overfitting or underfitting the data. For example, Bayesian and SVM algorithms have been employed for the classification of text data[83]. In addition, decision trees are highly interpretable classifiers but can be prone to overfitting[84]; therefore, ensemble methods, including aggregated decision trees (such as RF), may be better as a starting point for classification.

Hyperparameters and parameters. After algorithm selection, certain variables need to be defined. Each ML model is characterized by hyperparameters and parameters[85]. Hyperparameters are set before model construction, either as the default values of the software algorithm or those that are input by the user[85,86]. For example, *k* in *k*-NN and the number of trees in RF are hyperparameters[85,86]. Hyperparameters can be tuned during model construction to improve model accuracy. Some approaches to hyperparameter tuning involve the evaluation of model performance with every combination of hyperparameters within user-defined boundaries[87,88]. For example, when constructing a *k*-NN model, the user must select two common hyperparameters: the number of clusters (*k*) and the distance measurement method (p). To tune the hyperparameters, model performance in each combination of *k* and p is evaluated and the highest performing model is selected. Other strategies involve setting a sampling distribution for each hyperparameter[87]. Strategies for hyperparameter tuning have been reviewed elsewhere[88].

Conversely, parameters are determined during model construction based on the input data[85]. Parameters to be set are intrinsic to the type of algorithm employed, although their value will differ for models built on different datasets. During model construction, supervised model parameters are tuned based on the ability of the model to accurately predict the known outcome[85]. Unsupervised models are constructed on data for which the outcome is not known, and therefore with

Regularization algorithms
A type of supervised regression method that shrinks coefficient estimates to zero to avoid overfitting (for example, least absolute shrinkage and selection operator and ridge regression).

Hyperparameters
Variables that must be set prior to model construction by the user or by software default and can then be tuned during model construction to maximize accuracy.

Parameters
Variables that are 'learned' during model construction. Parameters differ between algorithms based on algorithm architecture.

unsupervised models, parameters are tuned to minimize the similarity between objects in different clusters and maximize the similarity between objects in the same clusters[32].

Training and validating models. After the data are preprocessed and the most appropriate algorithm is selected, model training begins. Model training is an iterative process by which the model 'learns' to classify, regress or cluster the outcome variable. Through iterations, the model tunes its parameters to minimize error. Supervised models have the benefit of tuning model parameters based on the ability to achieve the correct labelled outcome.

Ideally, supervised models are built and improved using two independent datasets: the training dataset and the validation dataset[6,89]. The training dataset is the initial input data on which the model is built. The validation dataset is either a part of the training dataset that was reserved for validation, or an entirely separate dataset containing data with the same independent and dependent variables. The validation dataset is used to assess the performance of the initial model by giving an estimate of how accurately the model predicts the outcome. The validation dataset can also be used to further tune model hyperparameters[6]. All ML models are built on a training dataset, but some lack a dataset on which they can be validated. In this case, the model serves only as a proof-of-concept. Even with a validation dataset, however, final model performance must be evaluated without tuning. If more than two datasets with labelled data are available, then the final model can be evaluated with a third dataset known as the test dataset[6,89]. Unlike the validation dataset, the test dataset is used to provide an unbiased evaluation of the final model performance without hyperparameter tuning. At present, third datasets are rare in RAIDs research.

When a validation dataset is unavailable, a portion of the training data can be withdrawn from the original dataset and used as a validation dataset. Splitting the training dataset into a training dataset and a validation dataset is known as holdout[90,91]. A common holdout approach uses 80% of the data for training and 20% of the data for validation. Another validation technique, *k*-fold cross-validation, optimizes splitting one dataset into a training and validation portion through resampling. With *k*-fold cross-validation, the data are split into a number of groups, *k*, and the model is run *k* times[92]. During each run, one of the groups is withheld as the validation dataset, and the remaining groups are combined and used as the training data. The model is evaluated during each run based on its ability to accurately predict the validation data. Additional validation strategies for limited sample sizes have been reviewed elsewhere[93].

Assessment of model performance. Several metrics are used to characterize supervised model performance. For classification, metrics include accuracy (the ratio of correct predictions to total predictions), sensitivity (the true positive rate) and specificity (the true negative rate)[29,94]. If the classification problem is binary (that is, there are two classes) these values are often represented using receiver operating characteristic (ROC) curves, where the false-positive rate (1 – specificity) is plotted on the *x*-axis and the true positive rate is plotted on the *y*-axis[95]. Each point on the ROC curve represents the false-positive rate and true positive rate at a specific classification threshold[95]. The ROC curve is a frequently employed evaluation tool for supervised modelling and describes a model's ability to distinguish between classes. From the ROC curve, the area under the curve (AUC; also known as the area under the ROC curve (AUROC)) is calculated and represents the probability that the model can distinguish correct and incorrect classes (FIG. 4). A perfectly performing classifier has an AUC of 1.0, whereas a score of 0.5 indicates that the model's performance is comparable to random chance (guessing). For regression, mean squared error[96–98], root mean squared error (RMSE)[98,99] and the coefficient of determination (R^2)[97,98,100] can be used to evaluate model performance. For clustering, it is difficult to validate the performance of these models consistently[101,102], as unsupervised models are not contingent on labelled data.

Examples of machine learning use in rheumatic autoimmune inflammatory diseases

The increased use of ML in the analysis of RAIDs data[6,21–23], cancer[19,20], genetics[103], transcriptomics[103], biological networks[104], drug discovery[105] and autoimmune diseases[106] demonstrates the potential utility and power of this analytical strategy. Below, we highlight the use of ML in RAIDs research to date, focusing on studies that demonstrate correct application of ML by preprocessing data, comparing ML models built using different algorithms, choosing the model with the best accuracy, cross-validating the model and/or validating the model in external datasets. For each type of data below, TABLE 1 provides an example of a study employing one or more of these 'good practice' strategies.

Patient classification using electronic health records and clinical data

EHRs contain a wealth of information about individual patients and collective information about diseases[107], and efforts are being made to analyse EHR data to determine characteristics of patient cohorts that could then enable prediction of future disease progression, drug utilization, comorbidities or additional medical needs. Although multiple studies have analysed RAIDs-related EHR data without the use of ML[108–112], many examples exist of ML-based EHR analysis in RAIDs[8–11,13,113–115]. EHRs are inherently noisy, and a large part of EHR-based ML begins with language data transformation so that the classifier algorithm can interpret the data. Another challenge with EHR data is incorrect class labels (that is, RA or SLE). For example, the use of ICD codes alone to determine class labels can oversimplify disease identity and may result in false positives[116,117].

A study compared multiple ML models to determine the most appropriate method for classifying patients with SLE from EHR data[8]. Medical notes from 662 patients (322 SLE, 340 healthy individuals) were obtained, filtered based on presence of the word 'rheumatol', and

Training dataset
The dataset used by supervised models to 'learn' to predict an outcome by viewing both the input and output variables in the data.

Validation dataset
A portion of the training dataset that is withdrawn to give an estimate of fit while tuning model parameters, or a separate dataset used to estimate model fit and tune parameters.

Holdout
The process of reserving some samples for training and some for validation from a single dataset.

***k*-fold cross-validation**
An extension of model validation that partitions the data into complementary subsets when training, to perform parallel analyses on each subset.

Sensitivity
The proportion of the actual positives that are correctly identified. Also known as the true positive rate.

Specificity
The proportion of the actual negatives that are correctly identified. Also known as the true negative rate.

Receiver operating characteristic (ROC) curves
(ROC curve). A plot of the sensitivity against the 1 – specificity that is used to assess the performance of a binary classifier.

Area under the curve
(AUC). Generally refers to the area under the ROC curve, so it can also be referred to as the area under the ROC (AUROC).

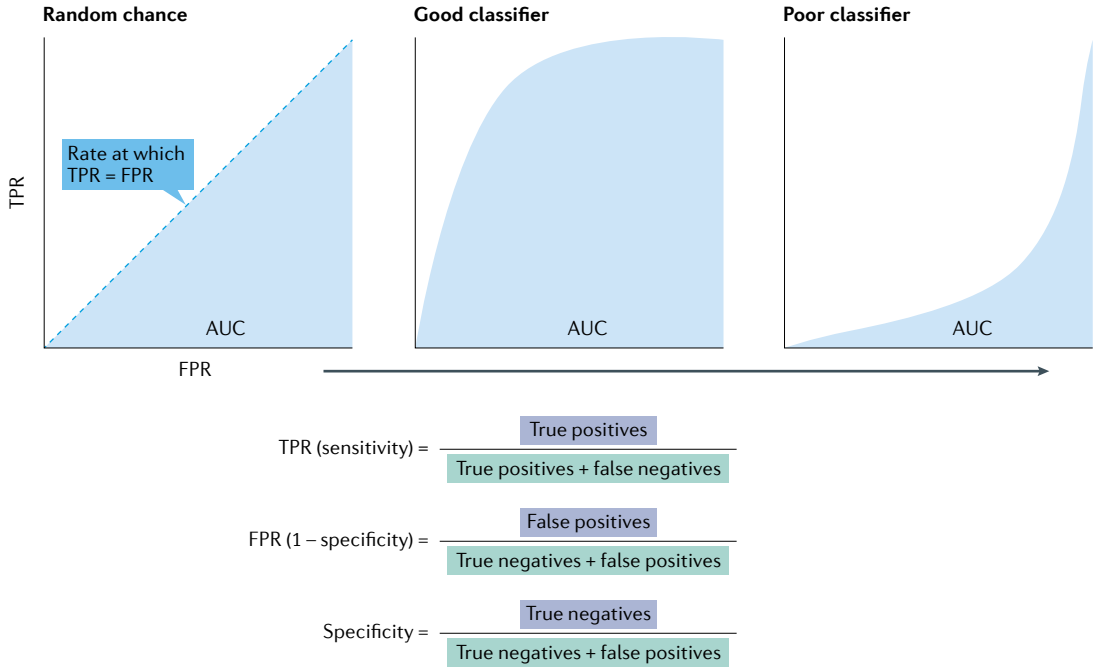


Fig. 4 | **Receiver operating characteristic curves are used to assess binary classification performance.** Receiver operating characteristic (ROC) curves plot the true positive rate (TPR; that is, the sensitivity) against the false positive rate (FPR, which is equivalent to 1 – specificity)[95,214]. The TPR is the ratio of positives that are correctly classified to all positive samples within the dataset. The FPR is the ratio of false positives that are identified to all negative identifications. ROC curves can illustrate that the classifier is no better than random chance because the TPR = FPR or that classification is good when the AUC >0.5 or poor when the AUC <0.5.

SLE samples were randomly sampled with replacement (bootstrapped) to create a balanced sample distribution. One hundred samples (50 SLE, 50 healthy individuals) were then removed from the balanced distribution to serve as the validation dataset and the remaining samples were used for training. Three independent NLP techniques (CUIs, BOWs and Word2Vec) were applied to the training data to extract relevant features. Once preprocessed, ML algorithms were applied in differing combinations, resulting in eight multistep models. Initial models from the training data were cross-validated and assessed for accuracy, and then assessed for accuracy in the validation dataset. Many of the ML models substantially outperformed traditional ICD code classification. Ultimately, the RF classifier with either BOWs or CUIs matrices as inputs performed best, with accuracy of either combination ≥95% and AUC >0.979. This study serves as an example of two important concepts: first, the use of ML in a multistep or combination approach by employing an initial ML algorithm for data transformation (preprocessing) and a subsequent ML algorithm for classification, and second, splitting the initial data into a training set and a validation set, when an external validation set is not available.

Another study classified patients with RA based on EHR-derived disease phenotypes and selected clinical codes (features) that had higher frequencies in patients with RA than in healthy individuals[10]. An RF feature selection method was employed to reduce the number of input predictors by ranking the importance of clinical codes (features) and selecting the top features[10]. This pipeline reduced the initial 43,100 features to 37 groups of clinical codes. Then, the most important model-selected features were aggregated with other codes, if determined to be sufficiently similar. A decision tree was then trained on these features to classify patients with RA. An external dataset was used for model validation; the decision tree model had 94.6% specificity, 86.2% sensitivity and 92.3% accuracy in its classification of patients with RA. This study appropriately utilizes feature selection and employs an external dataset for validation.

A series of studies demonstrated the potential power of ML classification in EHR data when the model is refined over time and can be evaluated in three independent datasets[118,119]. The initial EHR-derived penalized logistic regression model was built to classify RA using data derived from one independent dataset (500 samples for training and 29,432 samples for validation)[118], then later refined in the same dataset through cross-validation[119], and ultimately evaluated in two sufficiently large independent datasets, achieving AUCs >0.92 in both datasets[119]. The original penalized logistic regression model that compared ML classification was built with ICD codes or other codified data in EHRs alone, EHR data extracted by NLP alone or a combination of the two[118].

Others have attempted to develop a prognostic tool based on EHR data. For example, a predictive pipeline was developed that would improve diagnosis and early identification of patients with AS based on demographic, biomarker and clinical information for over 6,000 patients with AS and nearly 50,000 controls from the Truven database[114]. The authors first used mutual

Table 1 | **Examples of machine learning application to the classification of patients with rheumatic autoimmune inflammatory diseases**

Publication	Disease	Aim (analysis type)	Sample size	Type of Analysis	Input data type	Preprocessing	Algorithms compared	Final model selected	Validation	Assessment
Turner et al. (2017)[8]	SLE	Classify patients with SLE	Total: 322 patients with SLE; 340 controls Training: (bootstrapped to get equal numbers and then removed 100 for validation) 290 patients with SLE; 290 controls Validation: 50 patients with SLE; 50 controls	Classification	Medical notes	Bootstrap sampling and NLP (CUIs, BOWs and Word2Vec)	RF, naive Bayes, SVM and neural net	CUIs and RF or BOWs and RF	5-fold cross-validation, 100 samples removed before training for validation set	CUIs + RF: Accuracy = 95% AUC = 0.979 BOWs + RF: Accuracy = 95.25% AUC = 0.994
Gossec et al. (2019)[145]	RA and AS	Classify presence or absence of flare in patients with RA or AS	82 patients with RA; 73 patients with AS	Classification	Activity data (steps/minute)	Bootstrap sampling and normalization	Naive Bayes and RF	Naive Bayes	Split dataset: 70% for training and 30% for validation	Sensitivity = 96% Specificity = 97%
Robinson et al. (2020)[149]	SLE	Classify patients with juvenile-onset SLE	67 patients with SLE; 39 controls	Step 1: classification	Demographic data and PBMC flow cytometry data	Preprocessing not described	Balanced RF and sparse PLS-DA	Balanced RF	10-fold cross-validation; also confirmed significance of balanced RF-selected variables with logistic regression	Sensitivity = 89.6% Specificity = 82.1% AUC = 0.909
		Stratify patients with juvenile-onset SLE	67 patients with SLE	Step 2: clustering	Significant parameters (8/28 immune cell subsets from flow cytometry data) identified from balanced RF, sparse PLS-DA and logistic regression models	The step 1 model was used as feature selection	No comparison	*k*-means clustering	No formal validation because of unsupervised learning	Mathematical: none because unsupervised clustering Biological: group 1 from *k*-means clustering had more active disease over time, suggesting that clustering with the 8-cell type signature could further separate patients
Ceccarelli et al. (2018)[59]	SLE	Classify patients with and without erosive arthritis	120 patients with SLE	Classification	Demographic, clinical, laboratory and treatment data	Forward wrapper method to select features	Logistic regression and decision trees	Forward wrapper and logistic regression	Leave-one-out validation for hyperparameter tuning	AUC = 0.86

Table 1 (cont.) | **Examples of machine learning application to the classification of patients with rheumatic autoimmune inflammatory diseases**

Publication	Disease	Aim (analysis type)	Sample size	Type of Analysis	Input data type	Preprocessing	Algorithms compared	Final model selected	Validation	Assessment
Patrick et al. (2019)[186]	Psoriasis (and other inflammatory skin diseases such as SLE, scleroderma, myositis and myasthenia gravis)	Classify drugs for immune-mediated skin conditions	2,814 drugs mined from the literature	Classification	Drug-vocabulary matrices derived from PubMed abstracts (20 million abstracts, with 3.3 billion words)	NLP word embedding	GBM, logistic regression, RF, LASSO regression, nearest shrunken centroid and PLS-DA	PLS-DA	10-fold cross-validation; applied cross-validated model to a test dataset, and were able to confirm that the targets of model-predicted drugs were differentially expressed in psoriasis samples	AUC = 0.928
Figgett et al. (2019)[195]	SLE	Identify subsets of patients with SLE	161 patients with SLE	Step 1: clustering	Gene expression data (RNA-seq)	Normalization of gene expression data and batch correction; Gap and Davies–Bouldin clustering evaluations to determine the number of clusters	*k*-means clustering, PLS-DA and PCA	*k*-means clustering	No formal validation because of unsupervised learning, but PLS-DA achieved similar clusters to *k*-means clustering	Mathematical: none because unsupervised clustering Biological: gene set enrichment analysis to determine the transcriptomic profile of each cluster; different clinical features found in each cluster, suggesting the clusters could represent different clinical manifestations of disease
	SLE	Classify subsets of patients with SLE	Patients with SLE	Step 2: classification	Gene expression data (RNA-seq) with cluster labelling from *k*-means	Normalization of gene expression data and batch correction	ECOC SVM, RF	Both ECOC SVM and RF were equally accurate	ECOC SVM: validated in independent cases from additional datasets RF: repeated double cross-validation	ECOC SVM: accuracy = 88% RF: accuracy = 88%

AS, ankylosing spondylitis; AUC, area under the curve; BOWs, bag of words; CUIs, concept unique identifiers; ECOC, error-correcting output code; GBM, gradient boosting machine; LASSO, least absolute shrinkage and selection operator; ML, machine learning; NLP, natural language processing; PBMC, peripheral blood mononuclear cell; PCA, principal component analysis; PLS-DA, partial least squares discriminant analysis; RA, rheumatoid arthritis; RF, random forest; RNA-seq, RNA sequencing; SLE, systemic lupus erythematosus; SVM, support vector machine. [a]The table includes select examples of ML models built with each type of input data that demonstrate proper ML application by preprocessing data, comparing ML models built using different algorithms and choosing the model with the best accuracy, cross-validating the model, and/or validating the model in external datasets.

information, a statistic that measures the mutual dependence of two features[120], to select features for model input. After feature selection, they specified two analysis time periods (segments), so that they could predict whether patients from the first time period would be diagnosed with AS in the second time period. Patient data from segment 1 were split into the AS population and the matched control population, and used to train the first model (model A), which was evaluated with the remaining segment 1 data. Then, the second model (model B) was trained on the same AS cases as model A, but the false positives incorrectly labelled as AS cases by model A became the controls. As such, the second model was used to assess prediction accuracy with more challenging cases. Multiple ML classifiers were compared in order to determine the most accurate classifier for both model A and model B. When cross-validated, the gradient-boosting classifier was the top-performer for

Testing dataset
An independent dataset that is used to provide an unbiased evaluation of the final model fit.

model A, with an AUC of 0.81 when predicting AS. The extra trees model, another ensemble method[121], yielded the highest AUC of 0.79 for model B. Both trained models were developed to be applied sequentially on other independent datasets. The first model has the ability to exclude patients that are unlikely to have AS and the second has the ability to remove false positives. This approach of applying the results of one model to another model can help to mitigate the shortcomings of either model. The predictive accuracy of model A and model B was then evaluated by determining whether patients predicted to have AS in the first segment were correctly diagnosed with AS in the second time period. Together, model A and model B generated a positive predictive value more than twice that of a clinical model based on international AS diagnoses criteria, albeit the resulting positive predictive value was only 6.24%. This study demonstrates that good model accuracy in original training and validation data does not guarantee high performance in the testing dataset, despite comparison of multiple classifiers and the use of a multistep model, illustrating the necessity of external validation before proceeding to clinical application. However, it should be noted that the positive predictive value of 6.24% was improved compared to the diagnostic accuracy achieved using a model built on clinical parameters described by the Assessment of SpondyloArthritis international Society[114].

ML has also been used in RAIDs to classify patients with differing disease manifestations, using clinical data. In one example, linear discriminant analysis was used for selection of features that best separated patients with pSS based on fatigue levels (high or low fatigue)[122]. The nine selected clinical features were used to build an SVM classifier that achieved an AUC of 0.725, which was higher than the SVM classifier that included all 57 features[122], demonstrating the power of feature selection. Another study used medical records data (including demographic and laboratory data) as input for both a decision tree model and an RF model that were built to distinguish pSS from dry eye disease[123]. Using 89 features as input, both the decision tree and the RF models achieved high sensitivity, but the decision tree model had lower specificity. Although the RF model was able to classify patients with pSS accurately with all variables, it did not perform as well when only a subset of the variables was used[123], which demonstrates the need to compare models with differing numbers of features in order to achieve the best combination of variables, as prediction power might not always by improved by reducing the number of variables. Furthermore, a form of LASSO was utilized to classify RA disease activity in patients based on claims data only, claims and medication data, or a combination of claims, medication and laboratory data[107]. The number of variables differed between models. The model based on claims, medications and laboratory data achieved the highest sensitivity (83.1%), but the claims data model achieved the highest specificity (74.7%), again showing the benefit of comparing different models. Conversely, another RA study using patient characteristics, clinical outcomes, patient-reported outcomes, laboratory values and medication was input into a deep neural net to predict disease activity[124]. Similarly, LASSO–logistic regression was employed to classify patients with SLE from those with other rheumatological conditions, based on clinical and serological data[125]. The model, which was optimized with 10-fold cross-validation on the training data of 802 adults, performed well in a validation dataset comprising 502 patients with SLE and 143 control individuals. Finally, a logistic regression classifier was built to determine whether SLE pathogenesis is neutrophil- or lymphocyte-driven, using clinical variables[126]. Although the model was built with clinical data, the class labels (neutrophil-driven SLE or lymphocyte-driven SLE) were assigned to samples based on their gene expression profile[127]. The logistic regression model was trained on one cohort and cross-validated with 1,000 iterations before being tested on a separate cohort. Evaluation of the model with the test dataset demonstrated an AUC of 0.87. Overall, this study provides an example of correct employment of ML in which logistic regression seems to successfully classify patients with SLE as having neutrophil-driven or lymphocyte-driven SLE. Although models derived from different algorithms were not compared for classification, the investigators did examine different variable combinations as input for the logistic regression model, and optimized the number of iterations for cross-validation.

Patient classification using imaging and biometric data

ML can be employed to improve the accuracy of imaging-based diagnosis, evaluation and outcome prediction in RAIDs. Neural networks and deep learning have been employed in multiple studies aimed at diagnosing or grading RA disease activity based on imaging data[128–136]. Deep learning has also been used for image analysis in pSS to either classify the grade of pSS from salivary gland ultrasonography[137] or diagnose pSS from CT images[138]. In one study, feature selection was employed before model construction with an SVM algorithm and cross-validation to classify patients with neuropsychiatric SLE from controls using functional MRI data[139]. SVM has also been used to estimate the progression score of RA joints on X-ray images[140] and predict radiographic progression in AS[141]. In addition, ensemble algorithms have been investigated to classify patients with RA based on thermal measurement of hand joints[142] and to classify neuropsychiatric SLE based on connectivity disturbances measured by functional MRI[143].

The accuracy of many of these models has been assessed based on their agreement with the analysis of an expert rheumatologist. As diagnosis can be variable among rheumatologists, constructing a model that can consistently grade arthritis severity could help to reduce variability within the field[128]. To further reduce variability, future ML-guided image analysis models could be refined by comparing classification by the model with that of multiple pooled experts. Altogether, the present studies demonstrate the utility of training and validation, where the algorithm tunes its parameters to achieve the same diagnosis as that of the examining physician. However, if an ML model consistently disagrees with a physician's diagnosis, it would be pertinent

to examine whether model-observed patterns in the data could be used to stratify patient outcomes more effectively, as a clinician could be biased by current knowledge, or whether the model's predictions are a result of noise. Indeed, a cross-validated deep learning model was recently found to detect AS from CT images with a greater sensitivity and specificity than a radiologist with 9 years of experience[144].

Bayesian algorithms have been employed to assess the association between flares and physical activity in RA and AS[145]. Physical activity, recorded as steps per minute, was measured over 3 months by wearable activity trackers and used as input data to classify the presence or absence of a flare. The dataset was appropriately split into training and validation datasets and used to build a naive Bayes model that classified patient-reported flares with a sensitivity of 96% and specificity of 97%[145]. A similar study used ML to develop a single-sensor wearable diagnostic that could differentiate rest (sitting versus lying down) and activity (walking versus standing) data in patients with RA[146]. Both studies performed data preprocessing and selected features for model construction. In addition, both studies compared different models: the first analysed the data with naive Bayes and RF[145], and the second compared RF, SVM and deep learning for their ability to classify these activity patterns in patients with RA and healthy controls[146]. If either approach was validated in a large, independent population, physical activity tracking might provide a means of flare prediction or disease activity classification in RA and AS.

Patient classification using urinalysis, flow cytometry and genomics

Two studies used urinalysis data to build models for patient classification. In the first study, a diagnostic was developed to determine class, activity and chronicity of kidney disease in patients with lupus nephritis (LN) based on measurement of urine protein levels[147]. Neural networks were trained on normalized protein abundance and the resulting models were able to classify patients in each LN World Health Organization (WHO) histological class with a sensitivity of >86%[147]. Trained neural networks were able to predict the renal chronicity index with a Pearson's correlation coefficient of 0.87 (REF.[147]). In addition, this study compared the sensitivity of individual parameters (including individual protein spot, age, gender and race) in determining LN class, activity and chronicity. A combination of six protein spots (α-1 acid glycoprotein, zinc α-2 glycoprotein, IgG κ light chain and two spots of α_1 microglobulin) demonstrated high sensitivity for each output metric, but when employed individually were unable to classify patients[147]. The neural network classifier with the median performance following cross-validation was selected. Although this study compared ML outcomes with different numbers of input features and employed cross-validation, it was limited by a small sample size of 20 patients and lacked an external dataset for validation. The second study similarly sought to classify patients with LN into their WHO histological class based on clinical variables, such as levels of urinary *N*-acetyl-β-d-glucosaminidase enzyme, creatinine, C3 and serum urate[148]. In this study, an RF model could classify patients into their WHO histological class (II, III/IV and V) with accuracies of 51.3–63.7%, which was better than random classification (33.3%). Neither study compared different algorithms for classification, but the second study was less likely to overfit the data because of a substantially larger sample size.

Another study used both demographic data and flow cytometry data of 28 immune cell subsets from peripheral blood mononuclear cells to first classify and then cluster 67 patients with juvenile-onset SLE[149]. Two ML algorithms (balanced RF and sparse PLS-DA) were trained separately and cross-validated to classify the patients and determine the most important features. The balanced RF model achieved classification accuracy of 87.8%. Univariate logistic regression was used as a statistical technique to confirm the contribution of the immune cell subsets to juvenile-onset SLE, but not as an ML classifier. Important parameters identified from both supervised classification algorithms, as well as those identified by univariate logistic regression, were compared and the eight parameters in common among the three techniques were used as input to the *k*-means clustering algorithm. Four clusters were derived, and there were significant differences in T cell frequencies in the resulting patient clusters, whereas there were no statistically significant differences in demographic and treatment characteristics. This study serves to illustrate that parameters derived from supervised learning can be further used to find new associations or clusters in the data using unsupervised learning. Moreover, this study compares and combines different algorithms for parameter selection. In addition, this work reinforces the idea that it might be more advantageous to classify patients on the basis of biological or molecular disease characteristics than on clinical or demographic parameters[150]. However, the findings must still be validated in an independent dataset[150].

ML analysis of genomic data, especially because of its high dimensionality, is increasingly being employed. One study applied various ML algorithms to select a microRNA panel that classifies patients as having RA, SLE or neither disease[151]. Small RNA-sequencing (RNA-seq) data from 167 RA and 91 control samples were preprocessed and then used as input into an RF algorithm. The RF algorithm was used as a feature selection technique to identify microRNAs that were differentially expressed in RA samples, and then the ability of the differentially expressed microRNAs to classify patients with RA in the same cohort was evaluated using LASSO and logistic regression. Based on the results, six microRNAs were selected that classified RA in the RA validation cohort (32 patients with RA and 32 control individuals) with an AUC of 0.71 and SLE in the SLE validation cohort (12 patients with SLE) with an AUC of 0.80. However, the model could not reliably distinguish patients with SLE from those with RA. Altogether, this study serves as a good example of preprocessing, employing ML models in multiple steps (for both feature selection and classification), and then subsequent validation.

A further study used a generalized linear model (GLM), a type of linear regression, to classify whether

patients with SLE had active or inactive disease based on myeloid cell gene expression as compared with other cellular gene signatures[152]. Although the study was limited by small sample numbers and validation in an external dataset is needed, the myeloid cell signature outperformed the low-density granulocyte signature and was similar to the plasma cell signature in classifying patients by disease activity, as determined by the SLE Disease Activity Index. This study demonstrates the utility of investigator-based feature selection when there is a specific hypothesis to be tested. A third study compared GLM, RF, *k*-NN, and hierarchical clustering models built using scaled enrichment scores of gene modules derived from a gene co-expression approach with those built using raw gene expression data, for their ability to accurately classify patients as having active SLE or inactive SLE across independent data sets[153]. RF classification using input gene module enrichment scores achieved the highest accuracy among the supervised models following 10-fold cross-validation, although all supervised classifiers performed well. By contrast, attempts to discriminate active and inactive disease in patients using hierarchical clustering with the same input gene modules was unsuccessful. This study demonstrates the utility of comparing both differing preprocessing methods (raw gene expression versus gene module enrichment scores) and multiple supervised and/or unsupervised models to determine the most appropriate model for the dataset and desired outcome. However, external testing is still necessary.

ML has also been employed to predict ancestry of patients with SLE using gene expression data[28]. Gene expression values for 752 genes were used as input data for logistic regression, SVM and elastic GLM algorithms. These genes were chosen because they are members of gene modules with a demonstrated ability to separate patients by ancestry — an example of feature selection based on prior knowledge. After 10-fold cross-validation, SVM achieved an accuracy of 96%, although all models performed well. The 25 most important gene predictors in the SVM model demonstrated that the classification was based primarily on B cell gene expression differences between patients with SLE of African and European ancestry. Although not validated in an independent testing dataset, the 25 most important predictors align with previous reports of B cell perturbations in patients with SLE who are of African ancestry, which, coupled with the sufficiently large sample size, give more credence to this study.

An additional study used gene expression to refine histological disease subtyping of patients with RA[154]. First, expression levels for the 500 most variably expressed genes derived from 45 synovial biopsy samples (6 control individuals and 39 patients with RA) were used to cluster the patients. Consensus clustering, an ensemble clustering technique[155], revealed three clusters of patients, with the optimal number of clusters evaluated by statistical measures, and the appropriateness of the clusters confirmed by PCA. Patients from the three clusters were then labelled as either the low-, mixed- or high-inflammatory subtype based on histological evaluation of the biopsy sample from the same patient. The identity of each cluster, which was assigned as a different RA subtype, was confirmed with analysis of the functional enrichment of the differentially expressed genes that characterized each cluster[21,154]. The three subtypes derived from gene expression clustering were then used as the patient labels for classification[6,21,154]. Synovial histological features of the same 45 biopsy samples were used as input data to train several binary SVM classifiers to predict one labelled inflammatory subtype from the other two. The cross-validated model based on histological score input performed best (AUC = 0.88) when classifying the high-inflammatory subtype from the other two[154]. The model was less successful at classifying the mixed-inflammatory subtype (AUC = 0.59) from the others[154]. After cross-validation, the developed SVM models were used to predict the gene expression subtype of the remaining 82 synovial samples. However, there is still a need to confirm whether the subtype predictions were correct, as gene expression data of these samples were not available. The development of this model demonstrates that ML can be used to determine the histological features that are able to best predict the molecular phenotype of synovial samples and how histological features can then be used to predict gene expression profiles in patients with RA[21].

Classification of early or late flare in patients with SLE by gene expression alone has been compared with classification with other clinical features[156]. Four independent RF models were built using either blood gene expression data (enrichment of inflammatory gene modules in individual patients with SLE and control individuals), flow cytometry data (percentage of inflammatory cell populations in the blood), clinical data (SLE Disease Activity Index, clinical–serological markers of SLE) and soluble mediators (enzyme-linked immunosorbent assay and other analyte measurements of cytokines). Noisy variables were pruned and each model was built on 1–3 variables. In this example, the model based on flow cytometry measurements was the most accurate and that based on the gene expression module was the least accurate. Understanding which type of data is most useful for flare prediction could greatly improve future clinical practice; however, this proof-of-concept study was limited by a small sample size ($n = 21$ early flare, $n = 13$ late flare), cross-validation was not reported, and the models lacked independent test datasets.

Risk classification and outcome prediction

Genome-wide association studies have identified genetic susceptibility loci for SLE[27,157,158], RA[27,159], AS[27,160], PsA[27,161] and pSS[162]. However, the role of single nucleotide polymorphisms (SNPs) in disease progression is poorly understood. ML offers another possible modality for identifying disease risk markers using SNPs and genetic variants. In a Swedish study, ML was employed to classify patients with SLE using SNP data and subsequently predict genetic variants that confer risk of SLE[163]. An RF classifier was built using genotype data (134,523 SNPs) from 1,160 patients with SLE and 2,711 control individuals. The resulting cross-validated model classified patients with SLE with an AUC of 0.78. When the model was used for classification of LN alone, the classification

achieved an AUC of 0.91. Then, an RF model was developed to calculate gene importance scores and identify risk genes, based on the classification performance of those genes. The model identified 40 risk genes, 25 of which were already known to be linked to SLE[163]. The risk genes predicted by the model were compared with the top 40 genes determined by single-SNP association analysis and 15 genes were shared by the two methods. The overlap of these 15 genes between the ML approach and a more standard SNP analysis technique suggests that ML might aid in the identification of novel risk genes or the validation of previously identified loci. In another study, RF and logistic regression were employed to predict genetic interactions between identified risk genes in RA[164]. However, it has been suggested that although RF models could be used to prescreen risk loci, statistical analysis with logistic regression is necessary to determine genotype combinations that confer a high risk[165]. There are conflicting reports as to whether logistic regression or penalized logistic regression are advantageous compared with standard procedures to identify risk loci[166,167]. In addition, it is difficult to evaluate whether ML or the more standard approaches are most accurate for prediction of risk genes, but the RF method might be able to overcome linkage disequilibrium (the non-random association of alleles) by determining the importance of any given gene using many different decision trees[163]. Although ML has been employed in this context, it remains unclear whether RF or logistic regression is best suited for identification of risk genes. Instead, one might consider using ML to classify patients based on risk genes that have been identified by standard analytical methods.

ML can also be used to predict future outcomes, including resulting organ damage. For example, a supervised artificial neural network model was used to identify risk of chronic organ damage in SLE[78]. This model incorporated demographic data, laboratory parameters, and clinical parameters and was trained on two classes: 'controls', which include patients with SLE who had a baseline damage index of zero and did not develop damage, and 'cases', which include patients with SLE who had a baseline damage index of zero and subsequently developed damage. The model was trained until the AUC on the labelled data exceeded 0.95. The model's ability to generalize was also evaluated using eightfold validation, which yielded a final sensitivity of 74%. The same group developed ML models to classify patients with SLE with or without erosive arthritis[59]. Features were first selected using the forward wrapper method, and logistic regression or decision trees were employed for classification. Both ML models were chosen because of their interpretability and suitability for smaller sample sizes, unlike neural networks[59]. The forward wrapper and logistic regression model performed better than the forward wrapper and decision tree model and achieved an AUC of 0.806. However, both models achieved higher AUCs when implemented with the forward wrapper method in comparison with using only the classifier, demonstrating an appropriate use of feature selection.

Many patients with RAIDs are at risk of cardiovascular disease. Artificial neural networks have been employed to predict risk of cardiovascular events in pSS, using clinical and serological information as input data[168]. Other studies have similarly used ML to identify patients with PsA[169] or AS[170] who are at a high risk of cardiovascular disease and to predict risk of atherosclerosis[171] or hospital readmission[172] in patients with SLE. Of note, the RF ML model used in the prediction of cardiovascular risk for patients with AS was found to be more discriminative than all but one of the seven traditional predictors proposed by the EULAR[170]. In addition, ML was used to predict pregnancy outcomes in patients with SLE with preeclampsia ($n = 21$) and those with no known complications ($n = 45$)[173]. Eighteen ML models were built using one of three different inputs (transcriptomic data only, laboratory data or clinical data, or combined laboratory or clinical and transcriptomic data) and one of six different ML algorithms. Data preprocessing prior to application of ML algorithms involved the removal of unannotated transcripts and highly correlated genes. The six models built on transcriptomics data only had a mean accuracy of 74.2% compared with an average accuracy of 67.8% for those built on laboratory or clinical parameters. Use of both transcriptomic and clinical parameters for model construction increased prediction accuracy to 75.7%. The kernel PLS model achieved the highest AUC when only transcriptomic data were used, whereas the PLS regression for GLMs method performed slightly better for the combined input. Both models show potential for risk prediction in SLE, as they were appropriately compared with many different models and highly collinear features were removed; however, a larger sample size would provide more reliable results, which additionally need to be validated with an external cohort.

Finally, ML has been applied to predict renal flare in patients with LN who achieved response after initial therapy[174]. This study comprising 1,694 patients used demographic, clinical, serological, histological and therapeutic variables as input into an eXtreme Gradient Boosting (XGBoost) algorithm[174], which is an ensemble of decision trees and GBM[175]. The 1,694 patients with LN were split so that 70% were used for training and the remaining 30% were used for validation. In addition, fivefold cross-validation was employed. The validated model achieved an AUC of 0.819 and was used to identify important parameters for other clinical models. Although it is necessary to test this model in an independent dataset, the sample sizes of the training and holdout validation sets are sufficiently large, and the variables determined most important by the model have been previously implicated in LN[174], suggesting that the model is biologically accurate.

Predicting treatment response or candidates for treatment

In the era of expensive biologics, the ability to accurately predict non-responders to targeted therapies would be beneficial for patients and clinicians, so that focus could shift to treatments that are more likely to be effective. A 2014 challenge invited participants to design models that can predict response to TNF-neutralizing therapy in patients with RA[176,177]. Integrated demographic data,

treatment history, SNP data and baseline disease activity for 1,892 patients with RA were employed to train the models. One group separated patients based on the TNF-neutralizing treatment (adalimumab, etanercept or infliximab) they received and constructed multiple supervised ML models for each treatment to predict quantitative change in RA disease activity score[25]. The disease activity score value was then used to classify patients as treatment responders or non-responders. Models built from Gaussian process regression (a type of Bayesian regression), gradient-boosting, logistic regression or RF algorithms were compared. Following cross-validation, the Gaussian process regression model was most robust and could predict treatment non-response amongst patients with RA of European ancestry with an accuracy of 78%[25]. In an independent testing dataset of 680 patients, the model achieved an AUC of 0.615. The strengths and limitations of this study have been reviewed previously[176].

In another study, ML models were built to predict anti-TNF biologic therapy response in patients with RA using either gene expression or methylation data[178]. The RF model built using differentially expressed genes from RA peripheral blood mononuclear cells as features was most accurate in predicting response to adalimumab, whereas the RF model built using differentially methylated CpG positions was most accurate at predicting response to etanercept. However, external validation with a larger sample size is needed[179], as the authors performed validation with only nine external samples, and the study, like most performed in RAIDs, is limited by a small sample size, with only 40 patients in each treatment group used in model construction[178]. Others have also proposed the application of ML to aid in treatment of RA with biologics. For example, in a conference proceeding, SVM and deep learning were used to predict whether patients with RA received increased doses of infliximab[180]. Similarly, a neural net model was used to predict patients with AS who would use TNF inhibitors early in disease duration[181]. In addition, ML is being used to identify treatment response gene signatures in clinical trials of ustekinumab in SLE[182] and has been used to predict additional benefit from treatment with secukinumab in patients with PsA[183].

As a first step in developing a decision support tool to inform LN therapy, ML was employed to identify biomarkers of LN treatment response[184]. Urine cytokines were measured in a subset of patients with biopsy-confirmed LN before induction therapy. Univariate associations between clinical factors (such as demographics and drugs) and novel urine protein biomarkers were calculated using logistic regression to determine their individual predictive power, as assessed by the AUC. These univariate models were compared with RF classifiers generated with either clinical variables only or combined data (clinical variables and urinary biomarkers). Classification with single biomarkers was generally unsuccessful. Classification with RF models, built using a panel of biomarkers, achieved greater sensitivity than the univariate models. Of note, four of the five statistically significant biomarkers (except glomerular filtration rate) in the univariate models were also in the top five most important features for the RF model[184]. RF models using both traditional and urinary biomarkers as features had improved sensitivity compared with those built with clinical biomarkers only. Although not validated in an external dataset, the authors report that when the data were separated into training and testing portions, the RF model did not overfit[184].

Similarly, ML was employed to predict flares following drug tapering in 41 patients with RA[185]. Patient characteristics, disease activity data, medication data and laboratory data derived from a clinical trial were used as input data. Although limited by a small sample size, the study evaluated multiple ML models (including logistic regression, naive Bayes, *k*-NN and RF) for their ability to accurately predict flares. In addition, the study employed a multistep (stacking) model, in which the results of the four classifiers were used as variables in a logistic regression model. After cross-validation, the stacking logistic regression model achieved the highest AUC of 0.81. Importantly, this study illustrates that ML has the potential to aid in the analysis of clinical trial data, yet, as suggested by the authors, the models need to be validated in a larger cohort and, ideally, employed in a clinical care setting, to determine whether they are applicable in settings outside of tightly controlled trials[185].

ML techniques have also been used to predict drug repurposing candidates for immune-mediated cutaneous diseases, including psoriasis, SLE, scleroderma, myositis and myasthenia gravis[186]. These models involved a multistep ML process. An NLP word-embedding approach was employed to generate features, allowing the authors to extract drug-vocabulary matrices from ~20 million PubMed abstracts. Many different ML algorithms were then employed to classify the drug–disease relationship from the input NLP-derived drug-vocabulary matrices. Following cross-validation, the PLS-DA model achieved the highest AUC. When tested on unlabelled drug-vocabulary matrices, the PLS-DA model successfully identified drugs currently used for psoriasis and predicted additional candidates. Of note, the targets of the predicted candidate drugs were identified to be enriched in differentially expressed genes from psoriasis[186], supporting the potential for ML to aid in drug discovery and generate new hypotheses.

Other studies have used ML to inform drug discovery[105,187,188] or predict drug repurposing candidates[189]. ML tools have been used to probe the structural similarity of drugs or targets[24] or create drug networks containing drugs that are predicted to have similar mechanisms of action[190]. As these networks grow and incorporate multiple types of data, ML is crucial for the synthesis of multimodal information. ML algorithms have also been used to construct databases of drug–adverse effect pairs, primarily by text mining[191], or to predict potential drug-associated adverse events[192]. Many ML algorithms are being employed to understand pharmacological properties of drugs by identifying patterns in drug-induced transcriptomic profiles[193]. Additional repurposing techniques that incorporate ML and that are applicable to or have been utilized in RAIDs have been reviewed eslewhere[24,194].

Patient clustering to determine disease subtypes

Molecular subtyping of patients is often thought of as major step in achieving precision medicine. However, validating the subtypes, which may have been determined by supervised or clustering models, is difficult. ML models based on gene expression data derived from multiple different RNA-seq datasets have been used to cluster and classify patients with SLE[195]. PLS-DA, a supervised clustering algorithm, was applied to samples from four combined studies and the resulting model separated most patients with SLE from controls. The PLS-DA approach provided discriminatory genes with the potential to separate patients with SLE from controls[195,196], but also suggested heterogeneity among patients with SLE. Consequently, patients with SLE were clustered using *k*-means clustering, in which $k=4$ clusters were specified based on statistical evaluation of the appropriate number of clusters. Unsupervised clustering identified patterns in gene expression among patients, and further analysis of clusters revealed that each had specific enrichment for disease features or disease activity[195]. An SVM approach was also employed to classify patients into the four clusters (which were used as labels) derived from *k*-means clustering[21,195]. When the SVM model was trained on one dataset and tested on a combination of other datasets, the model showed 88% accuracy. A cross-validated RF classifier built with 141 patients with SLE from combined datasets was also able to classify the patients with 88% accuracy.

Numerous studies have employed clustering to probe gene expression data. *k*-means clustering of gene expression modules and soluble mediator measurements was employed on another SLE dataset[197]. Seven clusters were identified based on differences in interferon, lymphocyte and monocyte gene modules. Demographics were variable in the clusters defined by gene or protein expression. In another study, 143 patients with SLE were grouped into four clusters based on lymphocyte gene expression, and then the clinical features of patients in each cluster were examined[198]. Some clusters were characterized by higher incidence of tissue or organ manifestations such as nephritis or arthritis, whereas others had lower disease activity. Another clustering study used clinical and biological features to cluster patients with anti-Ku-antibody-positive myositis hierarchically into three clusters[199], although the approach was met with criticism about whether the outcomes were biologically meaningful[200]. The critical review claimed that there is no optimal solution to ML problems aimed at determining the number of clusters within a dataset[200], which may suggest that simple clustering methods are not ideal for forming clinically meaningful patient subsets. Thereby, although there is utility in clustering, it is important to ensure that the assumptions and criteria of the analytical choice are met.

Compared with less complex clustering approaches such as *k*-means clustering or PCA, artificial neural networks have proved to be a useful tool in identifying disease subsets. PCA was compared with an artificial neural network-based model in classifying populations of patients with pSS and determining the most important variables for predicting lymphoma development in patients with pSS[201]. The artificial neural network-based model could classify patients with pSS into two groups based on whether their glandular manifestations were severe or mild. In addition, after feature selection was used to determine the 15 variables most associated with lymphoma development, the artificial neural network was applied to predict whether patients would develop lymphoma. The proof-of-concept model built using these variables demonstrated a sensitivity and specificity of 92.5% and 98% respectively, although validation in an external dataset is still needed.

Lessons from machine learning use in rheumatology

When appropriately employed, ML is an effective and efficient technique for analysing high-dimensional data. ML is powerful, as it can recognize patterns in data that are not easily detected by humans. In addition, the consistent following of algorithm-determined rules and optimization schema in ML can reduce variability in classification or regression, which may be important for clinical standardization. Moreover, ML has been used to integrate multiple types of data for outcome prediction. In RAIDs, ML has been employed to classify patients in EHRs, classify disease activity based on imaging, detect differences in gene expression in active and inactive disease, and predict drug repurposing candidates. However, there are few prospective studies that have been designed to generate data appropriate for ML analysis and subsequently employ this analytical technique. Most ML studies to date have been retrospective, as they have used existing datasets, many of which are not sufficiently large or robust. Consequently, these studies are not suitable to fully appreciate the translational capacity or properly test the utility of ML. For example, the measured data may not be suitable for input into one or many classifiers. The application of ML to available data instead of de novo experimental design might in itself be considered a weakness of this approach at present. Nevertheless, prospective studies still prove beneficial, as they can be used to generate hypotheses about patient cohorts, which may be tested in new studies that are designed with appropriate data collection and sample stratification to achieve a particular goal.

Limitations of machine learning

Despite its advantages, ML can still prove challenging (BOX 1). With the availability of numerous ML algorithms, it is often difficult to determine the most appropriate algorithm a priori. As discussed here, different algorithms have been implemented by different investigators to achieve the same goal (for example, classify patients using EHRs or classify patients' disease activity status). Often, some algorithms may be unfeasible because of the size of the dataset. In addition, algorithm choice might be limited by the characteristics of available data, and the most appropriate model might not be apparent without comparison of several models. The most thorough solutions compare multiple models or combinations of models that could solve the problem. Moreover, the optimal model for the specified problem might require implementation of multiple ML models in

Box 1 | **Limitations of machine learning**

Algorithm selection is not straightforward

- There is no consensus on algorithm selection for specific applications.
- The most appropriate machine learning algorithm for a dataset is dependent on the desired outcome, type of data (numerical versus categorical), number of samples and number of input features.

Quality of the input data may be poor

- The predictive power of a model is dependent on the quality of the input data[21], but available data could be of poor quality or variable between data sources.
- It is difficult to mitigate errors and detect anomalies in data[203].

Accuracy of class labels

- Class labels are dependent on expert definition and improper labelling of data could lead to improper classification of new samples[204].
- Class labels might be dependent on current knowledge in the field and models could recognize additional criteria to subdivide individuals who might be dismissed as inaccurate.

Model fit and interpretation

- Some models could overfit the training data and then fail on new datasets.
- Other models might fail to capture underlying trends of the data.
- There is a trade-off between overfitting and underfitting the data (the bias–variance trade-off[206]) and between the performance of the model and the ability to understand it.
- ML models that are not validated with external datasets or built from small sample sizes must be interpreted with caution.

series (that is, one ML model for preprocessing and one for classification). As such, many studies employing ML evaluate numerous models built with different algorithms and ultimately choose the model yielding the greatest accuracy, as determined by sensitivity, specificity or other metrics at the end of the analysis. Nevertheless, general guidelines exist for selecting the most appropriate model based on the desired outcome and type of data.

Because ML can be thought of as a 'black box', it can be difficult to assess the quality of the model implementation and the validity of the results. Throughout this Review, we have enumerated the importance of the careful preprocessing of data, comparing multiple algorithms to find the one that is most appropriate for the dataset being analysed, and subsequently validating the algorithm using cross-validation and, most crucially, with an independent dataset. If these practices are employed, then one can be confident of the outcome, even if the accuracy or predictive power of the resulting model is less than optimal. A helpful framework for the assessment of model quality has been published, which provides key questions for assessing the employment of preprocessing, cross-validation, sample size, evaluation, hyperparameters and ensembling in ML model construction[202].

In addition, the predictive power of a model is dependent on the quality of the input data[21], which can be affected by data collection methods. Although validation systems can help to detect anomalies in the data, mitigating errors is difficult[203]. As a result, even strong models built on a single training and validation dataset can fail in additional datasets. Furthermore, a supervised model is only as accurate as its original labels. Improper labelling of patients as having 'RA' or 'SLE' in the training dataset for a model that intends to classify whether a new patient has one disease and not the other could lead to improper classification of new samples[204]. Moreover, classification of a patient may be based upon the current knowledge in the field, and the model may recognize additional criteria to subdivide individuals who may be dismissed as inaccurate. As with any kind of model, ML models are subject to differing degrees of accuracy. Overfit models are often characterized as having high variance, which means that they capture the variability of the data points on which the model was built and incorporate these variations into the resulting model structure[89]. Overfit models capture too much of the noise that is present in the data from which they were built[205] and can therefore fail when applied to new datasets, especially when the model encounters new data that are sufficiently different from the data used to build the model. Underfit models are often oversimplified because they make simplifying assumptions about the model's underlying mathematical function[6,89]. As a result, underfit models could fail to capture underlying trends of the data, making them biased. Errors related to model fit can be reduced, but usually at a cost, a phenomenon known as the bias–variance tradeoff[206]. The best ML algorithms balance bias and variance for implementation of a strong predictive model. There is power in the ability of ML to predict information or draw conclusions without the general deductive reasoning paradigm, but because of the issue of fitting, caution with ML outcome interpretation is also necessary.

Finally, there is a tradeoff between model interpretability and model performance. Interpretability refers to the ability to derive biological meaning from the outcome and/or understand the mathematical mechanisms driving the algorithm[207,208], whereas performance is generally assessed by model accuracy. Neural networks are often cited as the most accurate models, but often they are difficult to implement and interpret because of their complex rules and 'hidden layers'. Similarly, SVM models are known to be difficult to interpret[33]. Decision trees are often cited as useful tools because they balance the accuracy–interpretability trade-off well, but they can be improved by ensemble methods[209]. Even if the outcome of a model can be understood biologically, it still might not be 'explainable', meaning that the user does not understand the mathematical relationship between inputs and outputs[207,210]. As such, an entire field termed 'explainable artificial intelligence' has emerged[207], and guidelines for the transparency of clinical trials employing ML-based analysis have been established[211]. However, the most challenging aspect of ML application in practice is validation. ML models are prone to overfitting, and thereby application of the model to a new dataset might produce an entirely different result. In addition, when applied to biomedicine, ML models may extract relationships with no known biological meaning, so it is important to determine whether the ML model might have uncovered biological function that cannot be deducible from a reductionist approach or has uncovered patterns of noise in the data.

Sample size can also affect model accuracy and interpretability. Although there are no definitive cut-off limits

Box 2 | **Using machine learning to improve clinical decision-making and rheumatology research**

Electronic health record data analysis
Machine learning (ML) models built using electronic health record data might be able to relate the features of one patient to thousands of other patients at the same medical practice. This ability could enable prediction of future disease progression, drug utilization, comorbidities or additional patient medical needs. Another ML model using electronic health record data could be used to recognize patterns of organ involvement, laboratory test ordering or medication use that actually suggests a diagnosis before one can be made based on clinical presentation and/or specimen collection by a health-care professional. A third model might be used for exploratory analysis; that is, to find patterns in the data that could be used as a reference and later validated.

Imaging analysis
Using ML, it might be possible to standardize the examination and diagnosis of clinical features from images across multiple institutions. In addition, ML might be able to identify patterns in images that are undetectable to the human eye, which could be used to generate hypotheses or new knowledge about the features of organ or tissue manifestations in rheumatic autoimmune inflammatory diseases.

Biomarker analysis
If potential biomarkers were used as input features for ML-based classification of known groups, analysis of which biomarkers correctly identified and separated the classes might lead to the identification of novel disease biomarkers. In addition, examination of the common features of ML-determined clusters could generate hypotheses about potential biomarkers.

Biometric analysis
The analysis of daily activity data (collected with wearable activity trackers) or other biometric data by ML might allow for the detection of patterns that could aid in the classification of patients with flares or onset of other rheumatic autoimmune inflammatory disease symptoms.

Transcriptomic analysis
ML-enabled analysis of transcriptomic data might allow for identification of new subtypes of patients with similar molecular disease features. These subtypes could then be correlated with known clinical variables and specific, personalized profiles of these patients could be developed; personalized therapy might be proposed because of their unique molecular profile.

Therapeutic analysis
ML has been employed to inform drug discovery and predict drug repurposing candidates. ML analysis could lend insight into patterns of therapeutic response and allow for prediction of patients who do or do not respond to a certain therapy.

for sample size, the general understanding is that more samples allow for a more robust analysis. Unlike other statistical strategies, power analyses are not routinely employed in the design of ML models, as the expected outcome and effect size might be unknown. Although many studies discussed here demonstrate modest to good accuracy in classifying patients or predicting outcomes, further validation in larger cohorts is necessary before the conclusions can be employed in clinical practice.

Altogether, ML is becoming increasingly powerful, but much remains to be tested and discovered. First and foremost, there is a need for larger initial training datasets so that the original data have more variability and available validation and testing datasets so that models can also be evaluated under independent conditions. In addition, improved understanding of model parameters and equations is necessary to determine whether or not a result is biologically meaningful. Ultimately, however, the success of a ML approach will be gauged by its ability to improve clinical care and provide more effective personalized management.

Conclusions

The availability of data, including patient history data (EHR data), gene expression data (which in the future may be derived from an at-home clinical test), wearable diagnostic data (for example, from step trackers) and genetic susceptibility data, might lead to a transformation in rheumatology care by implementation of rapid, effective precision medicine[212]. Indeed, wearable diagnostics were discussed earlier[145] and are already being employed in a clinical trial for both SLE and RA as a means of both tracking activity and helping to reduce inactivity (NCT02554474). In addition, a mobile app has been developed to track patient-reported outcomes in SLE on a daily basis (NCT03142711). Moreover, another clinical trial (NCT04306939) is collecting observational data for RA and other diseases to be analysed with ML in order to develop disease models that define risk factors. The application of ML to the analysis of these and other data is beginning to change the landscape of RAIDs research[213]. The studies explored here illuminate many advances in improving understanding of individual patients' disease using ML and, thereby, the knowledge gained might assist with future efforts in personalized medicine. ML-enabled diagnoses or flare prediction might help to prevent organ damage or aid in the management of disease chronicity. Patient classification and subtyping of gene expression with ML might give an insight into the most appropriate treatment for each patient. Furthermore, examination of a patient's genetic data by ML could help to classify patients on the basis of disease risk and ML analysis of EHR data could improve identification of patients with RAIDs (BOX 2). This progress, coupled with the prevalence of ML use by healthcare systems, demonstrate the rapid expansion and implementation of these techniques. Together, these advances reveal the current momentum in improving patient identification, treatment and personalized medicine with the application of ML when used appropriately.

Published online 2 November 2021

1. Jordan, M. I. & Mitchell, T. M. Machine learning: trends, perspectives, and prospects. *Science* **349**, 255–260 (2015).
2. Samuel, A. L. Some studies in machine learning using the game of checkers IBM journals & magazine. *IBM J. Res. Dev.* **3**, 210–229 (1959).
3. Bhavsar, P., Safro, I., Bouaynaya, N., Polikar, R. & Dera, D. Machine learning in transportation data analytics in Data Analytics for Intelligent Transportation Systems (eds Chowdhury, M., Apon, A. & Dey, K.) 283–307 (Elsevier Inc., 2017).
4. Kubat, M. An Introduction to Machine Learning. (Springer International Publishing, 2017).
5. Hand, D. Statistics and data mining: intersecting disciplines. *ACM SIGKDD Explor. Newsl.* **1**, 16–19 (1999).
6. Kim, K.-J. & Tagkopoulos, I. Application of machine learning in rheumatic disease research. *Korean J. Intern. Med.* **34**, 708–722 (2019).
7. Liao, K. P. et al. Development of phenotype algorithms using electronic medical records and incorporating natural language processing. *BMJ* **350**, h1885 (2015).
8. Turner, C. A. et al. Word2Vec inversion and traditional text classifiers for phenotyping lupus. *BMC Med. Inform. Decis. Mak.* **17**, 126 (2017).
9. Jorge, A. et al. Identifying lupus patients in electronic health records: development and validation of machine learning algorithms and application of rule-based algorithms. *Semin. Arthritis Rheum.* **49**, 84–90 (2019).
10. Zhou, S. M. et al. Defining disease phenotypes in primary care electronic health records by a machine learning approach: a case study in identifying rheumatoid arthritis. *PLoS One* **11**, 1–14 (2016).
11. Norgeot, B. et al. Assessment of a deep learning model based on electronic health record data to forecast clinical outcomes in patients with rheumatoid arthritis. *JAMA Netw. Open.* **2**, e190606 (2019).

12. Walsh, J. A. et al. Identifying axial spondyloarthritis in electronic medical records of US Veterans. *Arthritis Care Res.* **69**, 1414–1420 (2017).
13. Odgers, D. J., Tellis, N., Hall, H. & Dumontier, M. Using LASSO regression to predict rheumatoid arthritis treatment efficacy. *AMIA Jt. Summits Transl. Sci. Proc.* **2016**, 176–83 (2016).
14. Lockshin, M. D., Barbhaiya, M., Izmirly, P., Buyon, J. P. & Crow, M. K. SLE: Reconciling heterogeneity. *Lupus Sci. Med.* **6**, e000280 (2019).
15. McInnes, I. B. Psoriatic arthritis: embracing pathogenetic and clinical heterogeneity? *Clin. Exp. Rheumatol.* **34**, 9–11 (2016).
16. Weyand, C. M., Klimiuk, P. A. & Goronzy, J. J. Heterogeneity of rheumatoid arthritis: from phenotypes to genotypes. *Springer Semin. Immunopathol.* **20**, 5–22 (1998).
17. de Bruijne, M. Machine learning approaches in medical image analysis: From detection to diagnosis. *Med. Image Anal.* **33**, 94–97 (2016).
18. Deeb, S. J. et al. Machine learning-based classification of diffuse large B-cell lymphoma patients by their protein expression profiles. *Mol. Cell. Proteom.* **14**, 2947–60 (2015).
19. Ali, M. & Aittokallio, T. Machine learning and feature selection for drug response prediction in precision oncology applications. *Biophys. Rev.* **11**, 31–39 (2019).
20. Lou, B. et al. An image-based deep learning framework for individualising radiotherapy dose: a retrospective analysis of outcome prediction. *Lancet Digit. Heal.* **1**, e136–e147 (2019).
21. Jiang, M. et al. Machine learning in rheumatic diseases. *Clin. Rev. Allergy Immunol.* **60**, 96–110 (2021).
22. Hügle, M., Omoumi, P., van Laar, J. M., Boedecker, J. & Hügle, T. Applied machine learning and artificial intelligence in rheumatology. Rheumatol. *Adv. Pract.* **4**, rkaa005 (2020).
23. Stoel, B. Use of artificial intelligence in imaging in rheumatology-current status and future perspectives. *RMD Open* **6**, e001063 (2020).
24. Kingsmore, K. M., Grammer, A. C. & Lipsky, P. E. Drug repurposing to improve treatment of rheumatic autoimmune inflammatory diseases. *Nat. Rev. Rheumatol.* **16**, 32–52 (2020).
25. Guan, Y. et al. Machine learning to predict anti-TNF drug responses of rheumatoid arthritis patients by integrating clinical and genetic markers. *Arthritis Rheumatol.* **71**, 1987–1996 (2019).
26. Fautrel, B. et al. Choice of second-line disease-modifying antirheumatic drugs after failure of methotrexate therapy for rheumatoid arthritis: a decision tree for clinical practice based on rheumatologists' preferences. *Arthritis Care Res.* **61**, 425–434 (2009).
27. Eyre, S., Orozco, G. & Worthington, J. The genetics revolution in rheumatology: large scale genomic arrays and genetic mapping. *Nat. Rev. Rheumatol.* **13**, 421–432 (2017).
28. Catalina, M. D. et al. Patient ancestry significantly contributes to molecular heterogeneity of systemic lupus erythematosus. *JCI Insight* **5**, e140380 (2020).
29. Provost, F. & Kohavi., R. Glossary of Terms. *J. Mach. Learn.* **30**, 271–274 (1998).
30. Zhu, X. & Goldberg, A. Introduction to semi-supervised learning. Synth. Lect. Artif. *Intell. Mach. Learn.* **6**, 1–116 (2009).
31. Haldorai, A., Ramu, A. & Suriya, M. Organization internet of things (IoTs): supervised, unsupervised, and reinforcement learning. in EAI/Springer Innovations in Communication and Computing 27–53 (Springer, 2020).
32. Jain, A. K. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* **31**, 651–666 (2010).
33. Kotsiantis, S. B., Zaharakis, I. D. & Pintelas, P. E. Machine learning: a review of classification and combining techniques. *Artif. Intell. Rev.* **26**, 159–190 (2006).
34. Ayodele, T. O. Types of Machine Learning Algorithms. in New Advances in Machine Learning (ed. Zhang, Y.) 19–49 (InTech, 2010).
35. Alasadi, S. A. & Bhaya, W. S. Review of data preprocessing techniques in data mining. *J. Eng. Appl. Sci.* **12**, 4102–4107 (2017).
36. Zhang, Z. Missing data imputation: focusing on single imputation. Ann. Transl. Med. **4**, 9 (2016).
37. Cao, X. H., Stojkovic, I. & Obradovic, Z. A robust data scaling algorithm to improve classification accuracies in biomedical data. *BMC Bioinforma.* **17**, 359 (2016).
38. Han, J., Kamber, M. & Pei, J. Data Transformation and Data Discretization. in Data mining: Concepts and Techniques 111–119 (Elsevier, 2012).
39. Saeys, Y., Inza, I. & Larranaga, P. A review of feature selection techniques in bioinformatics. *Bioinformatics* **23**, 2507–2517 (2007).
40. Tuikkala, J., Elo, L. L., Nevalainen, O. S. & Aittokallio, T. Missing value imputation improves clustering and interpretation of gene expression microarray data. *BMC Bioinforma.* **9**, 202 (2008).
41. Aljuaid, T. & Sasi, S. Proper imputation techniques for missing values in data sets. in Proceedings of the 2016 International Conference on Data Science and Engineering ICDSE 2016 (Institute of Electrical and Electronics Engineers Inc., 2017)
42. Rahman, M. M. & Davis, D. N. Machine Learning-Based Missing Value Imputation Method for Clinical Datasets. in Lecture Notes in Electrical Engineering 245–257 (Springer, Dordrecht, 2013).
43. Raja, P. S. & Thangavel, K. Missing value imputation using unsupervised machine learning techniques. *Soft Comput.* **24**, 4361–4392 (2020).
44. Phung, S., Kumar, A. & Kim, J. A deep learning technique for imputing missing healthcare data. in Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 6513–6516 (Institute of Electrical and Electronics Engineers Inc., 2019).
45. Chowdhury, G. G. Natural language processing. *Annu. Rev. Inf. Sci. Technol.* **37**, 51–89 (2005).
46. Zhang, Y., Jin, R. & Zhou, Z. H. Understanding bag-of-words model: a statistical framework. *Int. J. Mach. Learn. Cybern.* **1**, 43–52 (2010).
47. Kozlowski, A. C., Taddy, M. & Evans, J. A. The geometry of culture: analyzing the meanings of class through word embeddings. *Am. Sociol. Rev.* **84**, 905–949 (2019).
48. McInnes, B. T., Pedersen, T. & Carlis, J. Using UMLS Concept Unique Identifiers (CUIs) for word sense disambiguation in the biomedical domain. *AMIA Annu. Symp. Proc.* **2007**, 533–537 (2007).
49. El Bouchefry, K. & de Souza, R. S. Learning in Big Data: Introduction to Machine Learning. in Knowledge Discovery in Big Data from Astronomy and Earth Observation 225–249 (Elsevier, 2020).
50. Lever, J., Krzywinski, M. & Altman, N. Principal component analysis. *Nat. Methods* **14**, 641–642 (2017).
51. Anowar, F., Sadaoui, S. & Selim, B. Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE). Comput. *Sci. Rev.* **40**, 100378 (2021).
52. Velliangiri, S., Alagumuthukrishnan, S. & Thankumar Joseph, S. I. A review of dimensionality reduction techniques for efficient computation. *Procedia Comput. Sci.* **165**, 104–111 (2019).
53. Guyon, I. & Elisseefl, A. An introduction to feature extraction. in Studies in Fuzziness and Soft Computing Vol. 207 1–25 (Springer, 2006).
54. Kubat, M. Some Practical Aspects to Know About. in An Introduction to Machine Learning 191–210 (Springer International Publishing, 2017).
55. Elashoff, J. C., Elashoff, R. M. & Goldman, G. E. On the choice of variables in classification problems with dichotomous variables. *Biometrika* **54**, 668–670 (1967).
56. Toussaint, G. T. Note on optimal selection of independent binary-valued features for pattern recognition. *IEEE Trans. Inf. Theory* **17**, 618 (1971).
57. Dormann, C. F. et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 27–46 (2013).
58. Stańczyk, U. Feature evaluation by filter, Wrapper and embedded approaches. *Stud. Comput. Intell.* **584**, 29–44 (2015).
59. Ceccarelli, F. et al. Biomarkers of erosive arthritis in systemic lupus erythematosus: application of machine learning models. *PLoS One* **13**, e0207926 (2018).
60. Guyon, I. & Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **3**, 27–46 (2003).
61. Tuv, E. et al. Feature selection with ensembles, artificial variables, and redundancy elimination. *J. Mach. Learn. Res.* **10**, 1341–1366 (2009).
62. Altman, N. & Krzywinski, M. Points of significance: clustering. *Nat. Methods* **14**, 545–546 (2017).
63. Tuv, E. Ensemble learning. in Studies in Fuzziness and Soft Computing (eds Guyon, I., Nikravesh, M., Nikravesh, M. Gunn, S. & Zadeh, L. A.) Vol. 207, 187–204 (Springer, 2006).
64. Dietterich, T. G. Ensemble methods in machine learning. in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) Vol. 1857 1–15 (Springer, 2000).
65. Breiman, L. Bagging predictors. *Mach. Learn.* **24**, 123–140 (1996).
66. Altman, N. & Krzywinski, M. Points of significance: ensemble methods: bagging and random forests. *Nat. Methods* **14**, 933–934 (2017).
67. Drucker, H. Improving regressors using boosting techniques. in 14th International Conference on Machine Learning 107–115 (1997).
68. Natekin, A. & Knoll, A. Gradient boosting machines, a tutorial. *Front. Neurorobot.* **7**, 21 (2013).
69. Schapire, R. E. The Boosting Approach to Machine Learning: An Overview. in Lecture Notes in Statistics 149–171 (Springer, 2003).
70. Snoek, J., Larochelle, H. & Adams, R. P. Practical Bayesian optimization of machine learning algorithms. in Advances in Neural Information Processing Systems Vol. 4 2951–2959 (ACM, 2012).
71. Kubat, M. Probabilities: Bayesian Classifiers. in An Introduction to Machine Learning 19–42 (Springer International Publishing, 2017).
72. Aha, D. W., Kibler, D., Albert, M. K. & Quinian, J. R. Instance-based learning algorithms. *Mach. Learn.* **6**, 37–66 (1991).
73. Brownlee, J. Master machine learning algorithms discover how they work and implement them from scratch. *Mach. Learn. Master.* **1**, 11 (2016).
74. Fu, W. J. Penalized regressions: the bridge versus the lasso? *J. Comput. Graph. Stat.* **7**, 397–416 (1998).
75. Tharwat, A., Gaber, T., Ibrahim, A. & Hassanien, A. E. Linear discriminant analysis: a detailed tutorial. *AI Commun.* **30**, 169–190 (2017).
76. Krogh, A. What are artificial neural networks? *Nat. Biotechnol.* **26**, 195–197 (2008).
77. Cross, S. S., Harrison, R. F. & Kennedy, R. L. Introduction to neural networks. *Lancet* **346**, 1075–1079 (1995).
78. Ceccarelli, F. et al. Prediction of chronic damage in systemic lupus erythematosus by using machine-learning models. *PLoS One* **12**, e0174200 (2017).
79. Ruder, S. An overview of gradient descent optimization algorithms. Preprint at *arXiv* **1609**, 04747 (2016).
80. O'Shea, K. & Nash, R. An introduction to convolutional neural networks. Preprint at *arXiv* **1511**, 08458v2 (2015).
81. Medsker, L. R. & Jaub, L. C. Recurrent Neural Networks: Design and Applications (CRC Press, 2001).
82. Arnold, L., Rebecchi, S., Chevallier, S. & Paugam-Moisy, H. An introduction to deep learning. in ESANN 2011 proceedings, 19th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning 477–488 (IEEE, 2010).
83. Ikonomakis, M., Kotsiantis, S. & Tampakas, V. Text classification using machine learning techniques. *WSEAS Trans. Comput.* **4**, 966–974 (2005).
84. Kubat, M. Decision Trees. in An Introduction to Machine Learning 113–136 (Springer International Publishing, 2017).
85. Luo, G. A review of automatic selection methods for machine learning algorithms and hyper-parameter values. *Netw. Model. Anal. Heal. Inform. Bioinforma.* **5**, 18 (2016).
86. Probst, P. & Bischl, B. Tunability: importance of hyperparameters of machine learning algorithms. *J. Mach. Learn. Res.* **20**, 1–32 (2019).
87. Bergstra, J. & Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **13**, 281–305 (2012).
88. Feurer, M. & Hutter, F. Hyperparameter Optimization. in Automated Machine Learning: Methods, Systems, Challenges 3–33 (Springer, 2019).
89. Lever, J., Krzywinski, M. & Altman, N. Points of Significance: model selection and overfitting. *Nat. Methods* **13**, 703–704 (2016).
90. Kim, J. H. Estimating classification error rate: repeated cross-validation, repeated hold-out and bootstrap. *Comput. Stat. Data Anal.* **53**, 3735–3745 (2009).
91. Schneider, J. Cross validation. Definitions https://www.cs.cmu.edu/~schneide/tut5/node42.html (1997).
92. Ross, K. A. et al. Cross-validation. in Encyclopedia of Database Systems 532–538 (Springer US, 2009).
93. Vabalas, A., Gowen, E., Poliakoff, E. & Casson, A. J. Machine learning algorithm validation with a limited sample size. *PLoS One* **14**, e0224365 (2019).

94. Lever, J., Krzywinski, M. & Altman, N. Points of significance: classification evaluation. *Nat. Methods* **13**, 603–604 (2016).
95. Kumar, R. & Indrayan, A. Receiver operating characteristic (ROC) curve for medical researchers. *Indian Pediatrics* **48**, 277–287 (2011).
96. Altman, N. & Krzywinski, M. Points of significance: regression diagnostics. *Nat. Methods* **13**, 385–386 (2016).
97. Handelman, G. S. et al. Peering into the black box of artificial intelligence: evaluation metrics of machine learning methods. *Am. J. Roentgenol.* **212**, 38–43 (2019).
98. Nantasenamat, C. How to build a machine learning model. Towards Data Science. https://towardsdatascience.com/how-to-build-a-machine-learning-model-439ab8fb3fb1 (2018).
99. Chai, T. & Draxler, R. R. Root mean square error (RMSE) or mean absolute error (MAE)? — arguments against avoiding RMSE in the literature. *Geosci. Model. Dev.* **7**, 1247–1250 (2014).
100. Chicco, D., Warrens, M. J. & Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peer J. Comput. Sci.* **7**, e623 (2021).
101. Alpaydin, E. Introduction to Machine Learning (Adaptive Computation and Machine Learning series) (The MIT Press, 2009).
102. Bas¸tanlar, Y. & Özuysal, M. Introduction to machine learning. *Methods Mol. Biol.* **1107**, 105–128 (2014).
103. Libbrecht, M. W. & Noble, W. S. Machine learning applications in genetics and genomics. *Nat. Rev. Genet.* **16**, 321–332 (2015).
104. Camacho, D. M., Collins, K. M., Powers, R. K., Costello, J. C. & Collins, J. J. Next-generation machine learning for biological networks. *Cell* **173**, 1581–1592 (2018).
105. Vamathevan, J. et al. Applications of machine learning in drug discovery and development. *Nat. Rev. Drug. Discov.* **18**, 463–477 (2019).
106. Stafford, I. S. et al. A systematic review of the applications of artificial intelligence and machine learning in autoimmune diseases. NPJ Digit. Med. **3**, 30 (2020).
107. Feldman, C. H. et al. Supplementing claims data with electronic medical records to improve estimation and classification of rheumatoid arthritis disease activity: a machine learning approach. *ACR Open. Rheumatol.* **1**, 552–559 (2019).
108. Barnado, A. et al. Developing electronic health record algorithms that accurately identify patients with systemic lupus erythematosus. *Arthritis Care Res.* **69**, 687–693 (2017).
109. Xiong, W. W. et al. Real-world electronic health record identifies antimalarial underprescribing in patients with lupus nephritis. *Lupus* **28**, 977–985 (2019).
110. Barnado, A. et al. Phenome-wide association study identifies dsDNA as a driver of major organ involvement in systemic lupus erythematosus. *Lupus* **28**, 66–76 (2019).
111. Barnado, A. et al. Phenome-wide association studies uncover a novel association of increased atrial fibrillation in male patients with systemic lupus erythematosus. *Arthritis Care Res.* **70**, 1630–1636 (2018).
112. Doss, J., Mo, H., Carroll, R. J., Crofford, L. J. & Denny, J. C. Phenome-wide association study of rheumatoid arthritis subgroups identifies association between seronegative disease and fibromyalgia. *Arthritis Rheumatol.* **69**, 291–300 (2017).
113. Zhao, S. S. et al. Incorporating natural language processing to improve classification of axial spondyloarthritis using electronic health records. *Rheumatology* **59**, 1059–1065 (2020).
114. Deodhar, A. et al. Use of machine learning techniques in the development and refinement of a predictive model for early diagnosis of ankylosing spondylitis. *Clin. Rheumatol.* **39**, 975–982 (2020).
115. Walsh, J. A., Rozycki, M., Yi, E. & Park, Y. Application of machine learning in the diagnosis of axial spondyloarthritis. *Curr. Opin. Rheumatol.* **31**, 362–367 (2019).
116. Moores, K. G. & Sathe, N. A. A systematic review of validated methods for identifying systemic lupus erythematosus (SLE) using administrative or claims data. *Vaccine* **31**, K62–73 (2013).
117. Murray, S. G., Avati, A., Schmajuk, G. & Yazdany, J. Automated and flexible identification of complex disease: building a model for systemic lupus erythematosus using noisy labeling. *J. Am. Med. Inform. Assoc.* **26**, 61–65 (2019).
118. Liao, K. P. et al. Electronic medical records for discovery research in rheumatoid arthritis. *Arthritis Care Res.* **62**, 1120–1127 (2010).
119. Carroll, R. J. et al. Portability of an algorithm to identify rheumatoid arthritis in electronic health records. *J. Am. Med. Inform. Assoc.* **19**, e162–9 (2012).
120. Ross, B. C. Mutual information between discrete and continuous data sets. *PLoS One* **9**, e87357 (2014).
121. Geurts, P., Ernst, D. & Wehenkel, L. Extremely randomized trees. *Mach. Learn.* **63**, 3–42 (2006).
122. Bellou, E., James, K., Ng, W. F. & Hallinan, J. Machine learning of fatigue-related clinical features in primary Sjogren's Syndrome. *Int. Symp. Sjogrens Syndr.* **81**, 363–364 (2015).
123. Donelle, J. A., Wang, S. X. & Caffery, B. Differentiating between Sjogren's syndrome and dry eye disease: an analysis using random forests. *J. Math.* **5**, 22–36 (2012).
124. Kalweit, M. et al. Personalized prediction of disease activity in patients with rheumatoid arthritis using an adaptive deep neural network. *PLoS One* **16**, e0252289 (2021).
125. Adamichou, C. et al. Lupus or not? SLE Risk Probability Index (SLERPI): a simple, clinician-friendly machine learning-based model to assist the diagnosis of systemic lupus erythematosus. *Ann. Rheum. Dis.* **80**, 758–766 (2021).
126. Toro-Domínguez, D. et al. Differential treatments based on drug-induced gene expression signatures and longitudinal systemic lupus erythematosus stratification. *Sci. Rep.* **9**, 15502 (2019).
127. Toro-Domínguez, D. et al. Stratification of systemic lupus erythematosus patients into three groups of disease activity progression according to longitudinal gene expression. *Arthritis Rheumatol.* **70**, 2025–2035 (2018).
128. Andersen, J. K. H. et al. Neural networks for automatic scoring of arthritis disease activity on ultrasound images. *RMD Open* **5**, e000891 (2019).
129. Tang, J. et al. Grading of rheumatoid arthritis on ultrasound images with deep convolutional neural network. in IEEE International Ultrasonics Symposium (IEEE Computer Society, 2018).
130. Tang, J. et al. Enhancing convolutional neural network scheme for rheumatoid arthritis grading with limited clinical data. *Chin. Phys. B* **28**, 038701 (2019).
131. Üreten, K., Erbay, H. & Maras¸, H. H. Detection of rheumatoid arthritis from hand radiographs using a convolutional neural network. *Clin. Rheumatol.* **39**, 969–974 (2020).
132. Murakami, S., Hatano, K., Tan, J., Kim, H. & Aoki, T. Automatic identification of bone erosions in rheumatoid arthritis from hand radiographs based on deep convolutional neural network. *Multimed. Tools Appl.* **77**, 10921–10937 (2018).
133. Rohrbach, J., Reinhard, T., Sick, B. & Dürr, O. Bone erosion scoring for rheumatoid arthritis with deep convolutional neural networks. *Comput. Electr. Eng.* **78**, 472–481 (2019).
134. Betancourt-Hernández, M., Viera-López, G. & Serrano-Muñoz, A. Automatic diagnosis of rheumatoid arthritis from hand radiographs using convolutional neural networks. *Rev. Cuba. Fis.* **35**, 39–43 (2018).
135. Hemalatha, R. J., Vijaybaskar, V. & Thamizhvani, T. R. Automatic localization of anatomical regions in medical ultrasound images of rheumatoid arthritis using deep learning. *Proc. Inst. Mech. Eng. Part. H. J. Eng. Med.* **233**, 657–667 (2019).
136. Dehghani, H., Feng, Y., Lighter, D., Zhang, L. & Wang, Y. Deep neural networks improve diagnostic accuracy of rheumatoid arthritis using diffuse optical tomography. in Optics InfoBase Conference Papers (SPIE-Intl Soc Optical Eng, 2019).
137. Vukicevic, A., Zabotti, A., de Vita, S. & Filipovic, N. Assessment of machine learning algorithms for the purpose of primary Sjögren's syndrome grade classification from segmented ultrasonography images. in Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering. *LNICST* **241**, 239–245 (2018).
138. Kise, Y. et al. Preliminary study on the application of deep learning system to diagnosis of Sjögren's syndrome on CT images. *Dentomaxillofacial Radiol.* **48**, 20190019 (2019).
139. Simos, N. J. et al. Machine learning classification of neuropsychiatric systemic lupus erythematosus patients using resting-state fmri functional connectivity. in IST 2019 — IEEE International Conference on Imaging Systems and Techniques, Proceedings (Institute of Electrical and Electronics Engineers Inc., 2019).
140. Morita, K., Tashita, A., Nii, M. & Kobashi, S. Computer-aided diagnosis system for Rheumatoid Arthritis using machine learning. in Proceedings of 2017 International Conference on Machine Learning and Cybernetics Vol. 2 357–360 (IEEE, 2017).
141. Joo, Y. B., Baek, I. W., Park, Y. J., Park, K. S. & Kim, K. J. Machine learning-based prediction of radiographic progression in patients with axial spondyloarthritis. *Clin. Rheumatol.* **39**, 983–991 (2020).
142. Sharon, H., Elamvazuthi, I., Lu, C. K., Parasuraman, S. & Natarajan, E. Development of rheumatoid arthritis classification from electronic image sensor using ensemble method. *Sensors* **20**, 167 (2020).
143. Simos, N. J. et al. Quantitative identification of functional connectivity disturbances in neuropsychiatric lupus based on resting-state fMRI: a robust machine learning approach. *Brain Sci.* **10**, 777 (2020).
144. Castro-Zunti, R., Park, E. H., Choi, Y., Jin, G. Y. & Ko, S. B. Early detection of ankylosing spondylitis using texture features and statistical machine learning, and deep learning, with some patient age analysis. *Comput. Med. Imaging Graph.* **82**, 101718 (2020).
145. Gossec, L. et al. Detection of flares by decrease in physical activity, collected using wearable activity trackers in rheumatoid arthritis or axial spondyloarthritis: an application of machine learning analyses in rheumatology. *Arthritis Care Res.* **71**, 1336–1343 (2019).
146. Andreu-Perez, J. et al. Developing fine-grained actigraphies for rheumatoid arthritis patients from a single accelerometer using machine learning. *Sensors* **17**, 2113 (2017).
147. Oates, J. C. et al. Prediction of urinary protein markers in lupus nephritis. *Kidney Int.* **68**, 2588–2592 (2005).
148. Tang, Y. et al. Lupus nephritis pathology prediction with clinical indices. *Sci. Rep.* **8**, 10231 (2018).
149. Robinson, G. A. et al. Disease-associated and patient-specific immune cell signatures in juvenile-onset systemic lupus erythematosus: patient stratification using a machine-learning approach. *Lancet Rheumatol.* **2**, e485–e496 (2020).
150. Choi, M. Y. & Ma, C. Making a big impact with small datasets using machine-learning approaches. *Lancet Rheumatol.* **2**, e451–e452 (2020).
151. Ormseth, M. J. et al. Development and validation of a MicroRNA panel to differentiate between patients with rheumatoid arthritis or systemic lupus erythematosus and controls. *J. Rheumatol.* **47**, 188–196 (2020).
152. Labonte, A. C. et al. Identification of alterations in macrophage activation associated with disease activity in systemic lupus erythematosus. *PLoS One* **13**, e0208132 (2018).
153. Kegerreis, B. et al. Machine learning approaches to predict lupus disease activity from gene expression data. *Sci. Rep.* **9**, 9617 (2019).
154. Orange, D. E. et al. Identification of three rheumatoid arthritis disease subtypes by machine learning integration of synovial histologic features and RNA sequencing data. *Arthritis Rheumatol.* **70**, 690–701 (2018).
155. Ghosh, J. & Acharya, A. Cluster ensembles. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **1**, 305–315 (2011).
156. Lu, R. et al. Immunologic findings precede rapid lupus flare after transient steroid therapy. *Sci. Rep.* **9**, 8590 (2019).
157. Bentham, J. et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. *Nat. Genet.* **47**, 1457–1464 (2015).
158. Morris, D. L. et al. Genome-wide association meta-analysis in Chinese and European individuals identifies ten new loci associated with systemic lupus erythematosus. *Nat. Genet.* **48**, 940–946 (2016).
159. Stahl, E. A. et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat. Genet.* **42**, 508–514 (2010).
160. International Genetics of Ankylosing Spondylitis Consortium (IGAS). et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat. Genet.* **45**, 730–8 (2013).
161. Bowes, J. et al. Dense genotyping of immune-related susceptibility loci reveals new insights into the genetics of psoriatic arthritis. *Nat. Commun.* **6**, 6046 (2015).

162. Li, Y. et al. A genome-wide association study in Han Chinese identifies a susceptibility locus for primary Sjögren's syndrome at 7q11.23. *Nat. Genet.* **45**, 1361–1365 (2013).
163. Almlöf, J. C. et al. Novel risk genes for systemic lupus erythematosus predicted by random forest classification. *Sci. Rep.* **7**, 6236 (2017).
164. Briggs, F. B. S. et al. Supervised machine learning and logistic regression identifies novel epistatic risk factors with PTPN22 for rheumatoid arthritis. *Genes. Immun.* **11**, 199–208 (2010).
165. Glaser, B. et al. Analyses of single marker and pairwise effects of candidate loci for rheumatoid arthritis using logistic regression and random forests. *BMC Proc.* **1**, S54 (2007).
166. Croiseau, P. & Cordell, H. J. Analysis of North American Rheumatoid Arthritis Consortium data using a penalized logistic regression approach. *BMC Proc.* **3**, S61 (2009).
167. Vignal, C. M., Bansal, A. T. & Balding, D. J. Using penalised logistic regression to fine map HLA variants for rheumatoid arthritis. *Ann. Hum. Genet.* **75**, 655–664 (2011).
168. Bartoloni, E. et al. Application of artificial neural network analysis in the evaluation of cardiovascular risk in primary Sjögren's syndrome: a novel pathogenetic scenario? *Clin. Exp. Rheumatol.* **37**, S133–S139 (2019).
169. Navarini, L. et al. A machine-learning approach to cardiovascular risk prediction in psoriatic arthritis. *Rheumatology* **59**, 1767–1769 (2020).
170. Navarini, L. et al. Cardiovascular risk prediction in ankylosing spondylitis: from traditional scores to machine learning assessment. *Rheumatol. Ther.* **7**, 867–882 (2020).
171. Ravenell, R. L. et al. Premature atherosclerosis is associated with hypovitaminosis D and angiotensin-converting enzyme inhibitor non-use in lupus patients. *Am. J. Med. Sci.* **344**, 268–273 (2012).
172. Reddy, B. K. & Delen, D. Predicting hospital readmission for lupus patients: an RNN-LSTM-based deep-learning methodology. *Comput. Biol. Med.* **101**, 199–209 (2018).
173. Hong, S. et al. Longitudinal profiling of human blood transcriptome in healthy and lupus pregnancy. *J. Exp. Med.* **216**, 1154–1169 (2019).
174. Chen, Y. et al. Machine learning for prediction and risk stratification of lupus nephritis renal flare. *Am. J. Nephrol.* **52**, 152–160 (2021).
175. Babajide Mustapha, I. & Saeed, F. Bioactive molecule prediction using extreme gradient boosting. *Molecules* **21**, 983 (2016).
176. Nair, N. & Wilson, A. G. Can machine learning predict responses to TNF inhibitors? *Nat. Rev. Rheumatol.* **15**, 702–704 (2019).
177. Plenge, R. M. et al. Crowdsourcing genetic prediction of clinical utility in the rheumatoid arthritis responder challenge. *Nat. Genet.* **45**, 468–469 (2013).
178. Tao, W. et al. Multiomics and machine learning accurately predict clinical response to adalimumab and etanercept therapy in patients with rheumatoid arthritis. *Arthritis Rheumatol.* **73**, 212–222 (2021).
179. Plant, D. & Barton, A. Machine learning in precision medicine: lessons to learn. *Nat. Rev. Rheumatol.* **17**, 5–6 (2021).
180. Van Looy, D. et al. Comparing statistics with machine learning models to predict dose increase of infliximab for rheumatoid arthritis patients. in Proc. 9th IASTED Int. Conf. Artif. Intell. Soft Computing, ASC 195–200 (ACTA Press, 2005).
181. Lee, S. et al. Machine learning to predict early TNF inhibitor users in patients with ankylosing spondylitis. *Sci. Rep.* **10**, 20299 (2020).
182. Seridi, L. et al. OP0161 association of baseline cytotoxic gene expression with ustekinumab response in systemic lupus erythematosus. *Ann. Rheum. Dis.* **79**, 101–102 (2020).
183. Gottlieb, A. B. et al. Secukinumab efficacy in psoriatic arthritis. *JCR* **27**, 239–247 (2021).
184. Wolf, B. J. et al. Development of biomarker models to predict outcomes in lupus nephritis. *Arthritis Rheumatol.* **68**, 1955–1963 (2016).
185. Vodencarevic, A. et al. Advanced machine learning for predicting individual risk of flares in rheumatoid arthritis patients tapering biologic drugs. *Arthritis Res. Ther.* **23**, 67 (2021).
186. Patrick, M. T. et al. Drug repurposing prediction for immune-mediated cutaneous diseases using a word-embedding-based machine learning approach. *J. Invest. Dermatol.* **139**, 683–691 (2019).
187. Ekins, S. et al. Exploiting machine learning for end-to-end drug discovery and development. *Nat. Mater.* **18**, 435–441 (2019).
188. Lavecchia, A. Deep learning in drug discovery: opportunities, challenges and future prospects. *Drug. Discov. Today* **24**, 2017–2032 (2019).
189. Kuang, Z. et al. A machine-learning-based drug repurposing approach using baseline regularization. *Methods Mol. Biol.* **1903**, 255–267 (2019).
190. Zeng, X. et al. DeepDR: a network-based deep learning approach to in silico drug repositioning. *Bioinformatics* **35**, 5191–5198 (2019).
191. Xu, R. & Wang, Q. Q. Automatic construction of a large-scale and accurate drug-side-effect association knowledge base from biomedical literature. *J. Biomed. Inform.* **51**, 191–199 (2014).
192. Bresso, E. et al. Integrative relational machine-learning for understanding drug side-effect profiles. *BMC Bioinforma.* **14**, 207 (2013).
193. Aliper, A. et al. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. *Mol. Pharm.* **13**, 2524–2530 (2016).
194. Grammer, A. C. & Lipsky, P. E. Drug repositioning strategies for the identification of novel therapies for rheumatic autoimmune inflammatory diseases. *Rheum. Dis. Clin. North. Am.* **43**, 467–480 (2017).
195. Figgett, W. A. et al. Machine learning applied to whole-blood RNA-sequencing data uncovers distinct subsets of patients with systemic lupus erythematosus. *Clin. Transl. Immunol.* **8**, e01093 (2019).
196. Catalina, M. D., Owen, K. A., Labonte, A. C., Grammer, A. C. & Lipsky, P. E. The pathogenesis of systemic lupus erythematosus: harnessing big data to understand the molecular basis of lupus. *J. Autoimmun.* **110**, 102359 (2020).
197. Guthridge, J. M. et al. Adults with systemic lupus exhibit distinct molecular phenotypes in a cross-sectional study. *EClinicalMedicine* **20**, 100291 (2020).
198. Lu, Z., Li, W., Tang, Y., Da, Z. & Li, X. Lymphocyte subset clustering analysis in treatment-naive patients with systemic lupus erythematosus. *Clin. Rheumatol.* **40**, 1835–1842 (2021).
199. Spielmann, L. et al. Anti-Ku syndrome with elevated CK and anti-Ku syndrome with anti-dsDNA are two distinct entities with different outcomes. *Ann. Rheum. Dis.* **78**, 1101–1106 (2019).
200. Pinal-Fernandez, I. & Mammen, A. L. On using machine learning algorithms to define clinically meaningful patient subgroups. *Ann. Rheum. Dis.* **79**, e128 (2020).
201. Baldini, C., Ferro, F., Luciano, N., Bombardieri, S. & Grossi, E. Artificial neural networks help to identify disease subsets and to predict lymphoma in primary Sjögren's syndrome. *Clin. Exp. Rheumatol.* **36**, S137–S144 (2018).
202. Delgadillo, J. Machine learning: a primer for psychotherapy researchers. *Psychother. Res.* **31**, 1–4 (2021).
203. Breck, E., Polyzotis, N., Roy, S., Whang, S. E. & Zinkevich, M. Data Validation for Machine Learning. in Proceedings of the 2nd SysML Conference (Palo Alto Networks, 2019).
204. Kubat, M. A Simple Machine-Learning Task. in An Introduction to Machine Learning (Springer International Publishing, 2017).
205. Van Der Aalst, W. M. P. et al. Process mining: a two-step approach to balance between underfitting and overfitting. *Softw. Syst. Model.* **9**, 87–111 (2010).
206. Schaffer, C. Overfitting avoidance as bias. *Mach. Learn.* **10**, 153–178 (1993).
207. Adadi, A. & Berrada, M. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE Access.* **6**, 52138–52160 (2018).
208. Tjoa, E. & Guan, C. A Survey on explainable artificial intelligence (XAI): toward medical XAI. *IEEE Trans. Neural Netw. Learn Syst*. https://doi.org/10.1109/TNNLS.2020.3027314 (2020).
209. Kingsford, C. & Salzberg, S. L. What are decision trees? *Nat. Biotechnol.* **26**, 1011–1012 (2008).
210. Doran, D., Schulz, S. & Besold, T. R. What does explainable AI really mean? A new conceptualization of perspectives. in CEUR Workshop Proceedings Vol. 2071 (CEUR-WS, 2018).
211. Cruz Rivera, S. et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Lancet Dig. Health* **2**, e549–e560 (2020).
212. Burmester, G. R. Rheumatology 4.0: big data, wearables and diagnosis by computer. *Ann. Rheum. Dis.* **77**, 963–965 (2018).
213. Pandit, A. & Radstake, T. R. D. J. Machine learning in rheumatology approaches the clinic. *Nat. Rev. Rheumatol.* **16**, 69–70 (2020).
214. Yang, S. & Berdine, G. The receiver operating characteristic (ROC) curve. *Southwest. Respir. Crit. Care Chron.* **5**, 34 (2017).

Acknowledgements
The authors thank P. Bachali, S. Shrotri, K. Bell, and J. Kain for helpful discussion about machine learning concepts. The authors thank Dr. C. Nantasenamat for allowing us to modify his figure about the workflow of ML. This work was supported by funding from the RILITE Foundation.

Author contributions
K. M. K. and C. E. P. researched data for the article. K. M. K., C. E. P., A. C. G. and P. E. L. contributed substantially to discussion of the content. K. M. K., C. E. P. and P. E. L. wrote the article. K. M. K. C.E. P. and P. E. L. reviewed and/or edited the manuscript before submission.

Competing interests
K.M.K. and C.E.P. were employed by AMPEL BioSolutions, LLC, during the preparation of this work. K.M.K. was additionally employed by the RILITE Research Institute during the preparation of this work. A.C.G. and P.E.L. are the founders of AMPEL BioSolutions, LLC. The authors declare that the content of this manuscript is not related to AMPEL BioSolutions, LLC's commercial activities. AMPEL uses machine learning as one technique in our analyses pipelines, but does not have a proprietary interest in machine learning as a technology or commercial interest in a specific classifier, regressor or clustering approach. All of the material described in the manuscript is freely available in the public domain.

Peer review information
Nature Reviews Rheumatology thanks M. Krusche and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information
The online version contains supplementary material available at https://doi.org/10.1038/s41584-021-00708-w.

Multisystem inflammatory syndrome in children and Kawasaki disease: a critical comparison

Chetan Sharma[1,12]✉, *Madhusudan Ganigara*[2,12], *Caroline Galeotti*[3], *Joseph Burns*[4], *Fernando M. Berganza*[5], *Denise A. Hayes*[6], *Davinder Singh-Grewal*[7], *Suman Bharath*[8], *Sujata Sajjan*[9] *and Jagadeesh Bayry*[10,11]✉

Abstract | Children and adolescents infected with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) are predominantly asymptomatic or have mild symptoms compared with the more severe coronavirus disease 2019 (COVID-19) described in adults. However, SARS-CoV-2 is also associated with a widely reported but poorly understood paediatric systemic vasculitis. This multisystem inflammatory syndrome in children (MIS-C) has features that overlap with myocarditis, toxic-shock syndrome and Kawasaki disease. Current evidence indicates that MIS-C is the result of an exaggerated innate and adaptive immune response, characterized by a cytokine storm, and that it is triggered by prior SARS-CoV-2 exposure. Epidemiological, clinical and immunological differences classify MIS-C as being distinct from Kawasaki disease. Differences include the age range, and the geographical and ethnic distribution of patients. MIS-C is associated with prominent gastrointestinal and cardiovascular system involvement, admission to intensive care unit, neutrophilia, lymphopenia, high levels of IFNγ and low counts of naive $CD4^+$ T cells, with a high proportion of activated memory T cells. Further investigation of MIS-C will continue to enhance our understanding of similar conditions associated with a cytokine storm.

Our understanding of the coronavirus disease 2019 (COVID-19) pandemic caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has improved greatly since the first human cases were reported in December 2019 in Wuhan City, China[1,2]. COVID-19 is known to involve multiple organ systems, with the major disease burden resulting from respiratory, cardiovascular, thrombotic and neurological complications[3–5]. Cellular entry of SARS-CoV-2 depends on binding of the viral spike (S) protein to cellular receptors such as angiotensin-converting enzyme 2 (ACE2) receptor, which is expressed in multiple organ systems[6,7], and on S-protein priming by host-cell proteases[8,9]. In some individuals these steps are followed by a cascade of inflammatory events, resulting in a 'cytokine storm'[7]. This massive pro-inflammatory cellular and cytokine response is a feature of patients with severe COVID-19 disease[10].

Although morbidity and mortality from primary COVID-19 infection have remained limited in children, we have witnessed the emergence of a new inflammatory disorder associated with COVID-19, termed multisystem inflammatory syndrome in children (MIS-C) in the USA and paediatric inflammatory multisystemic syndrome (PIMS) in Europe[11–15]. Current evidence suggests that MIS-C is a post-infectious, immunologically mediated disorder related to prior SARS-CoV-2 exposure or infection[16–18]. Epidemiological, clinical and immunological investigations have revealed that MIS-C has phenotypic similarities to Kawasaki disease, a childhood inflammatory vasculitis, and it has been suggested that SARS-CoV-2 acts as an additional infectious trigger of Kawasaki disease, leading to an exaggerated phenotype along the same disease spectrum. However, in this Review we present evidence that although MIS-C has some features that overlap with Kawasaki disease, they are distinct syndromes that differ in degrees of hyperinflammation and dysregulated immune responses (TABLE 1).

Overview of Kawasaki disease

Kawasaki disease is a paediatric, self-limited, systemic inflammatory vasculitis that was first described in 1967 in Japan by Dr Tomisaku Kawasaki[19]. The most important long-term sequelae of Kawasaki disease relate to abnormalities of the coronary artery, and it is now the

✉e-mail: Chetan.Sharma@bcm.edu; bayry@iitpkd.ac.in

https://doi.org/10.1038/s41584-021-00709-9

Key points

- Multisystem inflammatory syndrome in children (MIS-C) is characterized by exaggerated innate and adaptive immune responses following infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in predisposed children.
- Clinical presentation of MIS-C involves multiple organ systems, with prominent involvement of the gastrointestinal and cardiovascular systems.
- The factors that trigger the development of MIS-C in children exposed to or infected with SARS-CoV-2 are not yet known.
- Results from epidemiological, clinical and immunological investigations have revealed that although MIS-C has phenotypic similarities to Kawasaki disease, they are different syndromes.
- The approach to treatment of MIS-C aims to mute the augmented inflammatory response.

most common cause of acquired heart disease in children in the developed world[20]. A diagnostic feature of Kawasaki disease is fever that persists for more than 5 days when untreated. Additional typical clinical features include polymorphic skin rash (erythema), involvement of lips and oral mucosa (lip fissures, strawberry tongue), lymphadenopathy (cervical, often unilateral), non-exudative bilateral conjunctivitis and extremity changes (erythema and oedema of palms and soles that desquamate after 2–3 weeks, usually seen in the subacute phase)[20].

Aetiology

The aetiology of Kawasaki disease is uncertain, and there is no single specific diagnostic test. The general consensus, based on results from multiple studies, is that Kawasaki disease is an immune-mediated disease triggered by infection (or infections) in patients with a genetic predisposition[21–25]. Some epidemiological features offer clues to the pathogenesis of Kawasaki disease. It is typically noted in children between the ages of 6 months and 5 years, with an estimated incidence of 25 cases per 100,000 children younger than 5 years in North America[20,26]. It is believed that children younger than 6 months, who have immature immune systems, are protected by passive immunity provided by the transplacental transfer of maternal antibodies, whereas children older than 5 years have developed protective antibody responses to the ubiquitous antigens that most encounter uneventfully in early childhood[27]. There is a male predominance (~1.5:1) in the incidence of Kawasaki disease, a feature that is shared by many common childhood infectious diseases[28,29].

Seasonal variation in the incidence of Kawasaki disease has been noted, with peak incidence occurring in winter and spring in the USA and UK, and in summer in China and Korea[30–34]. Seasonal variation is least evident in Japan, the country with the highest incidence of Kawasaki disease[35,36]. Geographical variation and clustering in the incidence of Kawasaki disease also occurs, with the highest incidence reported in Japan, China, South Korea and Taiwan[35,37–39]. These epidemiological features point towards a transmissible infectious agent, which tends to occur in certain regions of the world with a seasonal variation in its incidence. Evidence exists for the presence of concurrent infections (with bacteria or common respiratory viruses, including coronaviruses) in patients with Kawasaki disease[40–42]. Immunohistochemistry analyses have shown infiltration of IgA plasma cells indicative of the antigen-driven immune response in inflamed tissues and the presence of cytoplasmic antigens suggestive of an infectious aetiology in bronchial and vascular endothelial cells and macrophages[43]. However, to date no single organism has been directly proved to cause Kawasaki disease[44,45].

Involvement of superantigens

The potential pathogenic role of superantigens has been evaluated, on the basis of observations of preferential expression of T cell receptor (TCR) β genes encoding variable regions Vβ2 and Vβ8.1 in the peripheral blood lymphocytes of patients with acute Kawasaki disease[46–48]. Superantigen activity has been identified in the gut microbiota of such patients, and culture supernatants of these bacteria contain a heat shock protein (Hsp60, also known as GroEL) that induces T cell division and production of pro-inflammatory cytokines[49]. However, in studies using flow cytometry in large series of patients with Kawasaki disease, TCR skewing and over-presentation of the described TCR clones has not been found, and it is currently believed that Kawasaki disease is a result of T cell activation by a conventional antigen[50,51].

Involvement of nutritional disorders

The role of nutritional disorders, including vitamin D deficiency, in the pathogenesis of Kawasaki disease is subject to debate[52]. Vitamin D has an anti-inflammatory effect mediated through elevation of expression of IL-10 and inhibition of expression of vascular endothelial growth factor[53,54]. Results from a German population-based study showed that vitamin D supplementation has a protective effect against the development of Kawasaki disease[55]. Low serum concentrations of vitamin D might contribute to the development of coronary artery complications in children with Kawasaki disease[56]. However, other results have

Author addresses

[1]Division of Paediatric Cardiology, Children's Hospital of San Antonio/Baylor College of Medicine, San Antonio, TX, USA.
[2]Division of Paediatric Cardiology, The University of Texas Southwestern Medical Center, Dallas, TX, USA.
[3]Service de Rhumatologie Pédiatrique, Centre de Référence des Maladies Auto-Inflammatoires Rares et des Amyloses, CHU de Bicêtre, le Kremlin Bicêtre, France.
[4]Division of Paediatrics, Cohen Children's Medical Center, New Hyde Park, NY, USA.
[5]Division of Paediatric Cardiology, Driscoll Children's Hospital, Driscoll, TX, USA.
[6]Division of Paediatric Cardiology, Cohen Children's Medical Center, New Hyde Park, NY, USA.
[7]Division of Paediatric Rheumatology, The Sydney Children's Hospitals Network, Sydney, NSW, Australia.
[8]Division of Neurology, John F Kennedy Medical Center/Hackensack Meridian University, Edison, NJ, USA.
[9]Division of Pathology and Laboratory Medicine, Northwell Health Laboratories, Lake Success, NY, USA.
[10]Institut National de la Santé et de la Recherche Médicale, Centre de Recherche des Cordeliers, Sorbonne Université, Université de Paris, Paris, France.
[11]Department of Biological Sciences & Engineering, Indian Institute of Technology Palakkad, Palakkad, India.
[12]These authors contributed equally: Chetan Sharma, Madhusudan Ganigara.

Table 1 | **Comparison of Kawasaki disease and MIS-C**

Comparison	Kawasaki disease	MIS-C
Demographics		
Age	6 months to 5 years	6–11 years
Sex	Male predominance (~1.5:1)	No apparent predominance
Race or ethnicity	Highest incidence in Japan, China, South Korea and Taiwan	Highest incidence in children of African and Hispanic heritage
Pathogenesis		
Trigger	Unknown but some data suggest possible preceding viral or bacterial infection	Onset ~3–6 weeks after SARS-CoV-2 exposure
Immunological characteristics		
Similarities	Enhancement of IL-1β$^+$ neutrophils and immature neutrophils	
Differences	T cell activation by a conventional antigen	SARS-CoV-2 viral spike (S) protein acts like a superantigen, triggering a cytokine storm
	High levels of IL-17	High levels of IL-15, IFNγ in severe cases
	Relatively less frequent MAS-like cytokine profile	>50% of patients with MIS-C have a MAS-like cytokine phenotype
	Lymphopenia is rare	Lymphopenia
	Anti-SARS-CoV-2 IgG not reported	Anti-SARS-CoV-2 IgG
Clinical features		
Similarities	Similar associations with fever, rash, cervical lymphadenopathy, neurological symptoms, extremity changes	
Differences	Relatively high incidence of conjunctival injection and oral mucous membrane changes	Relatively high incidence of gastrointestinal symptoms, myocarditis and shock, and coagulopathy
Management		
Common	IVIG, glucocorticoids, acetylsalicylic acid	IVIG, glucocorticoids, acetylsalicylic acid
Rare	Infliximab, ciclosporin and anakinra	Anakinra, tocilizumab

IVIG, intravenous immunoglobulin; MAS, macrophage activation syndrome; MIS-C; multisystem inflammatory syndrome in children; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

identified elevation of vitamin D levels during the acute phase of Kawasaki disease in children who subsequently developed coronary arterial lesions[57]. The contribution of other nutritional factors has also been suggested. For example, iron-deficiency anaemia is associated with development of coronary abnormalities in Kawasaki disease[58]. These varied results suggest the need for further investigation and research to elucidate the role of malnutrition in the pathogenesis of Kawasaki disease.

The role of microbiota

Disturbances in the normal microbiota (dysbiosis) have been proposed to have a role in the pathogenesis of various autoimmune and inflammatory disorders, including Kawasaki disease[59–61]. Stools from children with Kawasaki disease contain higher numbers of Gram-positive bacteria from the *Streptococcus*, *Staphylococcus*, *Eubacterium* and *Peptostreptococcus* genera, as well as Hsp60-producing Gram-negative bacteria, and lower numbers of lactobacilli than stools from children with other febrile illnesses or healthy controls[49,61,62]. Dysbiosis is associated with reduction in the production of short-chain fatty acids (particularly butyrate) and is proposed to lead to aberrant immune responses that are associated with Kawasaki disease[63].

Genetic susceptibility

Epidemiological and genetic studies of Kawasaki disease have shed light on the role of genetic susceptibility in its development[64]. Kawasaki disease is prevalent in Japan, but also in children of Japanese ancestry living in Hawaii[65]. Siblings of children with Kawasaki disease have a 10-fold higher risk of development of the condition than children in the general population[66]. Several candidate genes have been identified through genome-wide association studies and linkage studies. The four major groups of genes that have been studied in Kawasaki disease are those associated with T cell activation (*ORAI1* and *STIM1*), B cell signalling (*CD40*, *BLK* and *FCGR2A*), apoptosis (*CASP3*) and transforming growth factor-β (TGFβ) signalling (*TGFB2*, *TGFBR2*, *MMP* and *SMAD*)[64]. *CASP3* encodes caspase 3, which is an effector caspase with a vital role in the execution phase of apoptosis. A single-nucleotide polymorphism in the *CASP3* gene is associated with susceptibility to Kawasaki disease[67]. TGFβ is another vital protein with a central role in immunoregulation that affects multiple populations of leukocytes. Abnormalities in TGFβ signalling resulting from genetic variation are involved in Kawasaki disease susceptibility and outcomes[68]. Understanding the roles of these genetic alterations has implications for potential therapeutic approaches[69]. In addition to these groups, mutations in *ITPKC*, which is involved in Ca^{2+} mobilization and activation of NLRP3 inflammasomes, could result in enhancement of IL-1β and IL-18 production, disease susceptibility, coronary abnormalities and resistance to treatment with intravenous immunoglobulin (IVIG)[70,71]. Notably, immunosuppressive agents such as ciclosporin, a T cell inhibitor that blocks the calcineurin–NFAT pathway, have shown promise in the treatment of high-risk IVIG-resistant Kawasaki disease[72]. The observed association of HLA polymorphisms with Kawasaki disease varies; the predominant variant in a Japanese cohort was *HLA-Bw54*, whereas *HLA-Bw51* was predominantly identified in white and Jewish populations[73–75]. Epigenetic regulation of inflammatory and immunoregulatory genes by factors such as methylation, microRNAs and long non-coding RNAs has been identified in Kawasaki disease, and might be relevant to pathogenesis and prognosis[76]. Additionally, single-nucleotide polymorphisms in cytokine genes, including *IL1*, *KCNN2*, *TIFAB*, *P2RY12* and *TNF*, are associated with Kawasaki disease and with risk of coronary artery lesions, as well as IVIG treatment failure[77–82].

Immunological aberrations

Innate and adaptive immune responses both have important roles in the development of Kawasaki disease[83]. An intense initial response driven by the innate immune system takes the form of neutrophilic leukocytosis, activation of monocytes, natural killer (NK) cells and γδ T cells, and elevation of production of acute-phase

reactants and cytokines, especially IL-1β, which contributes to activation of endothelial cells, inducing upregulation of expression of cell-adhesion molecules, *IL6* and *IL8* (REFS[76–88]). Pro-inflammatory IL-17 produced by type 17 T helper (T_H17) cells could activate immune cells such as neutrophils and monocytes, leading to production of other inflammatory cytokines, such as IL-6, TNF and IL-8, thereby contributing to the pathogenesis of many inflammatory disorders[89–92]. By contrast, $CD4^+CD25^+$ regulatory T (T_{reg}) cells contribute to immune tolerance through suppression of the hyperactivation of both innate and adaptive immune cells, via several mutually nonexclusive mechanisms. An imbalance in these pathways could lead to immune dysregulation, which could have a role in the pathogenesis of Kawasaki disease[93–96].

Neutrophils are activated in Kawasaki disease and release reactive oxygen species, leading to endothelial cell injury[83]. Release of neutrophil extracellular traps (NETs) is also implicated in the pathogenesis of Kawasaki disease[97]. Although NETs have a protective role against infections as components of the innate immune system, they also have pathogenic potential for immune dysregulation and promotion of inflammation and tissue injury[98]. NETs have been implicated in the development and progression of rheumatic diseases, including systemic lupus erythematosus, rheumatoid arthritis and autoimmune vasculitis[99–101]. Yoshida et al.[97] demonstrated elevation of NET formation in the sera of patients with Kawasaki disease, as well as neutrophil infiltration in the lesions of vasculitis in the coronary arteries and aorta in a mouse model of Kawasaki disease.

Autoimmune antibodies are thought to have a role in the pathogenesis of Kawasaki disease, particularly those against endothelial cell antigens[102,103]. Anti-endothelial cell antibodies could cause endothelial damage, with release of pro-inflammatory cytokines and a hypercoagulable state leading to vessel-wall injury and intravascular thrombosis[104]. However, not all results have demonstrated elevation of anti-endothelial cell antibodies in patients with Kawasaki disease[105].

Immune complexes might have a role in the development of Kawasaki disease[106]. They appear in the first 7 days of the disease and peak in the second week before declining[107]. Elevation of circulating levels of immune complexes in Kawasaki disease is related to adverse outcomes such as coronary artery abnormalities[108,109]. However, a causal relationship between immune complexes and the pathogenesis of Kawasaki disease has not been definitively established. Activation of the complement system has also been implicated in the pathogenesis of Kawasaki disease, via both the classical and the mannose-binding lectin pathways[110,111].

Kawasaki disease is considered by some to be a form of IgA vasculitis. In children with Kawasaki disease, intestinal permeability and levels of secretory IgA in the circulation are greater than in unaffected children, and in mouse models of the disease, elevation of levels of circulating secretory IgA and IgA deposition in the vasculature are observed[112,113]. Furthermore, pharmacological blockade of zonulin (a modulator of intestinal tight junctions) and administration of IVIG in these mouse models reduce intestinal permeability and cardiovascular inflammation compared with levels in untreated controls[114].

Therapeutic strategies

IVIG and acetylsalicylic acid have emerged as first-line therapies for the management of Kawasaki disease[20]. IVIG therapy leads to rapid improvement in the clinical symptoms of rash, fever and conjunctival injection in most patients. Although the exact mechanism of action of IVIG is not yet known, proposals include inhibition of activation of innate immune cells and inflammatory mediators, expansion of T_{reg} cells and suppression of T_H17 cells[93,115–118]. Evidence indicates that IVIG might target IL-$1\beta^+$ neutrophils via caspase-independent pathways[119]. Single-cell RNA sequencing of peripheral blood mononuclear cells in acute Kawasaki disease before and after IVIG therapy has revealed that genes encoding inflammatory mediators (including *TNF* and *IL1B*) are highly expressed in monocytes in untreated disease, with reduction of expression following therapy, along with significant enhancement of the plasma-cell population and induction of oligoclonal expansion of T cell receptors and IgG and IgA B cell receptors[120]. Mining of transcriptomic data by Boolean analysis has identified that several metabolic pathways might contribute to IVIG resistance in Kawasaki disease[121]. In high-risk patients with acute Kawasaki disease and in those who do not respond to IVIG therapy, steroid treatment can be considered, to prevent the occurrence of coronary artery abnormalities[20]. Additional therapeutic options for IVIG-resistant Kawasaki disease include infliximab (a monoclonal antibody to TNF), ciclosporin (a calcineurin inhibitor) and anakinra (an IL-1 receptor antagonist)[20].

Overview of MIS-C

Since April 2020, many reports have documented a new hyperinflammatory syndrome in children[11,12,122]. In May 2020, the US Centers for Disease Control and Prevention (CDC) issued an alert identifying MIS-C as a critical illness in children that was associated with SARS-CoV-2 infection[123]. Since then, more than 4,000 cases of MIS-C have been reported in the USA alone[124]. In 26 studies published in 2020 and 2021, documenting 1,136 cases of MIS-C (mostly occurring in the USA and Europe), the reported median ages of the affected children were 6–11 years, with no significant gender difference[14,15,122,125–140] (TABLE 2).

Patients with MIS-C have symptoms that resemble those of other hyperinflammatory syndromes, such as Kawasaki disease, toxic-shock syndrome (TSS) and macrophage activation syndrome (MAS), which is a type of secondary haemophagocytic lymphohistiocytosis[11,12]. To improve clarity and aid diagnosis, the CDC published a case definition, which includes age <21 years, fever, laboratory evidence of inflammation, hospital admission, multisystem (two or more) organ involvement (cardiac, renal, respiratory, haematological, gastrointestinal, dermatological or neurological), either laboratory confirmation of SARS-CoV-2 infection (by PCR with reverse transcription (RT–PCR), serology or antigen test) or known COVID-19 exposure up to 4 weeks

Table 2 | **Demographic features of MIS-C study populations**

Study	Cohort location	*N*	Median age, years (range or IQR)	Male:female (%)	Race or ethnicity	Ref.
Dufort et al.	USA	99	No median Distribution: 0–5, 31%; 6–12, 42%; 13–20, 26%	54:46	Black, 40%; white, 37%; Hispanic, 36%; other, 18%; Asian, 5%	14
Cheung et al.	USA	17	8 (1.8–16)	47:53	Ashkenazi Jewish, 35%; Black, 24%; Hispanic, 24%; white non-Hispanic, 12%; Asian, 6%	15
Belhadjer et al.	France, Switzerland	35	10 (2–16)	51:49	Not reported	122
Kaushik et al.	USA	33	10 (IQR 6–13)	61:39	Hispanic, 45%; Black, 38%; white, 9%; Asian, 3%; other, 3%	125
Davies et al.	UK	78	11 (IQR 8–14)	67:33	Afro-Caribbean, 47%; Asian, 28%; white, 22%; other, 3%	126
Pouletty et al.	France	16	10 (IQR 4.7–12.5)	50:50	Not reported	127
Toubiana et al.	France	21	7.9 (3.7–16.6)	43:57	Sub-Saharan African/Caribbean parentage, 57%; European parentage, 29%; Asian parentage, 10%; Middle Eastern parentage, 5%	128
Capone et al.	USA	33	8.6 (IQR 4.4–12.6)	61:39	Other, 45%; Black, 24%; Asian, 9%; white, 9%; unknown, 12% (Hispanic, 27%; non-Hispanic, 73%)	129
Hameed et al.	UK	35	11 (IQR 6–14)	77:23	Not reported	130
Whittaker et al.	UK	58	9 (IQR 5.7–14)	56:44	Black, 38%; Asian, 31%; white, 21%; other, 10%	131
Moraleda et al.	Spain	31	7.6 (IQR 4.5–11.5)	58:42	Not reported	132
Dhanalakshmi et al.	India	19	6 (1.1–16.9)	42:58	Not reported	133
Miller et al.	USA	44	7.3 (0.7–20)	45:55	Hispanic, 34%; not reported, 25%; white, 20.5%; Black, 20.5%	134
Belot et al.	France	108	8 (IQR 5–11)	49:51	Not reported	135
Lee et al.	USA	28	9 (0.1–17)	57:43	Hispanic, 43%; white, 36%; Black, 18%; not reported, 3%	136
Riollano-Cruz et al.	USA	15	No median Mean 12 (3–20)	73:27	Hispanic, 66%; non-Hispanic African American, 13%; non-Hispanic white, 13%; other, 8%	137
Ramcharan et al.	UK	15	8.8 (IQR 6.4–11.2)	73:27	African or Afro-Caribbean, 40%; South Asian, 40%; mixed, 13%; other, 7%	138
Grimaud et al.	France	20	10 (2–16)	50:50	Not reported	139
Perez-Toledo et al.	UK	8	9 (7–14)	63:37	Not reported	140
Jonat et al.	USA	54	7 (0.7–20)	46:54	White, 35%; unknown, 31%; other, 19%; African American, 15%	205
Feldstein et al.	USA	186	8.3 (IQR 3.3–12.5)	65:35	Hispanic, 31%; Black, 25%; unknown, 22%; white, 19%; other, 5%	13
Toubiana et al.	France	23	8.2	52:48	Not reported	236
García-Salido et al.	Spain	61	9.4 (IQR 5.5–11.8)	66:34	Not reported	145
Shobhavat et al.	India	21	7 (IQR 1.9–12.1)	47:53	Not reported	261
Niño-Taravilla et al.	Chile	26	6.5 (IQR 2–10.5)	58:42	Chilean, 73%; Venezuelan, 12%; Peruvian, 8%; Colombian, 4%; Haitian, 4%	262
Tolunay et al.	Turkey	52	9 (IQR 5–13)	38:62	Turkish, 86%; Syrian, 14%	263

IQR, interquartile range; MIS-C, multisystem inflammatory syndrome in children.

before symptom onset, with no alternative plausible diagnosis[123]. The WHO and the UK Royal College of Paediatrics and Child Health (RCPCH) have also published case definitions, which are largely similar to the CDC definition, except that the RCPCH does not require evidence of prior exposure to SARS-CoV-2 (REFS[141,142]). Despite the broad case definitions, and the considerable overlap with primary COVID-19 and other common childhood febrile illnesses, patients with MIS-C have distinct clinical presentation and levels of biomarkers, which aids in differential diagnosis[143–145].

Aetiology

Compared with adults, primary SARS-CoV-2 infection is relatively mild in children[146]. Evidence indicates that a temporal relationship exists between SARS-CoV-2

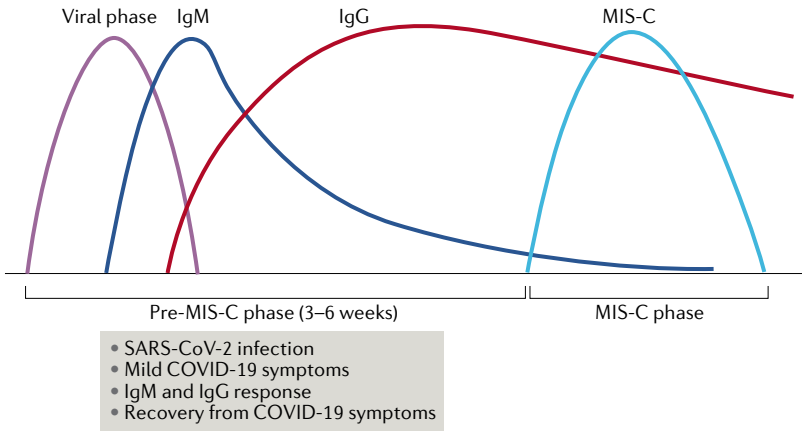


Fig. 1 | **The temporal relationship between SARS-CoV-2 infection and development of MIS-C.** Evidence suggests that a relationship exists between the timing of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection and development of multisystem inflammatory syndrome in children (MIS-C). Cases of MIS-C tend to be seen 3–6 weeks after the peak of SARS-CoV-2 transmission in a community. Because of this time lag, MIS-C is associated with a strong anti-spike protein IgG response, but a weak IgM response. It should be noted that implication of SARS-CoV-2 as a triggering factor for the development of MIS-C has yet to be firmly established.

exposure and development of MIS-C, as a spike in MIS-C cases occurs 3–6 weeks after the peak of SARS-CoV-2 transmission in a community[13,129,147] (FIG. 1). Median intervals of 21 and 25 days have been observed between the occurrence of COVID-19 symptoms and the onset of MIS-C[13,14]. Although 80–90% of patients with MIS-C have been found to be SARS-CoV-2 seropositive, positivity in PCR testing is only 20–40%, suggesting that the interval to the onset of MIS-C is sufficient for viral RNA levels to fall considerably[148,149]. Furthermore, nasopharyngeal aspirates from patients with MIS-C have higher SARS-CoV-2 real-time RT–PCR cycle thresholds (indicating lower levels of viral RNA) than those from patients with severe COVID-19 (REF.[150]). However, autopsy examinations for three individuals who had MIS-C identified SARS-CoV-2 in various tissues, including heart, kidneys, brain and intestine, which is consistent with multisystem organ involvement in MIS-C[151]. Notably, the prolonged presence of SARS-CoV-2 in children's intestines might cause zonulin-dependent loss of tight junctions, leading to leakage of viral antigens into the circulation, and to hyperinflammation and MIS-C[152]. By contrast, single-cell RNA sequencing of peripheral blood mononuclear cells from patients with acute MIS-C have revealed low viral and bacterial signatures in the immune cells, suggesting that active viral or bacterial infectious triggers are not contributing factors[153]. The accumulated evidence suggests that MIS-C might be the result of a combination of post-infectious immune dysregulation and virus-induced cytopathic effects and inflammation in multiple organ systems.

Paediatric patients with COVID-19 or MIS-C have strong IgG, but weak IgM antibody responses to the trimeric S glycoprotein of SARS-CoV-2, and weak responses to the nucleocapsid protein N, which is implicated in viral replication[140,154–158]. By contrast, adult COVID-19 patients have higher levels of anti-S antibodies, broader immunoglobulin response to SARS-CoV-2 with respect to specificity and isotype distribution (including IgG, IgM and IgA isotypes) and higher virus-neutralizing capacity[140,155,156,158]. The mild or asymptomatic nature of COVID-19 in children might be related to the extent of the antibody response. Nevertheless, IgG antibodies to S protein provide an important diagnostic criterion for MIS-C. Low IgM titres in MIS-C are consistent with its appearance several weeks after SARS-CoV-2 exposure.

Analyses from geographically diverse cohorts have demonstrated that 20–50% of people with no previous exposure to the virus have T cell reactivity against peptides corresponding to SARS-CoV-2 sequences[159], which might be related to $CD4^+$ T cell cross-reactivity with circulating seasonal human 'common cold' coronavirus (HCoV)[160]. Although this phenomenon has implications for the development of herd-immunity models and vaccine candidates, it is currently unclear whether the presence of prior cross-reactive $CD4^+$ T cells is protective or harmful in the pathogenesis of MIS-C. When tested for serological evidence of prior seasonal coronavirus infection, children with MIS-C and those hospitalized for non-COVID reasons had similar prevalence and levels of antibodies to HCoV[161]. Additionally, HCoV antibody levels did not correlate with the levels of SARS-CoV-2 antibodies, suggesting that prior HCoV infection neither provides protection nor worsens the course of paediatric SARS-CoV-2 infection or MIS-C.

SARS-CoV-2 S protein as a superantigen

SARS-CoV-2 viral S protein might behave like a superantigen, triggering a cytokine storm that results in the development of the TSS-like presentation of MIS-C[162] (FIG. 2). The S protein has a high-affinity motif for binding TCR, which is similar in structure to the staphylococcal enterotoxin B, a superantigen that mediates TSS by interacting with both TCR and MHC class II molecules. Computational modelling has shown that SARS-CoV-2 encodes a superantigen motif near the S1/S2 cleavage site, which interacts with both the TCR and CD28 (REF.[163]). TCR repertoire analysis of T cells in a small number of patients with MIS-C has identified skewing of TCR Vβ towards TRBV11-2 (Vβ21.3), which is associated with HLA class I alleles A02, B35 and C04 (REFS[163,164]). The CDR3-independent nature of TCR Vβ skewing suggested superantigen-mediated activation of T cells in MIS-C. Further evidence supports the enrichment of TRBV11-2 among T cells[153,165], although notably it has been observed in the absence of differential expression of a set of 'superantigen genes'[153]. Also, MIS-C is usually observed several weeks after primary SARS-CoV-2 exposure, in contrast to the acute illness and cytokine storm observed in TSS[166]. In most cases, SARS-CoV-2 is undetectable in patients with MIS-C during the acute phase of inflammation. Thus, the superantigenic property of SARS-CoV-2 S protein and its implication in MIS-C is not yet confirmed. As an RNA virus, SARS-CoV-2 undergoes constant mutation, and whether any particular variant of the virus contributes to MIS-C by triggering strong inflammatory signalling in the immune cells and endothelial cells of children with COVID-19 requires further exploration. Notably, the use of in silico techniques has demonstrated

that mutations in the binding region of SARS-CoV-2 S protein could influence the interaction with MHC class II molecules and TCR[163].

Involvement of nutritional disorders

Nutritional factors such as vitamin D deficiency might have a role in the development of MIS-C. Adults with vitamin D deficiency were noted to have a more severe form of COVID-19 with an increased risk of death than those without this deficiency[167]. Vitamin D supplementation has proved to be of some benefit in infections with other viruses, such as influenza A[168]. Whether a similar benefit could accrue in MIS-C has not been elucidated.

The role of microbiota

Another contributing factor in the development of MIS-C that warrants investigation is the role of gut and respiratory tract microbiota. Gastrointestinal microbes are important regulators of the gut immune system and inflammation, and influence the balance between T_H17 cells and T_{reg} cells[169]. Adult patients with COVID-19 display alteration of gut and upper respiratory microbiomes, and gut dysbiosis persists beyond the nasal clearance of SARS-CoV-2 (REFS[170–173]). Notably, faecal SARS-CoV-2 load is inversely correlated with the abundance of bacteria of the Bacteroidetes phylum, which suppress ACE2 in the mouse gut[173]. Preliminary data

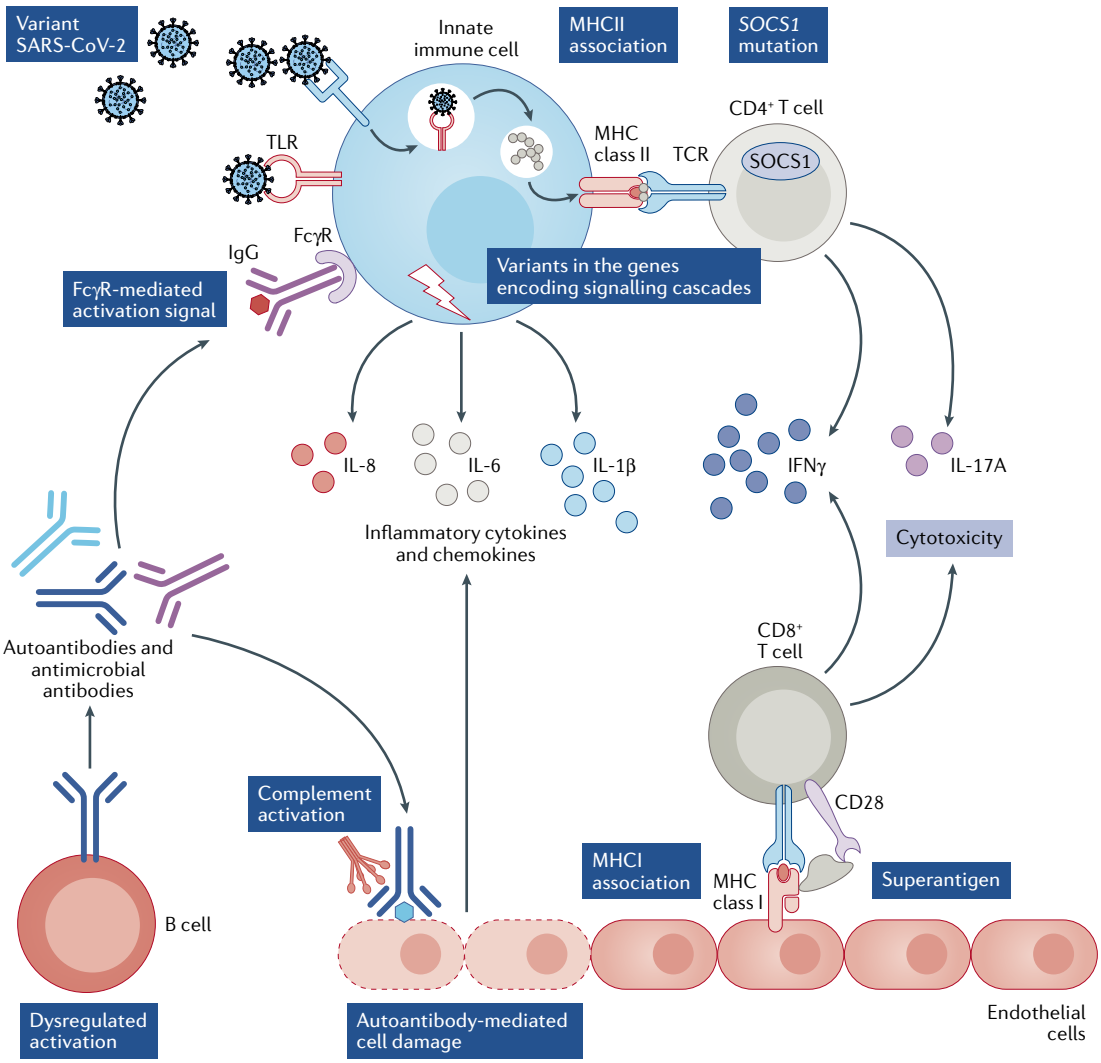


Fig. 2 | **Possible mechanisms implicated in aberrant activation of immune cells in MIS-C.** Clinical signs of multisystem inflammatory syndrome in children (MIS-C) mostly appear several weeks after severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection. MIS-C might be triggered by dysregulation of immune responses following viral infection. Aberrant activation of immune cells in patients with MIS-C could result from several factors. Infection with particular variants of SARS-CoV-2 might trigger hyperinflammatory responses. Genetic predisposition resulting from variants in the genes that encode pattern recognition receptors, Fcγ receptors and components of the signalling cascades of immune response, as well as mutations in genes such as *SOCS1*, which regulate inflammatory responses, could all contribute to enhancement of inflammatory responses to infection. Dysregulated activation of lymphocytes, with production of IgG corresponding to microbial pathogens or autoantigens, could cause immune-complex-mediated innate-cell activation by signalling via Fcγ receptors. Production of autoantibodies could also lead to complement activation and autoantibody-mediated endothelial damage. SARS-CoV-2 spike (S) protein might function as a superantigen, contributing to activation of T cells. SOCS1, suppressor of cytokine signalling 1; TLR, Toll-like receptor.

from peer reviewed and non-peer reviewed reports also suggest the persistence of microbiome dysbiosis in the upper respiratory tract and the gut in paediatric COVID-19 (REFS[174,175]).

Genetic susceptibility

The low incidence of MIS-C relative to COVID-19, and the similarity in antibody response to SARS-CoV-2 in paediatric patients with MIS-C and with COVID-19 (irrespective of the subsequent development of MIS-C) suggest that SARS-CoV-2 infection causes dysregulation of immune responses in a subgroup of predisposed children with particular genetic backgrounds[176]. Specifically, predisposition might be related to mutations and polymorphisms in the genes that encode pattern recognition molecules such as Toll-like receptors, components of the signalling cascades of the immune response and Fcγ receptors (FIG. 2). The incidence of MIS-C is higher in children of African and Hispanic heritage than in those of other ethnicities, although attribution of this finding to genetic differences is confounded by the contribution of socioeconomic factors to the risks of SARS-CoV-2 infection[177–180]. Among 145 HLA-A, HLA-B and HLA-C genotypes, *HLA-B*46:01* was associated with in silico prediction of the fewest SARS-CoV-2-binding peptides (suggesting particular vulnerability to COVID-19), whereas *HLA-B*15:03* was predicted to have the greatest capacity for coronavirus peptide presentation (suggesting protective T cell-based immunity)[181]. Monogenic loss-of-function variants affecting immunity in the type I interferon signalling pathway might confer a predisposition to severe COVID-19 manifestations[182,183]. In a study of two unrelated patients with infection-associated immune thrombocytopenia and autoimmune haemolytic anaemia, both had *SOCS1* haploinsufficiency and exhibited T cell activation and high levels of interferon signalling, and one developed MIS-C after SARS-CoV-2 infection. SOCS are negative regulators of interferon signalling, and silencing mutations might predispose the individuals to infection-associated hyperinflammatory states such as MIS-C[184] (FIG. 2). However, a clear genetic basis that explains why some children develop MIS-C after SARS-CoV-2 exposure is currently undetermined. Additional factors, such as epigenetic effects at the level of histones, DNA or microRNA might also contribute to the development of MIS-C.

Immunological aberrations

In general, the signatures of immune cells and inflammatory parameters of MIS-C closely overlap with those of adults with moderate-to-severe COVID-19 rather than with paediatric COVID-19, which is mostly mild or asymptomatic. Also, immune activation in MIS-C is transient and tends to reduce during recovery[154,185,186].

Pro-inflammatory mediators. Elevation of levels of pro-inflammatory cytokines such as IL-6, IL-10 and IL-17A, and chemokines such as CXCL5, CXCL11, CXCL1 and CXCL6 in MIS-C distinguishes it from paediatric COVID-19 (REFS[136,153,154,187]). In various cohorts, elevation of TNF, IL-1β, IFNγ, soluble IL-2R, CCL2, CCL3, CCL4, CXCL8 (IL-8) or IFNγ-induced chemokines CXCL9 and CXCL10 has been reported in the serum of patients with MIS-C relative to those with paediatric COVID-19 or healthy controls[136,150,153,165,186–189]. Overall, enhancement of these pro-inflammatory molecules in the circulation indicates inflammatory responses of myeloid and lymphoid cells. Endothelial cells could also contribute innate inflammatory mediators, as E-selectin, a marker of inflamed endothelial cells, shows elevation in the serum of patients with MIS-C[153]. The reasons for the absence of some pro-inflammatory mediators in particular cohorts of patients are not known. The mediators that were analysed could have differed from study to study, but also, the levels of inflammatory mediators might vary depending on the patients' genetic and epigenetic backgrounds, severity of the disease, geographical location and timing of the analyses. Results from a study of plasma proteomics in children with SARS-CoV-2 infection, which have not yet undergone peer review, suggest that IFNγ expression is heterogeneous among patients with MIS-C, and that patients have dysregulated response to IFNγ[190]. As the pandemic progresses, it will be important to have a consensus regarding the panel of cytokines and chemokines that should be analysed in relation to MIS-C, to facilitate our understanding of the molecular pathogenesis and heterogeneity of this complex disease, and to enable accurate prognosis and effective treatment.

Immune-cell profiles. Immune-cell profiling of children with MIS-C or primary COVID-19 infection reveals similarities as well as differences in their immune signatures[185]. They have similar proportions of eosinophils, immature granulocytes, monocytes and classic dendritic cells, but patients with MIS-C have elevation of neutrophils and reduction of plasmacytoid dendritic cells[153,185,188], which might contribute to the low levels of IFNα that are observed in the blood of patients with MIS-C relative to those with paediatric COVID-19 (REF.[188]).

Neutrophils and monocytes are activated in patients with MIS-C[186] and show upregulation of alarmin signatures (in particular S100A genes) and reduction of expression of antigen-presenting, antigen-processing and co-stimulatory molecules[153,165,186]. Compared with healthy children, those with MIS-C have greater expression of cytotoxicity genes and *CCL4* in NK cells, which might contribute to the occurrence of tissue damage[153]. Preliminary results suggest that plasma levels of IFNγ correlate with levels of NCR1 and IL-2RA, which are the soluble markers of activated NK and T cells, respectively[190].

MIS-C could have a common pathophysiology with Kawasaki disease involving NET formation, which has been described in the sera of adults with COVID-19 and with endothelial injuries or a prothrombotic state[191–193]. However, plasma levels of NETs and release of NETs from neutrophils are similar in children with mild or moderate COVID-19 or MIS-C and in healthy children[194]. Despite similarities between disorders associated with pathogenic NETs and MIS-C, the role of NETosis in the pathogenesis of MIS-C remains uncertain because of a lack of definitive evidence, and hence further studies are warranted.

Both MIS-C and paediatric COVID-19 present with general lymphopenia (affecting cells that include mucosa-associated invariant T cells, γδ T lymphocytes and $CD8^+$ T lymphocytes)[136,150,154,185–188]. Compared with paediatric COVID-19, in MIS-C there is more-pronounced $CD4^+$ T cell-biased lymphopenia, which is similar to the situation in severely ill adults with COVID-19 (REF.[185]). However, results from single-cell RNA sequencing analysis have revealed enhanced proliferation of $CD4^+$ T cells in patients with MIS-C compared with healthy individuals[153], suggesting that lymphopenia might be the result of homing of T cells to the inflamed tissues. Despite showing T cell lymphopenia, the relative distribution of various T cell subsets such as naive, central memory and effector memory cells in patients with MIS-C is similar to that in age-matched healthy individuals, indicating a pan-$CD4^+$/$CD8^+$ T cell lymphopenia, rather than a subset-specific effect[154,185].

A distinct feature of MIS-C compared with paediatric COVID-19 is the activation of $CX_3CR1^+CD8^+$ T cells ($CD8^+$ T cells that express vascular endothelium-homing CX_3CR1, also known as fractalkine receptor), which could have implications for development of vascular abnormalities and cardiovascular abnormalities[185]. This immunological phenotype is correlated with elevation of D-dimer, reduction of platelets and with the requirement for vasoactive medication. Although not as prominent as in NK cells, $CD8^+$ T cells in MIS-C also show increases in signatures of cytotoxicity compared with those in healthy children[153]. RNA sequencing in blood from children with MIS-C revealed aberrant NK and $CD8^+$ T cell regulation, with depletion of NK cells and an absence of NK cell-dependent exhaustion of effector $CD8^+$ T cells, which can lead to sustained inflammation[195]. Notably, the proportion of activated $CX_3CR1^+CD8^+$ T cells in patients with MIS-C decreases as the clinical status improves[185]. Thus, there seems to be a sustained activation and dysregulation of $CD8^+$ T cells, particularly those that express CX_3CR1.

Nonspecific activation of B cell clones and expansion of plasmablasts occurs in MIS-C[150,153,185,186]. Plasmablast elevation also occurs in children with COVID-19 (REF.[185]). However, the specificity of expanded B cells and plasmablasts might vary between the two conditions. Patients with MIS-C display pronounced autoreactivity signatures of plasma immunoglobulins compared with healthy children or adults and children with COVID-19 (REFS[103,153,154]). Also, patients with MIS-C have evidence of extrafollicular responses, as indicated by high frequencies of plasmablasts expressing the T box transcription factor T-bet[185]. Future research should aim to uncover the reasons for this B cell activation, and should compare the characteristics of expanded B cells and plasmablasts, and the specificities of immunoglobulins, in MIS-C and paediatric COVID-19. As both conditions are associated with nonspecific B cell activation and elevation of plasmablast frequencies[196], molecular mimicry between self-antigens and SARS-CoV-2 antigens (as described in a paper that has not yet been peer reviewed[197]) might not be entirely responsible for the appearance of autoreactivity in MIS-C, and instead a combination of molecular mimicry and dysfunctional immunoregulatory machinery could be involved.

Humoral features. Analysis of IgG by systems serology has identified that humoral features in patients with MIS-C, such as complement deposition and neutrophil phagocytosis, overlap with those in convalescent adults with COVID-19 (REF.[158]). However, patients with severe MIS-C have persistent levels of FcγR binding (and in particular activating FcγRIIA) and inflammatory monocyte/macrophage-activating IgG[158]. Although hypergammaglobulinaemia is not observed in patients with MIS-C, a selective expansion of the IgG repertoire to react not only to SARS-CoV-2, but also to other bacterial and viral pathogens, some of which are implicated in the triggering of Kawasaki disease, has been observed. The underlying reason for the enrichment of particular IgG specificities is not yet known, but many of the microbes have been identified in the respiratory tracts of patients with MIS-C[198], suggesting a role for an immune-complex-driven inflammatory response in the pathogenesis of MIS-C. IgG and IgA autoantibodies occur in patients with MIS-C, and recognize gastrointestinal, mucosal, immune-cell and endothelial antigens[153,154]. Although the functionality of these autoantibodies and their roles in the pathogenesis of MIS-C should be investigated, these results might explain at least in part the involvement of multiple organ systems in MIS-C and provide a pointer towards dysregulated activation of B lymphocytes, enhanced autoreactivity and immune-complex-mediated inflammatory responses (FIG. 2). Enhanced expression of CD64 (FcγR1), a high-affinity receptor for the Fc fragment of IgG, has been observed on neutrophils and monocytes of patients with MIS-C[154,186]. Furthermore, most of these patients respond to IVIG therapy[11,15,122,129,131], which provides additional indirect support for the implication of FcγR-mediated activation of innate immune cells by immune complexes formed by these IgGs.

A role for the complement system in the pathogenesis of MIS-C has been suggested. Patients with MIS-C or paediatric COVID-19 have elevated plasma levels of soluble C5b-9 compared with healthy controls[150,199]. Soluble C5b-9 is a biomarker to monitor the activity of the terminal pathway of complement, and elevated levels suggest complement activation and endothelial dysfunction. Notably, although patients with MIS-C and paediatric COVID-19 have similar levels of complement-activating IgG antibodies to S protein of SARS-CoV-2 (REFS[140,154–158]), those with MIS-C have enhanced autoreactive signatures of IgG[103,153,154]. As patients with MIS-C typically have minimal or no SARS-CoV-2 at the time of development of the disease, enhanced autoreactivity and immune-complex formation might contribute to the elevated levels of C5b-9. Consistent with complement activation, MIS-C is associated with clinical criteria for complement-mediated thrombotic microangiopathy, such as microangiopathic haemolytic anaemia, hypertension, thrombocytopenia, proteinuria and evidence of organ damage on the basis of lactate dehydrogenase elevation[199]. Compared with paediatric COVID-19, patients with MIS-C have higher incidence of thrombotic events[200]. Results from proteomics analyses of plasma samples, which have not yet been peer reviewed, suggest that phospholipase A2 (PLA2G2A) could be a biomarker for diagnosis of thrombotic

microangiopathy in MIS-C[190]. The lectin complement pathway might also have an important role in the pathogenesis of diseases associated with SARS-CoV-2, as a result of the carbohydrate-residue-rich surface structures of the virus[201–203].

Therapeutic strategies

Treatment approaches to MIS-C aim to mute the exaggerated inflammatory response. Multiple approaches, borrowed from Kawasaki disease and other hyperinflammatory syndromes, have been considered, ranging from IVIG to glucocorticoids and immunotherapy[204,205]. MIS-C treatment regimens described in 24 studies, involving 1,020 individuals, are summarized in TABLE 3, highlighting the many variations on the theme of attempting to calm overactive inflammatory responses[6,17,20,94,100–114,164]. In most studies, most (70–100%) of the patients were treated with IVIG as the first-line agent, with satisfactory results. Steroids were the second most common treatment employed for patients with MIS-C.

Shock and cardiovascular manifestations comprise a predominant mode of presentation of MIS-C, and high-dose glucocorticoids have been advocated for, and used successfully in, patients with shock. Widely followed guidance from the ACR recommends IVIG as first-line therapy in hospitalized patients with MIS-C, with addition of glucocorticoids in the presence of shock, organ-threatening disease or refractory disease[149]. In a study of 181 children with suspected MIS-C, IVIG alone had a higher failure rate than the use of IVIG with methylprednisolone (OR 0.25; 95% CI 0.09–0.70)[206]. By contrast, results from a multinational observational cohort study that involved 615 children with suspected MIS-C identified no difference in acute outcomes between primary treatment with IVIG alone, IVIG with steroids or steroids alone[207]. In view of the apparently important role of IL-1β in the pathogenesis of MIS-C, anakinra (an IL-1 receptor antagonist) has been used in MIS-C that is refractory to therapy with IVIG or steroids, extrapolating from its success in small groups of patients with IVIG-resistant Kawasaki disease[122,148,208,209].

Zonulin-dependent loss of intestinal mucosal permeability is implicated in mediation of the hyperinflammation observed in MIS-C, and accordingly, a patient who did not respond to anti-inflammatory therapies

Table 3 | **Treatment of MIS-C**

Study	Cohort location	*N*	IVIG (%)	Glucocorticoids (%)	Other treatments	Ref.
Dufort et al.	USA	99	70	64	NR	14
Cheung et al.	USA	17	77	82	Tocilizumab, 6%	15
Belhadjer et al.	France, Switzerland	35	72	34	Anakinra, 9%	122
Kaushik et al.	USA	33	54	51	Tocilizumab, 36%; remdesivir, 21%; anakinra, 12%; convalescent plasma therapy, 3%	125
Davies et al.	UK	78	76	73	Tocilizumab, 4%; anakinra, 10%; infliximab, 9%; rituximab, 1%	126
Pouletty et al.	France	16	94	18.8	Tocilizumab, 6%; anakinra, 6%; hydroxychloroquine, 6%	127
Toubiana et al.	France	21	100	33	NR	128
Capone et al.	USA	33	100	70	Tocilizumab, 9%; anakinra, 12%; infliximab, 3%	129
Hameed et al.	UK	35	100	100	NR	130
Whittaker et al.	UK	58	71	64	Anakinra, 5%; infliximab, 14%	131
Moraleda et al.	Spain	31	65	68	Remdesivir, 6%	132
Dhanalakshmi et al.	India	19	79	58	Tocilizumab, 5%	133
Miller et al.	USA	44	82	96	Anakinra, 18%	134
Lee et al.	USA	28	71	61	Anakinra, 18%	136
Riollano-Cruz et al.	USA	15	80	20	Tocilizumab, 80%; remdesivir, 13%; anakinra, 13%; convalescent plasma therapy, 6%	137
Ramcharan et al.	UK	15	66	33	NR	138
Grimaud et al.	France	20	100	10	Tocilizumab, 10%; anakinra, 10%	139
Jonat et al.	USA	54	83	79	NR	205
Feldstein et al.	USA	186	77	49	Anakinra, 13%	13
Toubiana et al.	France	23	100	61	NR	236
García-Salido et al.	Spain	61	45	80	Tocilizumab, 24%; hydroxychloroquine, 55%	145
Shobhavat et al.	India	21	52	86	Tocilizumab, 10%	261
Niño-Taravilla et al.	Chile	26	77	88	Tocilizumab, 12%; infliximab, 4%	262
Tolunay et al.	Turkey	52	93	71	Anakinra, 4%	263

IVIG, intravenous immunoglobulin; MIS-C, multisystem inflammatory syndrome in children; NR, not reported.

was treated with the zonulin antagonist larazotide, with a satisfactory outcome[152]. In a pooled meta-analysis, D-dimer was found to be elevated in 92% of patients (330 out of 356)[210]. Because of the associated risk of hypercoagulability, and extrapolating from the management of Kawasaki disease, the use of anticoagulants such as acetylsalicylic acid and/or enoxaparin has been reported[211,212].

MIS-C: distinct from Kawasaki disease?

Both Kawasaki disease and MIS-C have temporal associations with infectious diseases and are associated with immune-system alteration, systemic inflammation and cytokine storm. Myocardial dysfunction, which is seen in both pathologies, might be a consequence of systemic inflammation[213,214]. An artificial intelligence computational analysis based on viral pandemics and disease-severity gene signatures, and in particular induction of *IL15–IL15RA* genes, has placed Kawasaki disease and MIS-C on the same host-immune-response continuum (although these results have not yet been peer reviewed)[215]. Consistently, patients with MIS-C have significantly higher levels of IL-15 than paediatric patients with COVID-19 (REF.[189]). However, the intensity of the immune response is high in MIS-C, which places it further along the severity spectrum than Kawasaki disease[215].

A quarter to half of patients with MIS-C meet the full criteria for diagnosis of Kawasaki disease[17,148,149,212,216]. Without evidence of prior SARS-CoV-2 exposure in these patients, it might not be possible to differentiate them from those with classic Kawasaki disease. Commonly reported clinical features of MIS-C include fever, mucocutaneous findings, myocardial dysfunction with cardiogenic or vasoplegic shock, gastrointestinal symptoms and neurological features including headache and altered mental status (TABLE 4). Like Kawasaki disease, these clinical manifestations are not specific to MIS-C, and they could occur in other infectious or inflammatory conditions[20].

Epidemiological and clinical differences

Despite the apparent similarities between MIS-C and Kawasaki disease, there are important epidemiological and clinical differences[20,122,127]. Kawasaki disease is typically a disease of young children <5 years old, whereas MIS-C has been reported in a wide age range from 1.6 to 20 years, with a median age of 6–11 years[20,217,218] (TABLE 2). In sharp contrast to Kawasaki disease, there is a surprising lack of reports of MIS-C from Japan and East Asian countries[219,220]. In fact, published data from the USA and Europe suggest that MIS-C is most commonly encountered in children of African and Hispanic heritage[177,180]. These epidemiological differences suggest that although MIS-C has phenotypic similarities to Kawasaki disease, they are essentially distinct syndromes.

Cardiac involvement is more prevalent and severe in MIS-C than in Kawasaki disease. Although a quarter of untreated patients with Kawasaki disease will develop coronary artery abnormalities, in the current era with a high level of clinical suspicion as well as early diagnosis and treatment, the incidence of coronary artery abnormalities in Kawasaki disease is <10%[216,221–224]. By contrast, our understanding of coronary artery dilation in MIS-C is still evolving, and incidence rates of 14–48% have been reported in various patient populations[180,191,192] (FIG. 3). However, the adoption of standardized MIS-C management protocols has begun to reduce the rate of coronary artery involvement[149]. Cardiac MRI in MIS-C has demonstrated high signal intensity on T1-weighted and T2-weighted imaging, consistent with diffuse myocardial oedema, with no enhancement on late gadolinium imaging to suggest fibrosis[225]. Results from echocardiographic studies have demonstrated that global left ventricular longitudinal strain is significantly lower in individuals with MIS-C than in those with Kawasaki disease[226]. A longitudinal, single-centre study involving 15 children with MIS-C has demonstrated significant improvement towards normalization of both ventricular function and coronary artery size over a 30-day follow-up period[227].

Fewer than 10% of cases of Kawasaki disease manifest as Kawasaki disease-shock syndrome (KDSS), which requires the use of intravascular fluid resuscitation and vasoactive medication[228–231]. Patients with KDSS tend to be older, have longer duration of fever and higher levels of inflammatory markers, and have a higher incidence of IVIG resistance as well as coronary abnormalities than those without KDSS[232,233]. By contrast, shock and depressed left ventricular systolic function are more frequent with MIS-C, for which reports indicate that 40–80% of patients present with shock[210,234–236] (FIG. 3). In a retrospective comparison of a cohort of patients with KDSS with published data relating to MIS-C, individuals with KDSS were more likely than those with MIS-C to fulfil the diagnostic criteria for complete Kawasaki disease, with higher incidence of coronary artery aneurysms[237].

Kawasaki disease has been reported to occur with MAS[238,239]. In a retrospective analysis of 638 patients with Kawasaki disease, the incidence of MAS was <2%[240]. However, this figure is likely to be an underestimation of the true incidence, as a result of an absence of sensitive diagnostic criteria and a lack of awareness among health-care providers[241]. Patients with Kawasaki disease and MAS tend to have elevation of levels of IFNγ, TNF, serum neopterin, IL-18 and sTNFR-II[242]. A retrospective comparison of patients with MAS (as a complication of systemic-onset juvenile idiopathic arthritis) and MIS-C revealed that MAS was associated with lower levels of haemoglobin and fibrinogen, and higher ferritin and lactate dehydrogenase, whereas patients with MIS-C tended to have signs of shock and need of intensive care management[243]. Results that have not yet been peer reviewed, based on analyses of IFNγ and CXCL9 signalling characteristics, suggest that >50% of patients with MIS-C have a MAS-like cytokine phenotype[190], along with elevation of CD163, IL-2RA and ferritin (during the early period) in the plasma. However, although MAS has an association with neutropenia, patients with MIS-C, including those who meet the criteria for MAS, display neutrophilia. Thus, although KDSS and Kawasaki disease with MAS have overlapping clinical features with MIS-C, there are subtle differences that are likely to reflect the different cytokine profiles in these conditions.

Gastrointestinal and neurological symptoms are also more commonly encountered in MIS-C than in Kawasaki disease[134,235,236,244–247]. The gastrointestinal manifestations include abdominal pain, vomiting and diarrhoea[210,235,246], with rare presentations that resemble appendicitis requiring surgical exploration[248]. In a national US registry consisting of 1,695 children and adolescents with active COVID-19 infections including

Table 4 | **Reported clinical features of multisystem inflammatory syndrome in children**

Affected organ system	Symptoms	Frequency of involvement (%)	Refs
Cardiovascular	Shock	40–80	210,234–236
	Cardiac arrhythmias	2	235
	Abnormal ST- or T-wave segment	22	235
	Prolonged QT interval	2	235
	Pericardial effusion	13–28	235,246
	Decreased LVEF by echo	31–58	234,235
	Increased troponin	68–95	234,235
	Myocarditis	36–87	210,236,246
	Coronary artery dilation on CT	27	235
	Coronary artery aneurysm	14–48	234–236
	Mild	22	235
	Moderate	7	235
	Giant	1	235
Gastrointestinal	Gastrointestinal symptoms	60–100	234–236
	Diarrhoea	38–72	210,235
	Vomiting	51–68	210,246
	Abdominal pain	19–71	210,235
	Ascites	21	235
	Ileitis	9	235
	Colitis	4	235
Ophthalmological	Conjunctivitis	32–83	234–236,246,264
	Periorbital erythema and oedema	20	264
Nervous system	Neurological symptoms	13–35	210,235,236,249
	Severe symptoms, including encephalopathy, stroke, central nervous system infection/demyelination, Guillain–Barré syndrome and acute cerebral oedema	3	249
Integumentary	Rash	50–70	234,236,246
	Erythematous skin rash	62	235
	Hyperaemia, oedema or desquamation of extremities	26–51	235,264
	Malar erythema	17	264
	Skin eruptions	9–14	264
	Desquamation in groin	26	236
Respiratory	Upper respiratory tract infection	34	235
	Lower respiratory tract infection	22	235
	Pleural effusion on CT	20	235
	Lung involvement on CT (bilateral pulmonary consolidation and ground-glass opacity)	13	235
Mucosal	Oral mucosa hyperaemia	41	235
	Red and/or cracked lips	37–49	246,264
	Strawberry tongue	11–23	246,264
	Lips and oral-cavity changes	74	234,236
Other	Lymphadenopathy (cervical)	19–61	235,236,246
	Extremity changes	8–52	234,236,246

LVEF, left ventricular ejection fraction.

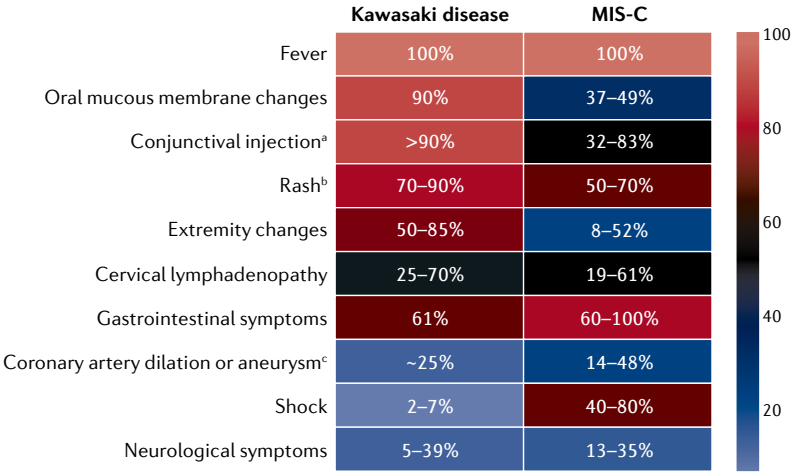


Fig. 3 | **Comparative incidence of clinical signs in MIS-C and Kawasaki disease.** Percentage incidence of particular symptoms in patients with multisystem inflammatory syndrome in children (MIS-C) or Kawasaki disease is shown, with the values derived from published reports[210,231,232,234–236,246,249,250,264–270]. Although some clinical signs, such as fever and cervical lymphadenopathy are equally prevalent in both MIS-C and Kawasaki disease, the incidence of other symptoms, including shock, coronary artery involvement and gastrointestinal symptoms (vomiting, diarrhoea or abdominal pain), are characteristic of MIS-C. [a]'Conjunctival injection' refers to bilateral non-exudative conjunctivitis in Kawasaki disease. [b]'Rash' refers to polymorphous rash in Kawasaki disease. [c]'Coronary artery dilation of aneurysm' refers to incidence in untreated cases of Kawasaki disease.

MIS-C (n = 616), neurological symptoms were noted in 22% of the patient population (n = 365), with most of those affected having transient symptoms[249]. Among these 365 patients, 126 met the criteria for MIS-C. In the patients with neurological involvement (n = 365), 43 (12%) had life-threatening neurological involvement (including encephalopathy, stroke, central nervous system infection and/or demyelination, Guillain–Barré syndrome and acute cerebral oedema) and among which 20 (47%) met the criteria for MIS-C. In a study of 286 children with MIS-C located in 55 centres in 17 European countries, neurological involvement was identified in 43 individuals (15%)[235]. In a pooled meta-analysis of data from 370 children with MIS-C, 133 (35.9%) had neurological symptoms[210] (TABLE 4). Neurological involvement in Kawasaki disease is variable, reportedly affecting 5–39% of patients[250,251].

In contrast to those with Kawasaki disease, patients with MIS-C tend to have a worse acute clinical course and multisystem involvement, as illustrated by an increased requirement forintensive care management. A large study of >1,000 patients with Kawasaki disease revealed that 2.4% of these children required intensive care[252]. In stark contrast, an analysis of 783 cases of MIS-C determined that 68% of patients required intensive care admission, 63% needed inotropic support, 28% had some form of respiratory support and 4% of patients required extra-corporeal membrane oxygenation[253]. Among 1,080 patients with MIS-C, intensive care admission was more likely in children aged >5 years old than in younger children, and in non-Hispanic Black patients than in non-Hispanic white patients, and coronary artery abnormalities were more common in boys than in girls[254]. Elevated acute-phase inflammatory markers, troponin, B-type natriuretic peptide and D-dimer levels also identified patents at risk of severe disease[254]. In a study evaluating 29 children with MIS-C in France, severe disease occurred in 52% of them and was associated with high persistent fever and high levels of inflammatory markers[255]. Although the reason for a more critical illness in the acute phase of MIS-C than in Kawasaki disease is unclear, it is thought to be linked to the cytokine storm in MIS-C[256,257].

Immunological differences

In several small studies, comparative immune profiling of children with MIS-C and Kawasaki disease has been performed, to differentiate between these two disease entities. MIS-C is associated with lymphopenia, lower white blood cell and naive CD4$^+$ T cell counts, and increased central and effector memory T cell subpopulations, compared with Kawasaki disease[103]. IL-17 is a mediator of inflammation in Kawasaki disease, but is less prominent in MIS-C[103]. In a comparison of cytokine profiles, levels of circulating IFNγ were significantly higher in patients with severe forms of MIS-C than in those with milder MIS-C or Kawasaki disease[258].

In a study of the immunological profiles of paediatric patients, 75% of those with MIS-C, but none with Kawasaki disease, TSS or COVID-19, displayed non-HLA-biased, SARS-CoV-2 non-reactive, polyclonal expansion of TCR Vβ 21.3$^+$ activated CD4$^+$ and CD8$^+$ T cells[165]. Notably, these Vβ 21.3$^+$ T cells had high expression of CX_3CR1, a marker previously identified on the activated CD8$^+$ T cells of patients with MIS-C[185]. The remarkable specificity of Vβ 21.3$^+$ T cell subset expansion noted in MIS-C is consistent with superantigen-mediated activation of the immune system[163], whereas in Kawasaki disease, evidence of a role of superantigens in pathogenesis is lacking[50].

Autoantibody profiles have been compared in patients with MIS-C and Kawasaki disease[103]. Levels of antibodies to some vascular endothelial cell proteins, such as endoglin, were higher in both groups of patients than in healthy controls, whereas some autoantibodies (such as that to EGF-like repeat and discoidin I-like domain-containing protein 3) were overexpressed in Kawasaki disease compared with MIS-C. To confound matters, plasma levels of endoglin were elevated in both sets of patients compared with healthy children, raising the possibility that antibodies to endothelial cells were the result, rather than the cause, of vascular damage. Another possibility is that the S protein superantigen of SARS-CoV-2 might cause aberrant activation of B cells[162].

Some laboratory parameters are important differentiators between Kawasaki disease and MIS-C. Although both syndromes involve a diffuse hyperinflammatory response, patients with MIS-C tend to have a lower platelet count, lower absolute lymphocyte count and higher levels of C-reactive protein, N-terminal pro-B-type natriuretic peptide, troponin and ferritin[16,17,20,94,100–114,164]. Additionally, coagulation abnormalities are common, including elevation of D-dimer and fibrinogen levels[11,12,257,259]. Hyponatraemia is another common laboratory finding in patients with MIS-C[127]. The presence

of burr cells and neutrophils with toxic granulation can also discriminate MIS-C from severe COVID-19 (REF.[150]). Finally, evidence of recent SARS-CoV-2 infection, particularly by positive serology, is a diagnostic indicator of MIS-C[148,149].

Conclusion

Epidemiological and clinical differences reveal that although MIS-C has phenotypic similarities to Kawasaki disease, they are different syndromes. They have varying degrees of hyperinflammation and dysregulated immune responses[131]. Children with MIS-C are, in general, more critically ill, with prominent gastrointestinal symptoms, cardiac involvement with shock, haematological abnormalities and elevated acute-phase reactants. They have positive SARS-CoV-2 serology, suggesting a link to prior clinical or subclinical infection or exposure. It is likely that a combination of pathogen and host factors is involved in the genesis of an intense aberrant activation of both innate and adaptive immune responses and subsequent cytokine storm[16,18]. Some of the pathogen-related factors include antigen mimicry of host antigens, and superantigen properties of viral proteins. Potential host factors include age and immune-system immaturity, altered intestinal microbiota, nutritional deficiencies and genetic (including inborn errors of immunity) and epigenetic predisposition[13]. However, to what extent each of these factors contributes, and how they interact to cause the clinical syndrome, are relative unknowns that need further exploration. Various lines of evidence based on inflammatory parameters, clinical signs or gene analyses have evoked the possibility that MIS-C is a heterogeneous complex disorder. Current data on the immune signatures in patients with MIS-C are based on small sample size, non-homogeneous cohorts. Hence, analysis of dynamic changes in the immune signatures of patients with MIS-C and their comparison with those with Kawasaki disease in a large homogeneous cohort is needed, to accurately determine the similarities and distinct features of the two disease entities. Nevertheless, with the evidence of elevation of signatures of autoimmunity in MIS-C, and reports of various post-COVID-19 conditions in adults, long-term follow-up of patients with MIS-C might be advisable, because of the possibility of relapse. Notably, however, relapse is rare in Kawasaki disease, which might suggest that recurrence is also unlikely in MIS-C.

While researchers and clinicians navigate the possibilities and evaluate the best treatment options for patients affected by COVID-19-related illnesses, it is imperative to establish registries and dedicated multidisciplinary research teams to investigate the pathogenesis and specific therapeutic strategies in MIS-C. An important step towards this end was the workshop that was convened by the NIH in June 2020, which aimed to bring together the experts on the subject, to initiate dialogue leading to future studies[260]. In conclusion, MIS-C has important epidemiological, clinical and immunological differences from Kawasaki disease, enabling its classification as a separate syndrome. Study of MIS-C will continue to enhance our understanding of these conditions that are related by their association with the cytokine storm phenomenon.

Published online 29 October 2021

1. Lai, C.-C. et al. Global epidemiology of coronavirus disease 2019 (COVID-19): disease incidence, daily cumulative index, mortality, and their association with country healthcare resources and economic status. *Int. J. Antimicrob. Agents* **55**, 105946 (2020).
2. Cucinotta, D. & Vanelli, M. WHO declares COVID-19 a pandemic. *Acta Bio Med. Atenei Parm.* **91**, 157–160 (2020).
3. Bikdeli, B. et al. COVID-19 and thrombotic or thromboembolic disease: implications for prevention, antithrombotic therapy, and follow-up: JACC state-of-the-art review. *J. Am. Coll. Cardiol.* **75**, 2950–2973 (2020).
4. Montalvan, V., Lee, J., Bueso, T., De Toledo, J. & Rivas, K. Neurological manifestations of COVID-19 and other coronavirus infections: a systematic review. *Clin. Neurol. Neurosurg.* **194**, 105921 (2020).
5. Clerkin, K. J. et al. COVID-19 and cardiovascular disease. *Circulation* **141**, 1648–1655 (2020).
6. Hoffman, J. I. E. & Kaplan, S. The incidence of congenital heart disease. *J. Am. Coll. Cardiol.* **39**, 1890–1900 (2002).
7. Coperchini, F., Chiovato, L., Croce, L., Magri, F. & Rotondi, M. The cytokine storm in COVID-19: an overview of the involvement of the chemokine/chemokine-receptor system. *Cytokine Growth Factor. Rev.* **53**, 25–32 (2020).
8. Ou, X. et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat. Commun.* **11**, 1620 (2020).
9. Hoffmann, M. et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and Is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280.e8 (2020).
10. Liu, J. et al. Longitudinal characteristics of lymphocyte responses and cytokine profiles in the peripheral blood of SARS-CoV-2 infected patients. *EBioMedicine* **55**, 102763 (2020).
11. Verdoni, L. et al. An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. *Lancet* **395**, 1771–1778 (2020).
An early report of MIS-C from the first wave of COVID-19 in mid-2020.
12. Riphagen, S., Gomez, X., Gonzalez-Martinez, C., Wilkinson, N. & Theocharis, P. Hyperinflammatory shock in children during COVID-19 pandemic. *Lancet* **395**, 1607–1608 (2020).
An early report of MIS-C from the first wave of COVID-19 in mid-2020.
13. Feldstein, L. R. et al. Multisystem inflammatory syndrome in U.S. children and adolescents. *N. Engl. J. Med.* **383**, 334–346 (2020).
An early report of MIS-C from the first wave of COVID-19 in mid-2020.
14. Dufort, E. M. et al. Multisystem inflammatory syndrome in children in New York State. *N. Engl. J. Med.* **383**, 347–358 (2020).
An early report of MIS-C from the first wave of COVID-19 in mid-2020.
15. Cheung, E. W. et al. Multisystem inflammatory syndrome related to COVID-19 in previously healthy children and adolescents in New York City. *JAMA* **324**, 294–296 (2020).
An early report of MIS-C from the first wave of COVID-19 in mid-2020.
16. Galeotti, C. & Bayry, J. Autoimmune and inflammatory diseases following COVID-19. *Nat. Rev. Rheumatol.* **16**, 413–414 (2020).
17. Simpson, J. M. & Newburger, J. W. Multi-system inflammatory syndrome in children in association with COVID-19. *Circulation* **142**, 437–440 (2020).
18. Levin, M. Childhood multisystem inflammatory syndrome — a new challenge in the pandemic. *N. Engl. J. Med.* **383**, 393–395 (2020).
19. Kawasaki, T. Acute febrile mucocutaneous syndrome with lymphoid involvement with specific desquamation of the fingers and toes in children. *Arerugi* **16**, 178–222 (1967).
20. McCrindle Brian, W. et al. Diagnosis, treatment, and long-term management of Kawasaki disease: a scientific statement for health professionals from the American Heart Association. *Circulation* **135**, e927–e999 (2017).
21. Burns, J. C. et al. Genetic variations in the receptor-ligand pair CCR5 and CCL3L1 are important determinants of susceptibility to Kawasaki disease. *J. Infect. Dis.* **192**, 344–349 (2005).
22. Burgner, D. et al. A genome-wide association study identifies novel and functionally related susceptibility loci for Kawasaki disease. *PLoS Genet.* **5**, e1000319 (2009).
23. Burgner, D. & Harnden, A. Kawasaki disease: what is the epidemiology telling us about the etiology? *Int. J. Infect. Dis.* **9**, 185–194 (2005).
24. Son, M. B. F. & Newburger, J. W. Kawasaki disease. *Pediatr. Rev.* **39**, 78–90 (2018).
25. Burns, J. C. The riddle of Kawasaki disease. *N. Engl. J. Med.* **356**, 659–661 (2007).
26. Holman, R. C., Curns, A. T., Belay, E. D., Steiner, C. A. & Schonberger, L. B. Kawasaki syndrome hospitalizations in the United States, 1997 and 2000. *Pediatrics* **112**, 495–501 (2003).
27. Kuijpers, T. W. et al. Kawasaki disease: a maturational defect in immune responsiveness. *J. Infect. Dis.* **180**, 1869–1877 (1999).
28. Green, M. S. The male predominance in the incidence of infectious diseases in children: a postulated explanation for disparities in the literature. *Int. J. Epidemiol.* **21**, 381–386 (1992).
29. Burgner, D. & Levin, M. Genetic susceptibility to infectious diseases. *Pediatr. Infect. Dis. J.* **22**, 1–6 (2003).
30. Harnden, A., Alves, B. & Sheikh, A. Rising incidence of Kawasaki disease in England: analysis of hospital admission data. *BMJ* **324**, 1424–1425 (2002).

31. Bell, D. M., Morens, D. M., Holman, R. C., Hurwitz, E. S. & Hunter, M. K. Kawasaki syndrome in the United States 1976 to 1980. *Am. J. Dis. Child.* **137**, 211–214 (1983).
32. Chang, R.-K. R. Hospitalizations for Kawasaki disease among children in the United States, 1988–1997. *Pediatrics* **109**, e87 (2002).
33. Du, Z.-D. et al. Epidemiologic picture of Kawasaki disease in Beijing from 1995 through 1999. *Pediatr. Infect. Dis. J.* **21**, 103–107 (2002).
34. Park, Y.-W. et al. Epidemiologic study of Kawasaki disease in Korea, 1997-1999: comparison with previous studies during 1991-1996. *J. Korean Med. Sci.* **17**, 453–456 (2002).
35. Uehara, R. & Belay, E. D. Epidemiology of Kawasaki disease in Asia, Europe, and the United States. *J. Epidemiol.* **22**, 79–85 (2012).
36. Yanagawa, H., Yashiro, M., Nakamura, Y., Kawasaki, T. & Kato, H. Results of 12 nationwide epidemiological incidence surveys of Kawasaki disease in Japan. *Arch. Pediatr. Adolesc. Med.* **149**, 779–783 (1995).
37. Burns, J. C. et al. Seasonality and temporal clustering of Kawasaki syndrome. *Epidemiology* **16**, 220–225 (2005).
38. Nakamura, Y., Yanagawa, I. & Kawasaki, T. Temporal and geographical clustering of Kawasaki disease in Japan. *Prog. Clin. Biol. Res.* **250**, 19–32 (1987).
39. Yanagawa, H., Nakamura, Y., Kawasaki, T. & Shigematsu, I. Nationwide epidemic of Kawasaki disease in Japan during winter of 1985-86. *Lancet* **2**, 1138–1139 (1986).
40. Chang, L.-Y. et al. Viral infections associated with Kawasaki disease. *J. Formos. Med. Assoc.* **113**, 148–154 (2014).
41. Benseler, S. M. et al. Infections and Kawasaki disease: implications for coronary artery outcome. *Pediatrics* **116**, e760–e766 (2005).
42. Esper, F. et al. Association between a novel human coronavirus and Kawasaki disease. *J. Infect. Dis.* **191**, 499–502 (2005).
43. Rowley, A. H. et al. Detection of antigen in bronchial epithelium and macrophages in acute Kawasaki disease by use of synthetic antibody. *J. Infect. Dis.* **190**, 856–865 (2004).
44. Cohen, E. & Sundel, R. Kawasaki disease at 50 years. *JAMA Pediatr.* **170**, 1093–1099 (2016).
45. Rowley, A. H. Is Kawasaki disease an infectious disorder? *Int. J. Rheum. Dis.* **21**, 20–25 (2018).
46. Abe, J. et al. Selective expansion of T cells expressing T-cell receptor variable regions Vβ2 and Vβ 8 in Kawasaki disease. *Proc. Natl Acad. Sci. USA* **89**, 4066–4070 (1992).
47. Abe, J. et al. Characterization of T cell repertoire changes in acute Kawasaki disease. *J. Exp. Med.* **177**, 791–796 (1993).
48. Curtis, N., Zheng, R., Lamb, J. R. & Levin, M. Evidence for a superantigen mediated process in Kawasaki disease. *Arch. Dis. Child.* **72**, 308–311 (1995).
49. Nagata, S. et al. Heat shock proteins and superantigenic properties of bacteria from the gastrointestinal tract of patients with Kawasaki disease. *Immunology* **128**, 511–520 (2009).
50. Pietra, B. A., Inocencio, J. D., Giannini, E. H. & Hirsch, R. TCR Vβ family repertoire and T cell activation markers in Kawasaki disease. *J. Immunol.* **153**, 1881–1888 (1994).
51. Mancia, L. et al. Characterization of the T-cell receptor V-β repertoire in Kawasaki disease. *Scand. J. Immunol.* **48**, 443–449 (1998).
52. Jun, J. S., Jung, Y. K. & Lee, D. W. Relationship between vitamin D levels and intravenous immunoglobulin resistance in Kawasaki disease. *Korean J. Pediatr.* **60**, 216–220 (2017).
53. Canning, M. O., Grotenhuis, K., de Wit, H., Ruwhof, C. & Drexhage, H. A. 1-α,25-Dihydroxyvitamin D3 (1,25(OH)(2)D(3)) hampers the maturation of fully active immature dendritic cells from monocytes. *Eur. J. Endocrinol.* **145**, 351–357 (2001).
54. Ben-Shoshan, M. et al. 1α,25-dihydroxyvitamin D3 (Calcitriol) inhibits hypoxia-inducible factor-1/vascular endothelial growth factor pathway in human cancer cells. *Mol. Cancer Ther.* **6**, 1433–1439 (2007).
55. Meyer, K. et al. Breastfeeding and vitamin D supplementation reduce the risk of Kawasaki disease in a German population-based case-control study. *BMC Pediatr.* **19**, 66 (2019).
56. Stagi, S., Rigante, D., Lepri, G., Matucci Cerinic, M. & Falcini, F. Severe vitamin D deficiency in patients with Kawasaki disease: a potential role in the risk to develop heart vascular abnormalities? *Clin. Rheumatol.* **35**, 1865–1872 (2016).
57. Chen, Y.-L., Wang, J.-L. & Li, W.-Q. Prediction of the risk of coronary arterial lesions in Kawasaki disease by serum 25-hydroxyvitamin D3. *Eur. J. Pediatr.* **173**, 1467–1471 (2014).
58. Kim, S. & Eun, L. Y. Iron deficiency anemia as a predictor of coronary artery abnormalities in Kawasaki disease. *Korean J. Pediatr.* **62**, 301–306 (2019).
59. Belkaid, Y. & Hand, T. Role of the microbiota in immunity and inflammation. *Cell* **157**, 121–141 (2014).
60. Kaneko, K., Akagawa, S., Akagawa, Y., Kimata, T. & Tsuji, S. Our evolving understanding of Kawasaki disease pathogenesis: role of the gut microbiota. *Front. Immunol.* **11**, 1616 (2020).
61. Yamashiro, Y., Nagata, S., Ohtsuka, Y., Oguchi, S. & Shimizu, T. Microbiologic studies on the small intestine in Kawasaki disease. *Pediatr. Res.* **39**, 622–624 (1996).
62. Takeshita, S., Kobayashi, I., Kawamura, Y., Tokutomi, T. & Sekine, I. Characteristic profile of intestinal microflora in Kawasaki disease. *Acta Paediatr.* **91**, 783–788 (2002).
63. Chen, J. et al. Altered gut microbiota correlated with systemic inflammation in children with Kawasaki disease. *Sci. Rep.* **10**, 14525 (2020).
64. Kumrah, R., Vignesh, P., Rawat, A. & Singh, S. Immunogenetics of Kawasaki disease. *Clin. Rev. Allergy Immunol.* **59**, 122–139 (2020).
65. Holman, R. C. et al. Kawasaki syndrome in Hawaii. *Pediatr. Infect. Dis. J.* **24**, 429–433 (2005).
66. Fujita, Y. et al. Kawasaki disease in families. *Pediatrics* **84**, 666–669 (1989).
67. Onouchi, Y. et al. Common variants in CASP3 confer susceptibility to Kawasaki disease. *Hum. Mol. Genet.* **19**, 2898–2906 (2010).
The results of this study indicate that multiple variants in the caspase 3 gene (*CASP3*) confer susceptibility to Kawasaki disease.
68. Shimizu, C. et al. Transforming growth factor-beta signaling pathway in patients with Kawasaki disease. *Circ. Cardiovasc. Genet.* **4**, 16–25 (2011).
69. Chang, L., Yang, H.-W., Lin, T.-Y. & Yang, K. D. Perspective of immunopathogenesis and immunotherapies for Kawasaki disease. *Front. Pediatr.* **9**, 697632 (2021).
70. Alphonse, M. P. et al. Inositol-triphosphate 3-kinase c mediates inflammasome activation and treatment response in Kawasaki disease. *J. Immunol.* **197**, 3481–3489 (2016).
71. Onouchi, Y. Molecular genetics of Kawasaki disease. *Pediatr. Res.* **65**, 46R–54R (2009).
72. Burns, J. C. Cyclosporine and coronary outcomes in Kawasaki disease. *J. Pediatr.* **210**, 239–242 (2019).
73. Matsuda, I., Hattori, S., Nagata, N., Fruse, A. & Nambu, H. HLA antigens in mucocutaneous lymph node syndrome. *Am. J. Dis. Child* **131**, 1417–1418 (1977).
74. Kato, S. et al. HLA antigens in Kawasaki disease. *Pediatrics* **61**, 252–255 (1978).
75. Krensky, A. M., Berenberg, W., Shanley, K. & Yunis, E. J. HLA antigens in mucocutaneous lymph node syndrome in New England. *Pediatrics* **67**, 741–743 (1981).
76. Sharma, K. et al. Epigenetics in Kawasaki disease. *Front. Pediatr.* **9**, 626 (2021).
77. Assari, R. et al. Pro-inflammatory cytokine single nucleotide polymorphisms in Kawasaki disease. *Int. J. Rheum. Dis.* **21**, 1120–1126 (2018).
78. Weng, K.-P. et al. IL-1B polymorphism in association with initial intravenous immunoglobulin treatment failure in Taiwanese children with Kawasaki disease. *Circ. J.* **74**, 544–551 (2010).
79. Kim, J.-J. et al. Identification of KCNN2 as a susceptibility locus for coronary artery aneurysms in Kawasaki disease using genome-wide association analysis. *J. Hum. Genet.* **58**, 521–525 (2013).
80. Kwon, Y.-C. et al. Identification of the TIFAB gene as a susceptibility locus for coronary artery aneurysm in patients with Kawasaki disease. *Pediatr. Cardiol.* **40**, 483–488 (2019).
81. Lu, Z. et al. P2RY12:rs7637803 TT variant genotype increases coronary artery aneurysm risk in Kawasaki disease in a southern Chinese population. *J. Gene Med.* **21**, e3066 (2019).
82. Fu, L. Y. et al. The IL-1B gene polymorphisms rs16944 and rs1143627 contribute to an increased risk of coronary artery lesions in southern Chinese children with Kawasaki disease. *J. Immunol. Res.* **2019**, 4730507 (2019).
83. Hara, T. et al. Kawasaki disease: a matter of innate immunity. *Clin. Exp. Immunol.* **186**, 134–143 (2016).
84. Lee, N. H., Choi, H. J. & Kim, Y. H. Clinical usefulness of serum procalcitonin level in distinguishing between Kawasaki disease and other infections in febrile children. *Korean J. Pediatr.* **60**, 112–117 (2017).
85. Tremoulet, A. H. et al. Increased incidence and severity of Kawasaki disease among Filipino-Americans in San Diego county. *Pediatr. Infect. Dis. J.* **30**, 909–911 (2011).
86. Burns, J. C. & Glodé, M. P. Kawasaki syndrome. *Lancet* **364**, 533–544 (2004).
87. Ikeda, K. et al. Unique activation status of peripheral blood mononuclear cells at acute phase of Kawasaki disease. *Clin. Exp. Immunol.* **160**, 246–255 (2010).
88. Armaroli, G. et al. Monocyte-derived interleukin-1β as the driver of S100A12-induced sterile inflammatory activation of human coronary artery endothelial cells: implications for the pathogenesis of Kawasaki disease. *Arthritis Rheumatol.* **71**, 792–804 (2019).
This article provides mechanistic insight on the roles of S100A12 and IL-1β in the pathogenesis of Kawasaki disease, and supports the therapeutic use of IL-1β inhibitors.
89. Zhang, K. et al. Increase in T helper type 17 cells in children with Kawasaki disease is NR4A2 dependent. *Eur. J. Inflamm.* **16**, 2058739218760945 (2018).
90. Wu, Y. et al. Interleukin-6 is prone to be a candidate biomarker for predicting incomplete and IVIG nonresponsive Kawasaki disease rather than coronary artery aneurysm. *Clin. Exp. Med.* **19**, 173–181 (2019).
91. Oharaseki, T. et al. The role of TNF-α in a murine model of Kawasaki disease arteritis induced with a *Candida albicans* cell wall polysaccharide. *Mod. Rheumatol.* **24**, 120–128 (2014).
92. Lin, C. Y., Lin, C. C., Hwang, B. & Chiang, B. Serial changes of serum interleukin-6, interleukin-8, and tumor necrosis factor alpha among patients with Kawasaki disease. *J. Pediatr.* **121**, 924–926 (1992).
93. Guo, M. M.-H. et al. Th17- and Treg-related cytokine and mRNA expression are associated with acute and resolving Kawasaki disease. *Allergy* **70**, 310–318 (2015).
94. Lin, C.-Y., Lin, C.-C., Hwang, B. & Chiang, B. N. Cytokines predict coronary aneurysm formation in Kawasaki disease patients. *Eur. J. Pediatr.* **152**, 309–312 (1993).
95. Rahmani, F. et al. Interleukin 10 and and transforming growth factor β polymorphisms as risk factors for Kawasaki disease: a case-control study and meta-analysis. *Avicenna J. Med. Biotechnol.* **11**, 325–333 (2019).
96. Jia, S., Li, C., Wang, G., Yang, J. & Zu, Y. The T helper type 17/regulatory T cell imbalance in patients with acute Kawasaki disease. *Clin. Exp. Immunol.* **162**, 131–137 (2010).
97. Yoshida, Y. et al. Enhanced formation of neutrophil extracellular traps in Kawasaki disease. *Pediatr. Res.* **87**, 998–1004 (2020).
98. Papayannopoulos, V. Neutrophil extracellular traps in immunity and disease. *Nat. Rev. Immunol.* **18**, 134–147 (2018).
99. Garcia-Romo, G. S. et al. Netting neutrophils are major inducers of type I IFN production in pediatric systemic lupus erythematosus. *Sci. Transl. Med.* **3**, 73ra20 (2011).
100. Kessenbrock, K. et al. Netting neutrophils in autoimmune small-vessel vasculitis. *Nat. Med.* **15**, 623–625 (2009).
101. Corsiero, E., Pratesi, F., Prediletto, E., Bombardieri, M. & Migliorini, P. NETosis as source of autoantigens in rheumatoid arthritis. *Front. Immunol.* **7**, 485 (2016).
102. Grunebaum, E. et al. The role of anti-endothelial cell antibodies in Kawasaki disease – in vitro and in vivo studies. *Clin. Exp. Immunol.* **130**, 233–240 (2002).
103. Consiglio, C. R. et al. The immunology of multisystem inflammatory syndrome in children with COVID-19. *Cell* **183**, 968–981.e7 (2020).
On the basis of systems-level analyses of various immune cells, cytokines and autoantibodies, this article identifies shared and distinct immunological features of MIS-C and Kawasaki disease.
104. Sakurai, Y. Autoimmune aspects of Kawasaki disease. *J. Investig. Allergol. Clin. Immunol.* **29**, 251–261 (2019).
105. Nash, M. C., Shah, V., Reader, J. A. & Dillon, M. J. Anti-neutrophil cytoplasmic antibodies and anti-endothelial cell antibodies are not increased in Kawasaki disease. *Rheumatology* **34**, 882–887 (1995).
106. Li, C. R., Yang, X. Q., Shen, J., Li, Y. B. & Jiang, L. P. Immunoglobulin G subclasses in serum and circulating immune complexes in patients with Kawasaki syndrome. *Pediatr. Infect. Dis. J.* **9**, 544–547 (1990).

107. Levin, M. et al. Platelet immune complex interaction in pathogenesis of Kawasaki disease and childhood polyarteritis. *Br. Med. J. Clin. Res. Ed.* **290**, 1456–1460 (1985).
108. Miyata, K. et al. Circulation immune complexes and granulocytes chemotaxis in Kawasaki disease: the 8th conference on prevention for rheumatic fever and rheumatic heart disease. *Jpn. Circ. J.* **48**, 1350–1353 (1984).
109. Ono, S., Onimaru, T., Kawakami, K., Hokonohara, M. & Miyata, K. Impaired granulocyte chemotaxis and increased circulating immune complexes in Kawasaki disease. *J. Pediatr.* **106**, 567–570 (1985).
110. Kohsaka, T., Abe, J., Asahina, T. & Kobayashi, N. Classical pathway complement activation in Kawasaki syndrome. *J. Allergy Clin. Immunol.* **93**, 520–525 (1994).
111. Biezeveld, M. H. et al. Polymorphisms in the mannose-binding lectin gene as determinants of age-defined risk of coronary artery lesions in Kawasaki disease. *Arthritis Rheum.* **54**, 369–376 (2006).
112. Ohshio, G. et al. High levels of IgA-containing circulating immune complex and secretory IgA in Kawasaki disease. *Microbiol. Immunol.* **31**, 891–898 (1987).
113. Noval Rivas, M. & Arditi, M. Kawasaki disease: pathophysiology and insights from mouse models. *Nat. Rev. Rheumatol.* **16**, 391–405 (2020).
114. Noval Rivas, M. et al. Intestinal permeability and IgA provoke immune vasculitis linked to cardiovascular inflammation. *Immunity* **51**, 508–521.e6 (2019).
115. Olivito, B. et al. Defective FOXP3 expression in patients with acute Kawasaki disease and restoration by intravenous immunoglobulin therapy. *Clin. Exp. Rheumatol.* **28**, 93–97 (2010).
116. Abe, J. et al. Gene expression profiling of the effect of high-dose intravenous Ig in patients with Kawasaki disease. *J. Immunol.* **174**, 5837–5845 (2005).
117. Yoshimura, K. et al. Increased nitric oxide production by neutrophils in early stage of Kawasaki disease. *Eur. J. Pediatr.* **168**, 1037–1041 (2009).
118. Wang, Y. et al. Evaluation of intravenous immunoglobulin resistance and coronary artery lesions in relation to Th1/Th2 cytokine profiles in patients with Kawasaki disease. *Arthritis Rheum.* **65**, 805–814 (2013).
119. Zhu, Y. P. et al. Immune response to intravenous immunoglobulin in patients with Kawasaki disease and MIS-C. *J. Clin. Invest.* https://doi.org/10.1172/JCI147076 (2021).
120. Wang, Z. et al. Single-cell RNA sequencing of peripheral blood mononuclear cells from acute Kawasaki disease patients. *Nat. Commun.* **12**, 5444 (2021).
This article applies single-cell RNA sequencing to report that monocytes are the major contributors of inflammatory responses, including IL-1β and TNF, in Kawasaki disease.
121. Rambabu, N., Mathew, M. J., Kaveri, S. V. & Bayry, J. Boolean analysis of the transcriptomic data to identify novel biomarkers of IVIG response. *Autoimmun. Rev.* **20**, 102850 (2021).
122. Belhadjer, Z. et al. Acute heart failure in multisystem inflammatory syndrome in children (MIS-C) in the context of global SARS-CoV-2 pandemic. *Circulation* **142**, 429–436 (2020).
123. CDC. Multisystem Inflammatory Syndrome in Children (MIS-C) Associated with Coronavirus Disease 2019 (COVID-19). *Health Alert Network* https://emergency.cdc.gov/han/2020/han00432.asp (2020).
124. CDC. COVID Data Tracker. https://covid.cdc.gov/covid-data-tracker (2020).
125. Kaushik, S. et al. Multisystem inflammatory syndrome in children associated with severe acute respiratory syndrome coronavirus 2 infection (MIS-C): a multi-institutional study from New York City. *J. Pediatr.* **224**, 24–29 (2020).
126. Davies, P. et al. Intensive care admissions of children with paediatric inflammatory multisystem syndrome temporally associated with SARS-CoV-2 (PIMS-TS) in the UK: a multicentre observational study. *Lancet Child Adolesc. Health* **4**, 669–677 (2020).
127. Pouletty, M. et al. Paediatric multisystem inflammatory syndrome temporally associated with SARS-CoV-2 mimicking Kawasaki disease (Kawa-COVID-19): a multicentre cohort. *Ann. Rheum. Dis.* **79**, 999–1006 (2020).
An early report of MIS-C from the first wave of COVID-19 in mid-2020.
128. Toubiana, J. et al. Kawasaki-like multisystem inflammatory syndrome in children during the covid-19 pandemic in Paris, France: prospective observational study. *BMJ* **369**, m2094 (2020).
An early report of MIS-C from the first wave of COVID-19 in mid-2020.
129. Capone, C. A. et al. Characteristics, cardiac involvement, and outcomes of multisystem inflammatory syndrome of childhood associated with severe acute respiratory syndrome coronavirus 2 Infection. *J. Pediatr.* **224**, 141–145 (2020).
130. Hameed, S. et al. Spectrum of imaging findings on chest radiographs, US, CT, and MRI images in multisystem inflammatory syndrome in children (MIS-C) associated with COVID-19. *Radiology* **298**, E1–E10 (2020).
131. Whittaker, E. et al. Clinical characteristics of 58 children with a pediatric inflammatory multisystem syndrome temporally associated with SARS-CoV-2. *JAMA* **324**, 259–269 (2020).
132. Moraleda, C. et al. Multi-inflammatory syndrome in children related to SARS-CoV-2 in Spain. *Clin. Infect. Dis.* **72**, e397–e401 (2020).
133. Dhanalakshmi, K. et al. Epidemiological and clinical profile of pediatric inflammatory multisystem syndrome — temporally associated with SARS-CoV-2 (PIMS-TS) in Indian children. *Indian. Pediatr.* **57**, 1010–1014 (2020).
134. Miller, J. et al. Gastrointestinal symptoms as a major presentation component of a novel multisystem inflammatory syndrome in children that is related to coronavirus disease 2019: a single center experience of 44 cases. *Gastroenterology* **159**, 1571–1574.e2 (2020).
135. Belot, A. et al. SARS-CoV-2-related paediatric inflammatory multisystem syndrome, an epidemiological study, France, 1 March to 17 May 2020. *Eurosurveillance* **25**, 2001010 (2020).
136. Lee, P. Y. et al. Distinct clinical and immunological features of SARS–CoV-2–induced multisystem inflammatory syndrome in children. *J. Clin. Invest.* **130**, 5942–5950 (2020).
137. Riollano-Cruz, M. et al. Multisystem inflammatory syndrome in children related to COVID-19: a New York City experience. *J. Med. Virol.* **93**, 424–433 (2021).
138. Ramcharan, T. et al. Paediatric inflammatory multisystem syndrome: temporally associated with SARS-CoV-2 (PIMS-TS): cardiac features, management and short-term outcomes at a UK tertiary paediatric hospital. *Pediatr. Cardiol.* **41**, 1391–1401 (2020).
139. Grimaud, M. et al. Acute myocarditis and multisystem inflammatory emerging disease following SARS-CoV-2 infection in critically ill children. *Ann. Intensive Care* **10**, 69 (2020).
140. Perez-Toledo, M. et al. SARS-CoV-2-specific IgG1/IgG3 but not IgM in children with pediatric inflammatory multi-system syndrome. *Pediatr. Allergy Immunol.* **32**, 1125–1129 (2021).
141. RCPCH. Paediatric multisystem inflammatory syndrome temporally associated with COVID-19. https://www.rcpch.ac.uk/resources/guidance-paediatric-multisystem-inflammatory-syndrome-temporally-associated-covid-19 (2020).
142. WHO. Multisystem inflammatory syndrome in children and adolescents temporally related to COVID-19. https://www.who.int/news-room/commentaries/detail/multisystem-inflammatory-syndrome-in-children-and-adolescents-with-covid-19 (2020).
143. Carlin, R. F. et al. Discriminating multisystem inflammatory syndrome in children requiring treatment from common febrile conditions in outpatient settings. *J. Pediatr.* **229**, 26–32.e2 (2021).
144. Roberts, J. E. et al. Differentiating multisystem inflammatory syndrome in children: a single-centre retrospective cohort study. *Arch. Dis. Child.* https://doi.org/10.1136/archdischild-2021-322290 (2021).
145. Garcīa-Salido, A. et al. Severe manifestations of SARS-CoV-2 in children and adolescents: from COVID-19 pneumonia to multisystem inflammatory syndrome: a multicentre study in pediatric intensive care units in Spain. *Crit. Care* **24**, 666 (2020).
146. Hoang, A. et al. COVID-19 in 7780 pediatric patients: a systematic review. *EClinicalMedicine* **24**, 100433 (2020).
147. Mahase, E. Covid-19: cases of inflammatory syndrome in children surge after urgent alert. *BMJ* **369**, m1990 (2020).
148. Henderson, L. A. et al. American College of Rheumatology clinical guidance for multisystem inflammatory syndrome in children associated with SARS–CoV-2 and Hyperinflammation in pediatric COVID-19: Version 1. *Arthritis Rheumatol.* **72**, 1791–1805 (2020).
Guidelines article on the management of MIS-C.
149. Henderson, L. A. et al. American College of Rheumatology clinical guidance for multisystem inflammatory syndrome in children associated with SARS-CoV-2 and hyperinflammation in pediatric COVID-19: Version 2. *Arthritis Rheumatol.* **73**, e13–e29 (2021).
Guidelines article on the management of MIS-C.
150. Diorio, C. et al. Multisystem inflammatory syndrome in children and COVID-19 are distinct presentations of SARS–CoV-2. *J. Clin. Invest.* **130**, 5967–5975 (2020).
151. Duarte-Neto, A. N. et al. An autopsy study of the spectrum of severe COVID-19 in children: from SARS to different phenotypes of MIS-C. *EclinicalMed* **35**, 100850 (2021).
152. Yonker, L. M. et al. Multisystem inflammatory syndrome in children is driven by zonulin-dependent loss of gut mucosal barrier. *J. Clin. Invest.* **131**, 149633 (2021).
A report on the implication of zonulin-dependent loss of intestinal mucosal permeability in the pathogenesis of MIS-C and the therapeutic application of zonulin antagonist larazotide.
153. Ramaswamy, A. et al. Immune dysregulation and autoreactivity correlate with disease severity in SARS-CoV-2-associated multisystem inflammatory syndrome in children. *Immunity* **54**, 1083–1095.e7 (2021).
Use of single-cell RNA sequencing, flow cytometry, unbiased serum proteomics and TCR repertoire analysis to identify the signatures of severity in MIS-C.
154. Gruber, C. N. et al. Mapping systemic inflammation and antibody responses in multisystem inflammatory syndrome in children (MIS-C). *Cell* **183**, 982–995.e14 (2020).
Mapping of systemic inflammation and autoantibodies in MIS-C.
155. Rostad, C. A. et al. Quantitative SARS-CoV-2 serology in children with multisystem inflammatory syndrome (MIS-C). *Pediatrics* **146**, e2020018242 (2020).
156. Weisberg, S. P. et al. Distinct antibody responses to SARS-CoV-2 in children and adults across the COVID-19 clinical spectrum. *Nat. Immunol.* **22**, 25–31 (2021).
157. Anderson, E. M. et al. SARS-CoV-2 antibody responses in children with MIS-C and mild and severe COVID-19. *J. Pediatr. Infect. Dis. Soc.* **10**, 669–673 (2020).
158. Bartsch, Y. C. et al. Humoral signatures of protective and pathological SARS-CoV-2 infection in children. *Nat. Med.* **27**, 454–462 (2021).
This article reports that children with MIS-C maintain FcγR-binding, inflammatory monocyte/macrophage-activating SARS-CoV-2 IgG antibodies.
159. Mateus, J. et al. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. *Science* **370**, 89–94 (2020).
160. Grifoni, A. et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* **181**, 1489–1501.e15 (2020).
161. Sermet-Gaudelus, I. et al. Prior infection by seasonal coronaviruses, as assessed by serology, does not prevent SARS-CoV-2 infection and disease in children, France, April to June 2020. *Eur. Surveill.* **26**, 2001782 (2021).
162. Rivas, M. N., Porritt, R. A., Cheng, M. H., Bahar, I. & Arditi, M. COVID-19–associated multisystem inflammatory syndrome in children (MIS-C): a novel disease that mimics toxic shock syndrome — the superantigen hypothesis. *J. Allergy Clin. Immunol.* **147**, 57–59 (2021).
163. Cheng, M. H. et al. Superantigenic character of an insert unique to SARS-CoV-2 spike supported by skewed TCR repertoire in patients with hyperinflammation. *Proc. Natl Acad. Sci. USA* **117**, 25254–25262 (2020).
Evidence that SARS-CoV-2 spike protein acts as a superantigen leading to HLA class I-associated expansion of TRBV11-2 T cells in MIS-C.
164. Porritt, R. A. et al. HLA class I-associated expansion of TRBV11-2 T cells in multisystem inflammatory syndrome in children. *J. Clin. Invest.* **131**, e146614 (2021).
Evidence that SARS-CoV-2 spike protein acts as a superantigen leading to HLA class I-associated expansion of TRBV11-2 T cells in MIS-C.
165. Moreews, M. et al. Polyclonal expansion of TCR Vβ 21.3^{+} CD4^{+} and CD8^{+} T cells is a hallmark of

multisystem inflammatory syndrome in children. *Sci. Immunol.* **6**, eabh1516 (2021).
Patients with MIS-C present a unique CD4⁺ and CD8⁺ T cell TCR repertoire with a polyclonal expansion of CX_3CR1-expressing activated (TRBV11-2)/Vβ21.3⁺ clonotype.
166. McCormick, J. K., Yarwood, J. M. & Schlievert, P. M. Toxic shock syndrome and bacterial superantigens: an update. *Annu. Rev. Microbiol.* **55**, 77–104 (2001).
167. Radujkovic, A. et al. Vitamin D deficiency and outcome of COVID-19 patients. *Nutrients* **12**, 2757 (2020).
168. Urashima, M. et al. Randomized trial of vitamin D supplementation to prevent seasonal influenza A in schoolchildren. *Am. J. Clin. Nutr.* **91**, 1255–1260 (2010).
169. Omenetti, S. & Pizarro, T. T. The T_{reg}/T_H17 axis: a dynamic balance regulated by the gut microbiome. *Front. Immunol.* **6**, 639 (2015).
170. Yeoh, Y. K. et al. Gut microbiota composition reflects disease severity and dysfunctional immune responses in patients with COVID-19. *Gut* **70**, 698–706 (2021).
171. Zuo, T. et al. Alterations in fecal fungal microbiome of patients with COVID-19 during time of hospitalization until discharge. *Gastroenterology* **159**, 1302–1310. e5 (2020).
172. Merenstein, C. et al. Signatures of COVID-19 severity and immune response in the respiratory tract microbiome. *mBio* **12**, e0177721 (2021).
173. Zuo, T. et al. Alterations in gut microbiota of patients with COVID-19 during time of hospitalization. *Gastroenterology* **159**, 944–955.e8 (2020).
174. Xu, R. et al. Progressive deterioration of the upper respiratory tract and the gut microbiomes in children during the early infection stages of COVID-19. *J. Genet. Genomics* https://doi.org/10.1016/j.jgg.2021.05.004 (2021).
175. Hurst, J. H. et al. Age-related changes in the upper respiratory microbiome are associated with SARS-CoV-2 susceptibility and illness severity. Preprint at *medRxiv* https://www.medrxiv.org/content/10.1101/2021.03.20.21252680v1 (2021).
176. Chou, J. et al. Mechanisms underlying genetic susceptibility to multisystem inflammatory syndrome in children (MIS-C). *J. Allergy Clin. Immunol.* **148**, 732–738.e1 (2021).
177. Belay, E. D. et al. Trends in geographic and temporal distribution of US children with multisystem inflammatory syndrome during the COVID-19 pandemic. *JAMA Pediatr.* **175**, 837–845 (2021).
178. Krishnan, L., Ogunwole, S. M. & Cooper, L. A. Historical insights on coronavirus disease 2019 (COVID-19), the 1918 influenza pandemic, and racial disparities: illuminating a path forward. *Ann. Intern. Med.* **173**, 474–481 (2020).
179. Alcendor, D. J. Racial disparities-associated COVID-19 mortality among minority populations in the US. *J. Clin. Med.* **9**, 2442 (2020).
180. Rafferty, M. S. et al. Multisystem inflammatory syndrome in children (MIS-C) and the coronavirus pandemic: current knowledge and implications for public health. *J. Infect. Public Health* **14**, 484–494 (2021).
181. Nguyen, A. et al. Human leukocyte antigen susceptibility map for severe acute respiratory syndrome coronavirus 2. *J. Virol.* **94**, e00510–e00520 (2020).
182. Zhang, Q. et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **370**, eabd4570v (2020).
183. Bastard, P. et al. Insufficient type I IFN immunity underlies life-threatening COVID-19 pneumonia. *C. R. Biol.* **344**, 19–25 (2021).
184. Lee, P. Y. et al. Immune dysregulation and multisystem inflammatory syndrome in children (MIS-C) in individuals with haploinsufficiency of SOCS1. *J. Allergy Clin. Immunol.* **146**, 1194–1200.e1 (2020).
This article provides the first evidence on genetic susceptibility to MIS-C.
185. Vella, L. A. et al. Deep immune profiling of MIS-C demonstrates marked but transient immune activation compared to adult and pediatric COVID-19. *Sci. Immunol.* **6**, eabf7570 (2021).
The article identifies a unique and strong activation of vascular patrolling CX_3CR1^+ CD8⁺ T cells in MIS-C.
186. Carter, M. J. et al. Peripheral immunophenotypes in children with multisystem inflammatory syndrome associated with SARS-CoV-2 infection. *Nat. Med.* **26**, 1701–1707 (2020).
Based on the immunoprofiling of MIS-C patients, this article suggests that MIS-C might represent a distinct immunopathogenic disease.
187. Pierce, C. A. et al. Immune responses to SARS-CoV-2 infection in hospitalized pediatric and adult patients. *Sci. Transl. Med.* **12**, eabd5487 (2020).
188. Caldarale, F. et al. Plasmacytoid dendritic cells depletion and elevation of IFN-γ dependent chemokines CXCL9 and CXCL10 in children with multisystem inflammatory syndrome. *Front. Immunol.* **12**, 654587 (2021).
189. Peart Akindele, N. et al. Distinct cytokine and chemokine dysregulation in hospitalized children with acute COVID-19 and multisystem inflammatory syndrome with similar levels of nasopharyngeal SARS-CoV-2 shedding. *J. Infect. Dis.* **224**, 606–615 (2021).
190. Diorio, C. et al. Proteomic profiling of MIS-C patients reveals heterogeneity relating to interferon gamma dysregulation and vascular endothelial dysfunction. Preprint at *medRxiv* https://www.medrxiv.org/content/10.1101/2021.04.13.21255439v1 (2021).
191. Zuo, Y. et al. Neutrophil extracellular traps in COVID-19. *JCI Insight* **5**, e138999 (2020).
192. Middleton, E. A. et al. Neutrophil extracellular traps contribute to immunothrombosis in COVID-19 acute respiratory distress syndrome. *Blood* **136**, 1169–1179 (2020).
193. Yamashita, K., Takaori-Kondo, A. & Mizugishi, K. Exaggerated neutrophil extracellular trap formation in Kawasaki disease: a key phenomenon behind the outbreak in western countries? *Ann. Rheum. Dis.* https://doi.org/10.1136/annrheumdis-2020-218593 (2020).
194. Seery, V. et al. Blood neutrophils from children with COVID-19 exhibit both inflammatory and anti-inflammatory markers. *EBioMedicine* **67**, 103357 (2021).
195. Beckmann, N. D. et al. Downregulation of exhausted cytotoxic T cells in gene expression networks of multisystem inflammatory syndrome in children. *Nat. Commun.* **12**, 4854 (2021).
196. Vojdani, A., Vojdani, E. & Kharrazian, D. Reaction of human monoclonal antibodies to SARS-CoV-2 proteins with tissue antigens: implications for autoimmune diseases. *Front. Immunol.* **11**, 617089 (2020).
197. An, H. & Park, J. Molecular mimicry map (3M) of SARS-CoV-2: prediction of potentially immunopathogenic SARS-CoV-2 epitopes via a novel immunoinformatic approach. Preprint at *bioRxiv* https://www.biorxiv.org/content/10.1101/2020.11.12.344424v1 (2020).
198. European Centre for Disease Prevention and Control. Rapid risk assessment: paediatric inflammatory multisystem syndrome and SARS-CoV-2 infection in children. https://www.ecdc.europa.eu/en/publications-data/paediatric-inflammatory-multisystem-syndrome-and-sars-cov-2-rapid-risk-assessment (2020).
199. Diorio, C. et al. Evidence of thrombotic microangiopathy in children with SARS-CoV-2 across the spectrum of clinical presentations. *Blood Adv.* **4**, 6051–6063 (2020).
200. Whitworth, H. B. et al. Rate of thrombosis in children and adolescents hospitalized with COVID-19 or MIS-C. *Blood* **138**, 190–198 (2021).
201. Endo, Y., Takahashi, M. & Fujita, T. Lectin complement system and pattern recognition. *Immunobiology* **211**, 283–293 (2006).
202. Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S. & Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* **369**, 330–333 (2020).
203. Polycarpou, A., Grigoriadou, S., Klavinskis, L. & Sacks, S. Does the lectin complement pathway link Kawasaki disease and SARS-CoV-2? *Front. Immunol.* **11**, 604512 (2020).
204. Dove, M. L. et al. Multisystem inflammatory syndrome in children: survey of protocols for early hospital evaluation and management. *J. Pediatr.* **229**, 33–40 (2021).
205. Jonat, B. et al. Multisystem inflammatory syndrome in children associated with coronavirus disease 2019 in a children's hospital in New York City: patient characteristics and an institutional protocol for evaluation, management, and follow-up. *Pediatr. Crit. Care Med.* **22**, e178–e191 (2021).
206. Ouldali, N. et al. Association of intravenous immunoglobulins plus methylprednisolone vs immunoglobulins alone with course of fever in multisystem inflammatory syndrome in children. *JAMA* **325**, 855–864 (2021).
207. McArdle, A. J. et al. Treatment of multisystem inflammatory syndrome in children. *N. Engl. J. Med.* **385**, 11–22 (2021).
This is an international observational cohort study aimed at supporting the treatment decisions in patients with MIS-C.
208. Guillaume, M.-P., Reumaux, H. & Dubos, F. Usefulness and safety of anakinra in refractory Kawasaki disease complicated by coronary artery aneurysm. *Cardiol. Young.* **28**, 739–742 (2018).
209. Kone-Paut, I. et al. The use of interleukin 1 receptor antagonist (anakinra) in Kawasaki disease: a retrospective cases series. *Autoimmun. Rev.* **17**, 768–774 (2018).
210. Jiang, L. et al. COVID-19 and multisystem inflammatory syndrome in children and adolescents. *Lancet Infect. Dis.* **20**, e276–e288 (2020).
211. Feldstein, L. R. et al. Characteristics and outcomes of US children and adolescents with multisystem inflammatory syndrome in children (MIS-C) compared with severe acute COVID-19. *JAMA* **325**, 1074–1087 (2021).
212. Hoste, L., Van Paemel, R. & Haerynck, F. Multisystem inflammatory syndrome in children related to COVID-19: a systematic review. *Eur. J. Pediatr.* **180**, 2019–2034 (2021).
213. Xu, Q.-Q. et al. Evaluation of left ventricular systolic strain in children with Kawasaki disease. *Pediatr. Cardiol.* **35**, 1191–1197 (2014).
214. Surve, S. V., Joseph, S., Gajbhiye, R. K., Mahale, S. D. & Modi, D. N. A systematic review on multisystem inflammatory syndrome in children (MIS-C) with COVID-19: development of a scoring system for clinical diagnosis. Preprint at *medRxiv* https://www.medrxiv.org/content/10.1101/2021.04.23.21255879v1 (2021).
215. Sahoo, D. et al. An AI-guided invariant signature places MIS-C with Kawasaki disease in a continuum of host immune responses. Preprint at *bioRxiv* https://www.biorxiv.org/content/10.1101/2021.04.11.439347v2 (2021).
216. Newburger, J. W. et al. A single intravenous infusion of gamma globulin as compared with four infusions in the treatment of acute Kawasaki syndrome. *N. Engl. J. Med.* **324**, 1633–1639 (1991).
217. Abrams, J. Y. et al. Multisystem inflammatory syndrome in children associated with severe acute respiratory syndrome coronavirus 2: a systematic review. *J. Pediatr.* **226**, 45–54.e1 (2020).
218. Ouldali, N. et al. Emergence of Kawasaki disease related to SARS-CoV-2 infection in an epicentre of the French COVID-19 epidemic: a time-series analysis. *Lancet Child. Adolesc. Health* **4**, 662–668 (2020).
219. Choe, Y. J. et al. Surveillance of COVID-19-associated multisystem inflammatory syndrome in children, South Korea. *Emerg. Infect. Dis.* **27**, 1196–1200 (2021).
220. Li, W., Tang, Y., Shi, Y., Chen, Y. & Liu, E. Why multisystem inflammatory syndrome in children has been less commonly described in Asia? *Transl. Pediatr.* **9**, 873–875 (2020).
221. Son, M. B. F. et al. Treatment of Kawasaki disease: analysis of 27 US pediatric hospitals from 2001 to 2006. *Pediatrics* **124**, 1–8 (2009).
222. Coon, E. R., Wilkes, J., Bratton, S. L. & Srivastava, R. Paediatric overdiagnosis modelled by coronary abnormality trends in Kawasaki disease. *Arch. Dis. Child.* **103**, 937–941 (2018).
223. Ghelani, S. J., Kwatra, N. S. & Spurney, C. F. Can coronary artery involvement in Kawasaki disease be predicted? *Diagnostics* **3**, 232–243 (2013).
224. Gong, G. W. K., McCrindle, B. W., Ching, J. C. & Yeung, R. S. M. Arthritis presenting during the acute phase of Kawasaki disease. *J. Pediatr.* **148**, 800–805 (2006).
225. Blondiaux, E. et al. Cardiac MRI in children with multisystem inflammatory syndrome associated with COVID-19. *Radiology* **297**, E283–E288 (2020).
226. Gaitonde, M. et al. COVID-19-related multisystem inflammatory syndrome in children affects left ventricular function and global strain compared with Kawasaki disease. *J. Am. Soc. Echocardiogr.* **33**, 1285–1287 (2020).
227. Jhaveri, S. et al. Longitudinal echocardiographic assessment of coronary arteries and left ventricular function following multisystem inflammatory syndrome in children. *J. Pediatr.* **228**, 290–293 (2020).
228. Gamez-Gonzalez, L. B. et al. Kawasaki disease shock syndrome: unique and severe subtype of Kawasaki disease. *Pediatr. Int.* **60**, 781–790 (2018).
229. Tsuda, E. et al. Changes in causes of sudden deaths by decade in patients with coronary arterial lesions due to Kawasaki disease. *Cardiol. Young.* **15**, 481–488 (2005).
230. Gatterre, P. et al. Kawasaki disease: an unexpected etiology of shock and multiple organ dysfunction syndrome. *Intensive Care Med.* **38**, 872–878 (2012).

231. Kanegaye, J. T. et al. Recognition of a Kawasaki disease shock syndrome. *Pediatrics* **123**, e783–e789 (2009).
232. Li, Y. et al. Kawasaki disease shock syndrome: clinical characteristics and possible use of IL-6, IL-10 and IFN-γ as biomarkers for early recognition. *Pediatr. Rheumatol.* **17**, 1 (2019).
233. Ma, L., Zhang, Y.-Y. & Yu, H.-G. Clinical manifestations of Kawasaki disease shock syndrome. *Clin. Pediatr.* **57**, 428–435 (2018).
234. Alsaied, T. et al. Review of cardiac involvement in multisystem inflammatory syndrome in children. *Circulation* **143**, 78–88 (2021).
235. Valverde, I. et al. Acute cardiovascular manifestations in 286 children with multisystem inflammatory syndrome associated with COVID-19 infection in Europe. *Circulation* **143**, 21–32 (2021).
236. Toubiana, J. et al. Distinctive features of Kawasaki disease following SARS-CoV-2 infection: a controlled study in Paris, France. *J. Clin. Immunol.* **41**, 526–535 (2021).
237. Suzuki, J. et al. Kawasaki disease shock syndrome in Japan and comparison with multisystem inflammatory syndrome in children in European countries. *Front. Pediatr.* **9**, 625456 (2021).
238. Kaneko, K., Takahashi, K., Fujiwara, S., Maruyama, T. & Obinata, K. Kawasaki disease followed by haemophagocytic syndrome. *Eur. J. Pediatr.* **157**, 610–611 (1998).
239. Cummings, C., McCarthy, P., van Hoff, J. & Porter, G. Kawasaki disease associated with reactive hemophagocytic lymphohistiocytosis. *Pediatr. Infect. Dis. J.* **27**, 1116–1118 (2008).
240. Latino, G. A., Manlhiot, C., Yeung, R. S. M., Chahal, N. & McCrindle, B. W. Macrophage activation syndrome in the acute phase of Kawasaki disease. *J. Pediatr. Hematol. Oncol.* **32**, 527–531 (2010).
241. Natoli, V., Rosina, S. & Ravelli, A. Is macrophage activation syndrome in Kawasaki disease underrecognized? *J. Rheumatol.* **48**, 162–164 (2021).
242. Jinkawa, A. et al. Cytokine profile of macrophage activation syndrome associated with Kawasaki disease. *Cytokine* **119**, 52–56 (2019).
243. Aydın, F. et al. Comparison of baseline laboratory findings of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis and multisystem inflammatory syndrome in children. *Int. J. Rheum. Dis.* **24**, 542–547 (2021).
244. Zulian, F. et al. Acute surgical abdomen as presenting manifestation of Kawasaki disease. *J. Pediatr.* **142**, 731–735 (2003).
245. Miyake, T. et al. Small bowel pseudo-obstruction in Kawasaki disease. *Pediatr. Radiol.* **17**, 383–386 (1987).
246. Halepas, S. et al. Oral manifestations of COVID-2019-related multisystem inflammatory syndrome in children: a review of 47 pediatric patients. *J. Am. Dent. Assoc.* **152**, 202–208 (2021).
247. Saguil, A., Fargo, M. & Grogan, S. Diagnosis and management of Kawasaki disease. *Am. Fam. Physician* **91**, 365–371 (2015).
248. Chen, T.-H., Kao, W.-T. & Tseng, Y.-H. Gastrointestinal involvements in children with COVID-related multisystem inflammatory syndrome. *Gastroenterology* **160**, 1887–1888 (2020).
249. LaRovere, K. L. et al. Neurologic involvement in children and adolescents hospitalized in the United States for COVID-19 or multisystem inflammatory syndrome. *JAMA Neurol.* **78**, 536–547 (2021).
250. Dengler, L. D. et al. Cerebrospinal fluid profile in patients with acute Kawasaki disease. *Pediatr. Infect. Dis. J.* **17**, 478–481 (1998).
251. Liu, X. et al. Neurological involvement in Kawasaki disease: a retrospective study. *Pediatr. Rheumatol. Online J.* **18**, 61 (2020).
252. Kuo, C.-C. et al. Characteristics of children with Kawasaki disease requiring intensive care: 10 years' experience at a tertiary pediatric hospital. *J. Microbiol. Immunol. Infect.* **51**, 184–190 (2018).
253. Radia, T. et al. Multi-system inflammatory syndrome in children & adolescents (MIS-C): a systematic review of clinical features and presentation. *Paediatr. Respir. Rev.* **38**, 51–57 (2020).
254. Abrams, J. Y. et al. Factors linked to severe outcomes in multisystem inflammatory syndrome in children (MIS-C) in the USA: a retrospective surveillance study. *Lancet Child. Adolesc. Health* **5**, 323–331 (2021).
255. Ouldali, N. et al. Factors associated with severe SARS-CoV-2 infection. *Pediatrics* **147**, e2020023432 (2021).
256. Ye, Q., Wang, B. & Mao, J. The pathogenesis and treatment of the 'cytokine storm' in COVID-19. *J. Infect.* **80**, 607–613 (2020).
257. Dolinger, M. T. et al. Pediatric Crohn disease and multisystem inflammatory syndrome in children (MIS-C) and COVID-19 treated with infliximab. *J. Pediatr. Gastroenterol. Nutr.* **71**, 153–155 (2020).
258. Esteve-Sole, A. et al. Similarities and differences between the immunopathogenesis of COVID-19-related pediatric multisystem inflammatory syndrome and Kawasaki disease. *J. Clin. Invest.* **131**, e144554 (2021).
A comparative analysis of circulating cytokines and chemokines to have an insight on the similarities and differences between MIS-C and Kawasaki disease.
259. Al-Ghafry, M. et al. Multisystem inflammatory syndrome in children (MIS-C) and the prothrombotic state: coagulation profiles and rotational thromboelastometry in a MIS-C cohort. *J. Thromb. Haemost.* **19**, 1764–1770 (2021).
260. Martinez, O. M., Bridges, N. D., Goldmuntz, E. & Pascual, V. The immune roadmap for understanding multi-system inflammatory syndrome in children: opportunities and challenges. *Nat. Med.* **26**, 1819–1824 (2020).
261. Shobhavat, L. et al. Multisystem inflammatory syndrome in children: clinical features and management-intensive care experience from a pediatric public hospital in western India. *Indian. J. Crit. Care Med. Peer Rev.* **24**, 1089–1094 (2020).
262. Niño-Taravilla, C., Otaola-Arca, H., Lara-Aguilera, N., Zuleta-Morales, Y. & Ortiz-Fritz, P. Multisystem inflammatory syndrome in children, Chile, May-August 2020. *Emerg. Infect. Dis.* **27**, 1457–1461 (2021).
263. Tolunay, O. et al. Multisystem inflammatory syndrome in children (MIS-C) associated with COVID-19: a case series experience in a tertiary care hospital of southern Turkey. *J. Trop. Pediatr.* **67**, fmab050 (2021).
264. Young, T. K. et al. Mucocutaneous manifestations of multisystem inflammatory syndrome in children during the COVID-19 pandemic. *JAMA Dermatol.* **157**, 207–212 (2021).
265. Newburger Jane, W. et al. Diagnosis, treatment, and long-term management of Kawasaki disease. *Circulation* **110**, 2747–2771 (2004).
266. Ozdemir, H. et al. Clinical and epidemiological characteristics of children with Kawasaki disease in Turkey. *J. Trop. Pediatr.* **56**, 260–262 (2010).
267. Fukushige, J., Takahashi, N., Ueda, Y. & Ueda, K. Incidence and clinical features of incomplete Kawasaki disease. *Acta Paediatr. Oslo. Nor.* **83**, 1057–1060 (1994).
268. Sung, R. Y. T. et al. Lack of association of cervical lymphadenopathy and coronary artery complications in Kawasaki disease. *Pediatr. Infect. Dis. J.* **25**, 521–525 (2006).
269. Smith, L. B., Newburger, J. W. & Burns, J. C. Kawasaki syndrome and the eye. *Pediatr. Infect. Dis. J.* **8**, 116–118 (1989).
270. Baker, A. L. et al. Associated symptoms in the ten days before diagnosis of Kawasaki disease. *J. Pediatr.* **154**, 592–595.e2 (2009).

Acknowledgements
The work of J.B. is supported by Agence Nationale de la Recherche, France (Flash COVID-19; ANR-20-COVI-0093-COVIMUNE).

Author contributions
C.S., M.G. and J.B. researched data for the article. C.S., M.G. and J.B. wrote the article. All authors provided substantial contribution to discussion of the content, and reviewed and approved the text before submission.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Rheumatology thanks H. Bassiri, who co-reviewed with A. Blatz (ECR), and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Interleukin-2 and regulatory T cells in rheumatic diseases

Antonios G. A. Kolios[1,2], *George C. Tsokos*[1] *and David Klatzmann*[3,4]✉

Abstract | Failure of regulatory T (T_{reg}) cells to properly control immune responses leads invariably to autoimmunity and organ damage. Decreased numbers or impaired function of T_{reg} cells, especially in the context of inflammation, has been documented in many human autoimmune diseases. Restoration of T_{reg} cell fitness and/or expansion of their numbers using low-dose natural IL-2, the main cytokine driving T_{reg} cell survival and function, has demonstrated clinical efficacy in early clinical trials. Genetically modified IL-2 with an extended half-life and increased selectivity for T_{reg} cells is now in clinical development. Administration of IL-2 combined with therapies targeting other pathways involved in the expression of autoimmune diseases should further enhance its therapeutic potential. Ongoing clinical efforts that capitalize on the early clinical success of IL-2 treatment should bring the use of this cytokine to the forefront of biological treatments for autoimmune diseases.

The maintenance of immune homeostasis is of paramount importance in the prevention of autoimmunity and the development of excessive effector immune response and host tissue injury[1]. In this respect, a potent T cell growth factor, later named interleukin-2 (IL-2), was discovered in 1976 in supernatants of activated T cell culture[2] and cloned in 1983 (REF.[3]). IL-2 at high doses seemed to be important for the differentiation of effector and memory effector T cells, which led to the initial development of this cytokine as an anticancer drug. It took 30 years to recognize that the main function of IL-2 is directed towards regulatory T (T_{reg}) cells rather than conventional T cells[4,5]. Indeed, IL-2 is the non-redundant key cytokine for differentiation, survival and function of T_{reg} cells[6–9]. T_{reg} cells were initially defined as $CD4^+CD25^{hi}$ T cells[10], which express high levels of the α chain of the IL-2 receptor (IL-2Rα, also known as CD25). In humans, the consensus definition of T_{reg} cells is that they are $CD4^+FOXP3^+CD25^+CD127^{low}$ T cells[11], and they account for approximately 5–10% of peripheral $CD4^+$ T cells[12–14]. However, in some diseases, such as systemic lupus erythematosus (SLE), CD25 and FOXP3 can be downregulated on T_{reg} cells, making the identification of T_{reg} cells difficult[15,16]. T_{reg} cells are crucial for the development and maintenance of self-tolerance. Most T_{reg} cells develop in the thymus, while some can develop from naive $CD4^+$ T cells in the periphery[17]. In addition, 'induced' T_{reg} cells can be generated in vitro from activated $CD4^+$ T cells by TGFβ treatment. Experimental ablation of T_{reg} cells in mice immediately triggers severe inflammation and the development of multi-organ autoimmune diseases[5]. This observation indicates that, in healthy individuals, conventional T cells can undergo exaggerated activation and attack normal tissues if not controlled by T_{reg} cells.

A small population of $CD8^+$ T_{reg} cells with a phenotype similar to that of $CD4^+$ T_{reg} cells (that is, a $CD8^+FOXP3^+CD25^+CD127^{low}$ phenotype) has been identified in mice and humans[18–20], although their function is poorly defined. Other T cells also display immunosuppressive activity, including TGFβ-producing T helper 3 (T_H3) cells, IL-10 producing type 1 T_{reg} (Tr1) cells[21], $CD8^+FOXP3^+CD45RC^{low}$ T cells and $CD8^+FOXP3^+CD25^+TNFR2^+$ T cells[22].

Mutation of molecules in the IL-2 signalling pathway in humans and mice is associated with systemic inflammation linked to T_{reg} cell deficiencies. Risk variants in the genes encoding IL-2, the IL-2R subunits IL-2Rα or IL-2Rβ, and the downstream signalling factors STAT5A, STAT5B, FOXP3 and PTPN2, lead to inflammation and/or autoimmunity[6,23]. T_{reg} cell deficiency or reduced function have been reported in rheumatic diseases, including rheumatoid arthritis (RA), SLE, primary Sjögren syndrome (pSS), ankylosing spondylitis[24], psoriatic arthritis[25], juvenile idiopathic arthritis[26], systemic sclerosis (SSc)[27,28], sarcoidosis and others[5,29].

Restoration of the immune balance by increasing T_{reg} cell number and/or function thus represents the basis of therapeutic efforts to restore tolerance and mitigate tissue inflammation and injury. In this Review, we discuss T_{reg} cell dysregulation in rheumatic diseases and the evolving therapeutic efforts to restore the function of these cells with IL-2 or IL-2-derived molecules.

[1]*Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA.*

[2]*Department of Dermatology, University Hospital Zurich, Zurich, Switzerland.*

[3]*Sorbonne Université, INSERM, Immunology-Immunopathology-Immunotherapy (i3), Paris, France.*

[4]*AP-HP, Hôpital Pitié-Salpêtrière, Clinical Investigation Center for Biotherapies (CIC-BTi) and Immunology-Inflammation-Infectiology and Dermatology Department (3iD), Paris, France.*

✉*e-mail:david.klatzmann@sorbonne-universite.fr*

https://doi.org/10.1038/s41584-021-00707-x

Key points

- Dysregulation of regulatory T (T_{reg}) cells and a consequent inability to correctly control immune responses leads invariably to autoimmunity and organ damage.
- Numerical or functional impairment of T_{reg} cells, especially in the context of inflammation, occurs in many human autoimmune diseases, such as rheumatoid arthritis and systemic lupus erythematosus (SLE).
- Restoration of the immune balance by restoring T_{reg} cell numbers and/or function is therefore the basis of therapeutic efforts to restore tolerance and mitigate tissue inflammation and injury.
- Low-dose IL-2 (0.3–3 MIU daily in humans) is now an established therapy in autoimmune diseases; it stimulates mainly T_{reg} cells and is well tolerated.
- Most studies of low-dose IL-2 in SLE have shown a clinically significant improvement (as assessed by the SLE Disease Activity Index (SLEDAI) or the SLE Responder Index 4 (SRI-4)) and a reduction in concomitant prednisone of ≥50% in 44–67% of patients.
- Genetically modified IL-2 proteins ('muteins') or IL-2–anti-IL-2R antibody complexes with an extended half-life and increased selectivity for T_{reg} cells are now in clinical development.

IL-2 biology

IL-2–IL-2R signalling. IL-2 is a four α-helix bundle cytokine that is produced mostly by conventional T cells after engagement of the T cell receptor (TCR) and the costimulatory molecule CD28 (REFS[6,30]). Activated $CD4^+$ T cells produce IL-2 in response to antigen-presenting dendritic cells that colocalize with T_{reg} cells in lymphoid organs[4,31–35]. IL-2 is also produced by activated $CD8^+$ T cells, natural killer (NK) cells and NK T cells and, in some conditions, dendritic cells and mast cells, albeit at lower levels than by $CD4^+$ T cells[5]. Of note, FOXP3 represses the expression of *Il2* in T_{reg} cells[36]. Consequently, T_{reg} cells cannot produce IL-2 and rely entirely on an exogenous supply of this cytokine for their survival and function.

IL-2R consists of three subunits: the α chain (IL-2Rα), the β chain (IL-2Rβ; also known as CD122) and the γ chain (IL-2Rγ; also known as the common γ chain (γ_c) or CD132)[37]. IL-2R occurs in monomeric (IL-2Rα), dimeric (IL-2Rβ–IL-2Rγ) or trimeric (IL-Rα–IL-2Rβ–IL-2Rγ) variants[4,38], which bind to IL-2 with increasing affinity[5]. Dimeric IL-2R is constitutively expressed at high levels on the cell surface of memory $CD8^+$ T cells and NK cells and at lower levels on memory $CD4^+$ T cells and naive T cells[23]. Trimeric IL-2R is constitutively expressed at high levels on T_{reg} cells[10] and at low levels on human and mouse B cells, type 2 innate lymphoid cells and activated conventional T cells[39]. The affinity of trimeric IL-2R for IL-2 ($K_d = 10^{-11}$ M) is two-log higher than that of dimeric IL-2R ($K_d = 10^{-9}$ M) and three-log higher than IL-2Rα ($K_d = 10^{-8}$ M) (FIG. 1).

Three major pathways are involved in downstream IL-2 signalling[40]: JAK–STAT, PI3K–AKT–mTOR and MAPK pathways[38,39,41,42]. JAK–STAT activation accounts for >90% of IL-2 signalling, especially STAT5 phosphorylation in T_{reg} cells[43]. After IL-2 binding, the IL-2R complex is internalized, at which point the fate of the IL-2R subunits diverge: while IL-2Rα is recycled, IL-2Rβ and IL-2Rγ are degraded[44].

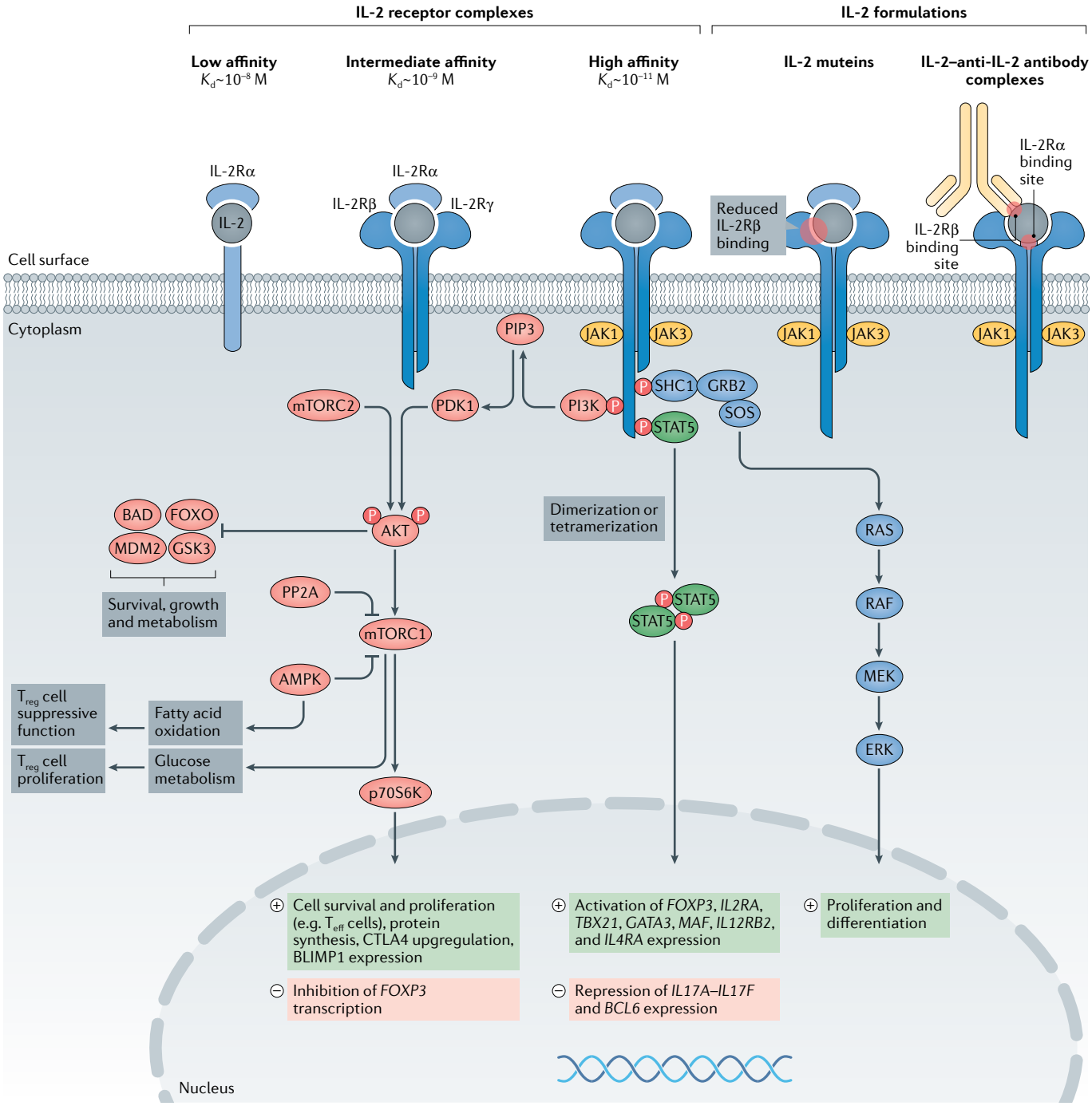
Various transcription factors are involved in expression of *Il2* upon TCR activation primarily in $CD4^+$ T cells: nuclear factor of activated T cells (NFAT) family members, activator protein 1 (AP-1), nuclear factor-κB (NF-κB), octamer-binding protein 1, and the high-mobility group proteins HMG-I and HMG-Y promote *IL2* expression, whereas NFAT members in complex with FOXP3 inhibit IL-2 expression[4,23,45]. Furthermore, positive regulators of *IL2* expression include SP1, early growth response protein 1 and GA-binding protein. Protein phosphatase 2A (PP2A) together with AMP-activated protein kinase are both upregulated in T_{reg} cells and constrain mTORC1 signalling, minimize glucose metabolism and promote FOXP3 expression[46,47]. Negative regulators of IL-2 production include zinc-finger E-box-binding homeobox 1, cAMP response element-binding protein and B lymphocyte-induced maturation protein 1 (BLIMP1)[38,48]. Control of PI3K by its negative regulator PTEN is important for T_{reg} cell lineage stability and homeostasis[49,50].

Fig. 1 | **IL-2 signalling pathways and formulations.** The IL-2 receptor (IL-2R) consists of three subunits (IL-2Rα, IL-2Rβ and IL-2Rγ) that form monomeric (IL-2Rα), dimeric (IL-2Rβ–IL-2Rγ) or trimeric (IL-Rα–IL-2Rβ–IL-2Rγ) receptors that bind to IL-2 with increasing affinity (K_d). Dimeric IL-2R is constitutively expressed on memory $CD8^+$ T cells and natural killer cells and trimeric IL-2R on regulatory T (T_{reg}) cells, indicating that T_{reg} cells are more sensitive to low levels of IL-2 than other cell types. On IL-2 binding, three major pathways are responsible for downstream signalling, namely, PI3K–AKT–mTOR (left), JAK–STAT (middle) and MAPK (right), which are schematically depicted for trimeric IL-2R. The JAK–STAT pathway accounts for 90% of IL-2–IL-2R signalling. IL-2 binding leads to heterodimerization of IL-2Rβ and IL-2Rγ, activating the tyrosine kinases JAK1 and JAK3, respectively, which phosphorylate tyrosine residues in IL-2Rβ. This promotes recruitment of signalling molecules such as PI3K, STAT5 or SHC1, which are phosphorylated by JAKs, resulting in specific pathway activation, nuclear translocation of transcription factors and finally targeted transcription regulation that induces cell activation, differentiation and proliferation. PI3K phosphorylates phosphatidylinositol 4,5-bisphosphate (PIP2), resulting in production of phosphatidylinositol-3,4,5-trisphosphate (PIP3), which promotes recruitment of phosphoinositide-dependent kinase l (PDK1) and AKT (also known as PKB) to the cell membrane. Phosphorylation of AKT by PDK1 and mTOR complex 2 (mTORC2) is necessary for full activation. AKT phosphorylation of tuberous sclerosis complex (TSC) proteins relieves TSC-mediated inhibition of RHEB (not shown) to activate mTORC1, which phosphorylates p70 ribosomal S6 kinase (p70S6K), a kinase that is important for survival, proliferation and protein translation. Tyrosine phosphorylation of STAT5 leads to its dimerization or tetramerization, nuclear translocation and transcription activation or repression. Phosphorylation of SHC1 promotes recruitment of GRB2 and SOS, forming a complex that catalyses GTP exchange on RAS and subsequent activation of the MAPK pathway. Depending on the concentration and duration of exposure, IL-2 induces different signals in conventional T cells compared with T_{reg} cells, which influences the outcome of a localized immune response in a pro-inflammatory setting. Aside from natural IL-2, enhanced IL-2 formulations such as muteins or IL-2–anti-IL-2 antibody complexes can be targeted to T_{reg} cells or conventional T cells in autoimmune or cancer settings, respectively, and, depending on modified binding properties, induce stronger IL-2 signalling. BLIMP1, B lymphocyte-induced maturation protein 1; PP2A, protein phosphatase 2A; T_{eff} cells, effector T cells.

IL-2 and T_{reg} cells. IL-2 is required for the differentiation, immunosuppressive function, homeostasis and survival of T_{reg} cells[6,23,39,41]. T_{reg} cells can suppress or modulate autoreactive T cells and other immune cells, such as antigen-presenting cells, through several mechanisms: secretion of anti-inflammatory cytokines (such as TGFβ, IL-10 and IL-35); stimulation of dendritic cells following transendocytosis of CD80/86, resulting in the production of the immunosuppressive enzyme indoleamine 2,3-dioxygenase[51,52]; conversion of ATP to adenosine by CD39 and CD73 (REF.[53]) consumption of local IL-2 by the trimeric IL-2R, thereby making it unavailable for activating conventional T cells and NK cells[54,55]; and direct cytotoxicity against $CD8^+$ T cells and NK cells[56,57].

Besides IL-2Rα, IL-2 stimulation leads to upregulated expression of numerous molecules that are important for T_{reg} cell function, including CTLA4, PD1, TIM3, LAG3, CD39 and TNF receptor superfamily members such as GITR, OX40 and TNF receptor 2 (TNFR2), all of which result in improved immunosuppressive capacity of T_{reg} cells[7,34,36,58–60]. Parallel upregulation of the genes encoding the anti-apoptotic molecules BCL-2 (REF.[34]) and FOXP1 might enhance FOXP3 binding to the promoters of target genes[61].

T_{reg} cells express high levels of TNFR2, and TNFR2 deficiency leads to a T_H17 cell-like phenotype, underlining the importance of TNFR2 for maintaining T_{reg} cell identity[62]. $TNFR2^+$ T_{reg} cells comprise ~40% of peripheral



T_{reg} cells in mice and have increased immunosuppressive functions[63,64]. In addition, $CD38^{high}$ T_{reg} cells seem to have increased immunosuppressive capacity in mice with experimental autoimmune encephalomyelitis (EAE)[65] and in patients with multiple myeloma[66].

In vitro assays measuring pSTAT5 phosphorylation during IL-2 signalling indicate that memory T_{reg} cells (defined as $CD4^+CD25^{hi}FoxP3^{hi}CD44^{hi}CD45RO^{hi}CD27^{hi}BLC2^{hi}CCR7^{low}$ T cells) are at least 20-fold more sensitive to IL-2 than effector memory $CD4^+$ T cells (defined as $CD4^+CD44^{hi}CD45RO^{hi}CD45RA^-CD127^{hi}CCR7^-$ T cells) or $CD56^{hi}$ NK cells, and 40-fold more sensitive than memory $CD8^+$ T cells. Furthermore, when examining the gene activation programme downstream of STAT5 phosphorylation, T_{reg} cells are at least 100-fold more sensitive to IL-2 than memory $CD4^+$ T cells[34]. This supports the unique ability of T_{reg} cells to be effectively modulated by low amounts of IL-2 via their high affinity trimeric IL-2R.

IL-2 and T_H17 cell differentiation. T_H17 cells are crucial for host defence against pathogens[67,68], and although they protect the intestine against inflammation, they are also involved in the pathogenesis of autoimmune diseases such as SLE and psoriasis[69]. T_H17 cell differentiation is induced by TGFβ, IL-21 and IL-6 through STAT3-mediated IL-23 signalling, which increases the expression of the nuclear receptor retinoic acid-related orphan receptor-γt (RORγt; encoded by *RORC*)[23,67]. IL-2 inhibits T_H17 cell differentiation by several mechanisms, including STAT5-mediated repression of RORγt expression, competition between STAT5 and STAT3 at the *Il17a* locus (reducing IL-17 production) and repression of the IL-6 receptor genes[70–72]. Thus, it is important to highlight that IL-2 is important not only for the activity of T_{reg} cells, and thereby regulation of conventional T cells, but also in inhibiting differentiation of T_H17 cells.

IL-2 and follicular T cells. T follicular helper (T_{FH}) cells foster the proliferation, survival and differentiation of germinal centre B cells by producing various cytokines, including IL-4, IL-9, IL-10 and IL-21, or by CD40 ligand co-stimulation. IL-2 inhibits T_{FH} cells at an early stage of their differentiation[73,74]. The effect of IL-2 on T_{FH} cells is mediated by downregulation of BCL-6 expression, repression of genes important for T_{FH} cell differentiation[75] and activation of AKT and mTORC1 signalling, which leads to BLIMP1 and T-bet expression[76], all of which result T_{FH} phenotype repression.

The origin of T follicular regulatory (T_{FR}) cells is not fully elucidated. Upon activation of naive T_{reg} cells that are $CXCR5^-$, a population of $CD25^+$ T_{FR} cells expressing intermediate levels of CXCR5 emerges. These T_{FR} cells can then upregulate CXCR5, which mediates their routing to the lymph node germinal centre (GCs)[77]. Terminally differentiated T_{FR} cells within the GCs do not express IL-2Rα[78]. The biology of T_{FR} cells is also less well understood than that of T_{FH} cells, but T_{FR} cells seem to inhibit the T_{FH} cell–B cell interaction by several mechanisms, including CTLA4 expression, reduction of glycolysis in B cells, production of inhibitory cytokines such as IL-10, TGFβ and granzyme B, and possibly the consumption of IL-1 through their expression of the IL-1 inhibitory receptors IL-1R2 and IL-1RA[78,79]. Altogether, because of the regulated CD25 expression on T_{FR} cells and the inhibition of T_{FH} differentiation, IL-2 should probably modulate antibody production. This effect, which could be important in the context of antibody-mediated autoimmunity, has not yet been studied in detail.

T_{reg} cells, IL-2 and tissue regeneration. In non-lymphoid tissues, T_{reg} cells have an important role in tissue regeneration. The gene expression profile of tissue-resident T_{reg} cells depends on the host tissue and environmental factors[80–82]. T_{reg} cells promote tissue regeneration in the intestine, skin, skeletal muscle, visceral adipose tissue, central nervous system, lung, liver and placenta, and protect and facilitate the stem cell niche in the bone marrow[1,83–87]. The characteristics of T_{reg} cell function in tissue repair vary depending on the context of the tissue and involve increased signalling through the IL-33–IL-1 receptor-like 1 (IL-1RL1; also known as ST2) axis (which enhances the tissue-protective ability of T_{reg} cells in mice) and production of the anti-inflammatory cytokine IL-10 and/or the T_{reg} cell effector amphiregulin (encoded by *AREG*)[88]. For example, T_{reg} cells account for up to 50% of $TCR\beta^+CD4^+$ T cells that are recruited to the muscle following injury[89]. In a mouse model of Duchenne muscular dystrophy, depletion of T_{reg} cells exacerbated muscle injury and inflammation, which was reversed by IL-2 treatment (in the form of an IL-2–anti-IL-2 antibody complex, see below)[90]. Specific deletion of *AREG* in T_{reg} cells leads to lung fibrosis after influenza virus infection, further highlighting the role of T_{reg} cells in tissue regeneration and homeostasis[91–93].

T_{reg} cells, IL-2 and metabolism. T_{reg} cells meet their metabolic requirements by fatty acid and pyruvate oxidation[47], whereas conventional T cells mainly use glycolysis[94]. However, mTORC1-driven glycolysis is required during T_{reg} cell activation and proliferation, which is dependent on IL-2. Because mTORC1 inhibits *FOXP3* expression, the suppressive capacity of T_{reg} cells is reduced temporarily[95,96]. During T_{reg} cell induction and maintenance, enhanced *FOXP3* expression reprogrammes peripheral T_{reg} cells to preferentially metabolize fatty acids[97,98]. This metabolic switch enables T_{reg} cells to proliferate in the presence of low levels of local leptin from adipocytes or short-chain fatty acids in the intestine[95,99]. Distinct metabolic pathways can influence the quality and extent of T cell responses, with glycolysis favouring conventional T cell generation and fatty acid oxidation and pyruvate oxidation favouring T_{reg} cell generation. Inhibition of certain metabolic enzymes could limit autoimmunity.

T_{reg} cells in the pro-inflammatory environment. While T_{reg} cells control the inflammatory response, they become less efficient within inflamed tissues[100] in which T_{reg} cells may even become unstable by losing *FOXP3* expression and converting to a phenotype that is more characteristic of conventional $CD4^+$ T cells; they are

then referred to as 'ex-T_{reg}' cells. Pro-inflammatory conditions defined by elevated IL-6 or TNF levels, as in the inflamed synovial tissue, inhibit the immunosuppressive function of T_{reg} cells. However, TNF has been claimed to enhance the immunosuppressive capacity of T_{reg} cells by signalling through TNFR2 (REF.[101]). In addition, T_{reg} cells positive for Helios, a marker of thymic T_{reg} cells, migrate to inflamed tissue and suppress inflammation in active SLE[102,103].

IL-2 and T_{reg} cells in rheumatic diseases

IL-2 and T_{reg} cells in RA. RA is a chronic, inflammatory, destructive joint disease[104], and autoantibodies are present in two-thirds of patients (seropositive RA)[105]. The inflamed synovium in RA contains increased numbers of synoviocytes[106] and $CD4^+$ memory T cells and diffuse or ectopic germinal centres with on-site affinity maturation and antibody production[107,108] (FIG. 2).

The immunosuppressive capacity and number of T_{reg} cells in patients with RA is controversial, as there are reports that these are decreased, normal or increased[62]. RA, like SLE and type 1 diabetes mellitus (T1DM), is associated with low IL-2 production that affects T_{reg} cell fitness[5,34,109–112]. Polymorphisms in *FOXP3* and *CTLA4* and upregulated expression of pro-inflammatory cytokines, including TNF, IL-6 and IL-17, have been reported in association with reduced T_{reg} cell activity or numbers, which correlate with increased disease activity[113]. Furthermore, direct killing of T_{reg} cells by autoreactive antigen-specific $CD8^+$ T cells in the synovial membrane has been reported[114]. In addition, the abundance of high-avidity $CD8^+$ T cells correlates with failure to respond to treatment, whereas low-avidity, TNF-producing $CD8^+$ T cells are present in patients with RA who respond to treatment with TNF inhibitors[114]. Diminished function and numbers of T_{reg} cells in RA could participate in disease pathogenesis.

IL-2 and T_{reg} cells in SLE. SLE is a heterogeneous, complex autoimmune disease with multifactorial aetiology, characterized by generalized loss of immune tolerance leading to autoantibody production and inflammation of multiple organs[115]. Autoimmunity in SLE has been linked to T cell and B cell abnormalities that are induced by environmental and hormonal triggers in genetically susceptible individuals[116] (FIG. 3).

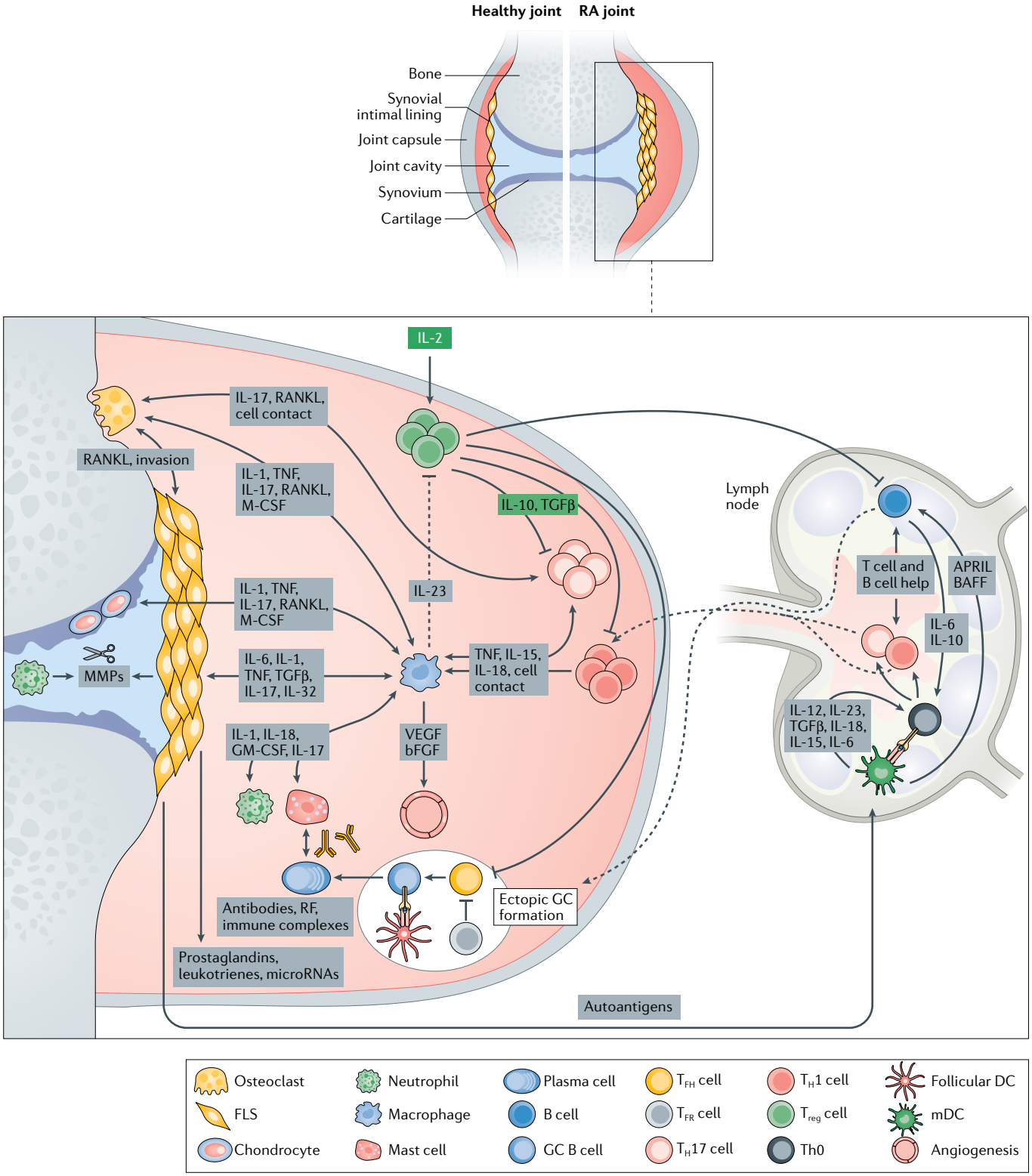
The role of T_{reg} cells and IL-2 in SLE has been extensively reviewed elsewhere[117–119]. The number and functional capacity of T_{reg} cells in SLE is unclear, with conflicting reports of these cells being impaired or normal, probably owing to diverse definitions of T_{reg} cells, different gating strategies in flow cytometry analysis and inclusion of activated $IL\text{-}2R\alpha^+CD4^+$ T cells[62,119]. Indeed, as a result of low IL-2 availability, T_{reg} cells from patients with SLE express lower to undetectable levels of IL-2Rα, which can be re-induced by in vitro stimulation with IL-2 (REFS[117–119]). This low IL-2Rα expression correlates with disease activity and circulating anti-dsDNA antibody levels, underlining its clinical relevance[118,119]. In agreement with these findings, T_{reg} cells from IL-2-deficient mice show significantly reduced IL-2Rα levels[7]. In this context of T_{reg} cell insufficiency, autoantibody production is driven by nuclear antigens that are released through cell apoptosis and stimulation of Toll-like receptors (TLRs) and the subsequent production of type I interferon[120]. Pathogenic differentiation of conventional T cells is mediated by abnormal TCR signalling, PI3K–Akt–mTOR pathway (including downstream molecules such as decreased PP2A, activated CaMK4, increased CD44, increased ROCK and decreased SRSF1) and JAK–STAT pathway activation, resulting in abnormal production of cytokines (increased production of pro-inflammatory cytokines such as interferons or IL-23/IL-17 and decreased production anti-inflammatory or regulatory cytokines such as IL-2)[121], and alterations in immune metabolism[95].

IL-2 production and signalling have been shown to be impaired in patients with SLE and mouse models of SLE[110,122–126] at the transcription level because of increased binding of cyclic AMP response element modulator-α (CREMα) to the *Il2* promoter (enabled by CaMK4 activation[127]), and decreased levels of SRSF1 (REF.[128]), NF-κB and AP-1 (REF.[129]). Decreased expression of IL-2Rα and STAT5 on T_{reg} cells in SLE also impairs IL-2 signalling[122,130].

IL-2 and T_{reg} cells in other rheumatic diseases. Osteoarthritis is a T_H1 cell-mediated low-grade, chronic inflammatory disease of the cartilage that is similar to but less severe than RA[131]. Increased numbers of synovial membrane T_{reg} cells, with high expression of CTLA4, PD1 and GITR compared with peripheral T_{reg} cells[132], have been reported in osteoarthritis. In addition, an increased abundance of $TIM3^-$ T_{reg} cells is associated with lower production of IL-10 and advanced stages of osteoarthritis[131].

Inflammatory spondyloarthropathies include ankylosing spondylitis[133] and psoriatic arthritis[134]. In ankylosing spondylitis, IL-2 signalling and the functional capacity and number of peripheral blood T_{reg} cells are decreased[24], whereas elevated T_{reg} cell numbers in the synovial fluid correlate with disease remission[135]. In juvenile idiopathic arthritis, a heterogeneous group of paediatric autoimmune arthritides, impaired immunosuppressive function of peripheral blood and synovial T_{reg} cells has been reported[26,136].

pSS is a systemic autoimmune epithelitis of exocrine glands, primarily the salivary and lacrimal glands. In pSS, elevated TLR3 expression in salivary glands, autoantigen presentation by apoptotic epithelial cells, pro-inflammatory cytokine production and differentiation of T_H1 cells, T_H17 cells and T_{FH} cells, and subsequent production of antibodies against SSA/Ro and/or SSB/La by B cells, causes epithelial hypofunction that presents clinically as sicca (dryness) of the eyes and mouth, pain and fatigue[137]. T_{reg} cell number but not function is reduced in salivary glands and peripheral blood of patients with pSS[138,139], along with reduced serum IL-2 levels[140]. In addition, elevated circulating levels of $Helios^+$ T_{reg} cells inversely correlate with IgG and IgM levels and are significantly higher ($P<0.05$) in anti-SSB/La seronegative patients, suggesting that these T_{reg} cells are important for suppression of autoantibody production in B cells[141]. Elevated levels of circulating



$CD4^+CD25^-GITR^+$ suppressive T cells in patients with inactive pSS have been interpreted as revealing a novel T_{reg} cell population[142]. Studies using only IL-$2R\alpha^+$ T cells to delineate T_{reg} cells are difficult to interpret because they most probably include activated T cells and exclude $CD25^-$ T_{reg} cells[139].

SSc is characterized by immune dysregulation, vasculopathy and fibrosis. TLR-mediated innate immunity drives a T_H2 cell-mediated immune response that involves production of IL-4, IL-5 and IL-13, simulating pro-fibrotic M2 monocytes/macrophages to secrete TGFβ[143]. In addition, T_H17 cells lead to increased collagen synthesis as well as $CD8^+$ T cells to increased endothelial cell injury, where the latter induces antigen-promoted B cell differentiation[28]. T_{reg} cells in SSc are functionally impaired, with suggested mechanisms

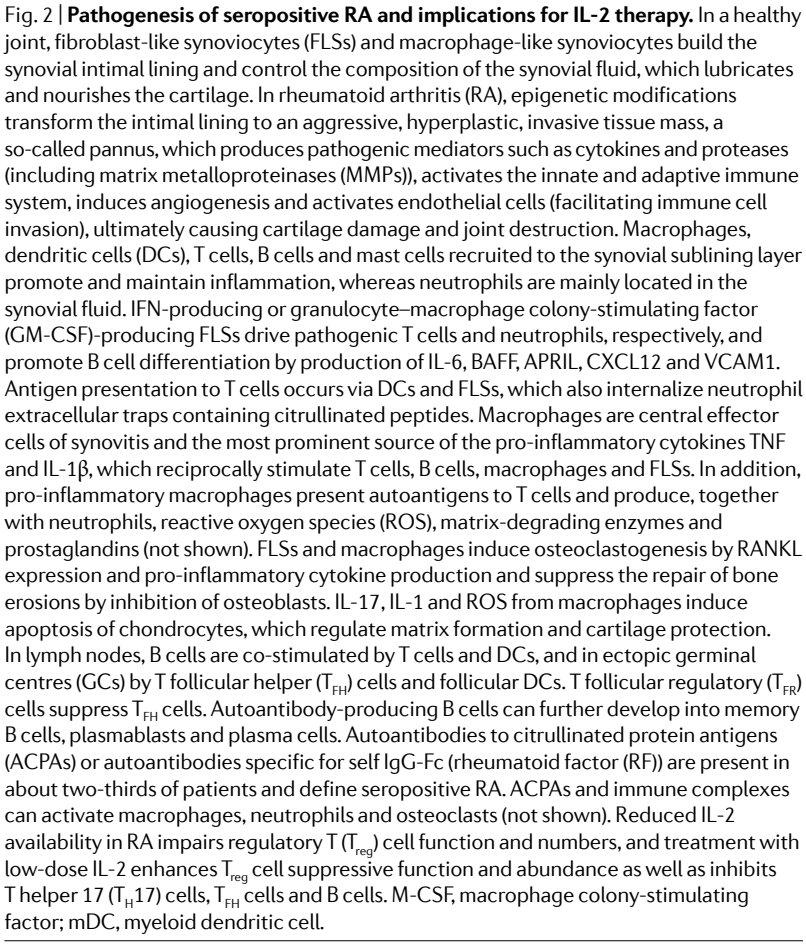


Fig. 2 | **Pathogenesis of seropositive RA and implications for IL-2 therapy.** In a healthy joint, fibroblast-like synoviocytes (FLSs) and macrophage-like synoviocytes build the synovial intimal lining and control the composition of the synovial fluid, which lubricates and nourishes the cartilage. In rheumatoid arthritis (RA), epigenetic modifications transform the intimal lining to an aggressive, hyperplastic, invasive tissue mass, a so-called pannus, which produces pathogenic mediators such as cytokines and proteases (including matrix metalloproteinases (MMPs)), activates the innate and adaptive immune system, induces angiogenesis and activates endothelial cells (facilitating immune cell invasion), ultimately causing cartilage damage and joint destruction. Macrophages, dendritic cells (DCs), T cells, B cells and mast cells recruited to the synovial sublining layer promote and maintain inflammation, whereas neutrophils are mainly located in the synovial fluid. IFN-producing or granulocyte–macrophage colony-stimulating factor (GM-CSF)-producing FLSs drive pathogenic T cells and neutrophils, respectively, and promote B cell differentiation by production of IL-6, BAFF, APRIL, CXCL12 and VCAM1. Antigen presentation to T cells occurs via DCs and FLSs, which also internalize neutrophil extracellular traps containing citrullinated peptides. Macrophages are central effector cells of synovitis and the most prominent source of the pro-inflammatory cytokines TNF and IL-1β, which reciprocally stimulate T cells, B cells, macrophages and FLSs. In addition, pro-inflammatory macrophages present autoantigens to T cells and produce, together with neutrophils, reactive oxygen species (ROS), matrix-degrading enzymes and prostaglandins (not shown). FLSs and macrophages induce osteoclastogenesis by RANKL expression and pro-inflammatory cytokine production and suppress the repair of bone erosions by inhibition of osteoblasts. IL-17, IL-1 and ROS from macrophages induce apoptosis of chondrocytes, which regulate matrix formation and cartilage protection. In lymph nodes, B cells are co-stimulated by T cells and DCs, and in ectopic germinal centres (GCs) by T follicular helper (T_{FH}) cells and follicular DCs. T follicular regulatory (T_{FR}) cells suppress T_{FH} cells. Autoantibody-producing B cells can further develop into memory B cells, plasmablasts and plasma cells. Autoantibodies to citrullinated protein antigens (ACPAs) or autoantibodies specific for self IgG-Fc (rheumatoid factor (RF)) are present in about two-thirds of patients and define seropositive RA. ACPAs and immune complexes can activate macrophages, neutrophils and osteoclasts (not shown). Reduced IL-2 availability in RA impairs regulatory T (T_{reg}) cell function and numbers, and treatment with low-dose IL-2 enhances T_{reg} cell suppressive function and abundance as well as inhibits T helper 17 (T_H17) cells, T_{FH} cells and B cells. M-CSF, macrophage colony-stimulating factor; mDC, myeloid dendritic cell.

including decreased *FOXP3* expression (by either hypermethylation of the *FOXP3* promoter or skewed X chromosome inactivation (the site of the *FOXP3* locus)), decreased expression of *RUNX1* (which controls *FOXP3* expression), low serum levels of semaphorin 3A (which is important for T_{reg} cell maintenance) and increased plasticity of T_{reg} cells towards a T_H17 cell phenotype[27,28]. Daily low-dose IL-2 administration in patients with chronic graft-versus-host disease (GVHD) improved fibrotic and sclerotic skin manifestations, attesting to the ability of T_{reg} cells to repair tissues[144,145]. The importance of T_{reg} cell impairment in sclerosing diseases is supported by a report of improved fibrotic and sclerotic skin manifestations following daily low-dose IL-2 administration in patients with chronic GVHD.

Impaired T_{reg} cell function has been reported in patients with different forms of small-vessel and large-vessel vasculitides including granulomatosis with polyangiitis[146], eosinophilic granulomatosis with polyangiitis[147,148], hepatitis C virus (HCV)-induced vasculitis[149,150], polyarteritis nodosa[151], Behçet disease[152], Kawasaki syndrome[153], Takayasu arteritis[154], giant cell arteritis and polymyalgia rheumatica[155].

Sarcoidosis is a systemic granulomatous disease characterized by granuloma formation in response to an unknown trigger, which induces an innate immune response (activation of dendritic cells, macrophages or epithelioid cells) and an adaptive immune response mediated by T_H1 cells and T_H17 cells. Although the number of T_{reg} cells in the circulation is elevated in individuals with sarcoidosis[156], the function[29] and survival[157] of these cells are impaired. Elevated TNFR2$^+$ T_{reg} cell numbers and soluble TNFR2 levels correlate with response to infliximab treatment in patients with sarcoidosis[156,158].

Gout is an innate immunity-driven, NLRP3 inflammasome-mediated intermittent inflammatory arthritis that develops in response to the formation of monosodium urate crystals[159]. Decreased T_{reg} cell to T_H17 cell ratios have been observed in rats induced to develop gout arthritis[160]. The current evidence of T_{reg} cell impairment in rheumatological autoimmune diseases summarized here provides a rationale for T_{reg} cell-inducing treatments.

IL-2 therapy in mouse disease models

Low-dose IL-2 treatment of mice corresponds to administration of ≤50,000 international units (IU) daily[39,161]. At these doses, treatment of C57BL/6 mice leads to a dose-dependent increase in T_{reg} cell number and an improvement in the conventional T cell–T_{reg} cell balance with an early recirculation of blood T_{reg} cells in secondary lymphoid organs[162]. Of note, long-term (up to 1 year) IL-2 treatment, accomplished by infection of mice with a recombinant adeno-associated viral vector (rAAV), increased and activated T_{reg} cells without impairing immune responses to infections, vaccination or cancer[163]. These data suggest that long-term administration of IL-2 is not associated with impaired immune responses, at least in mice.

IL-2 and T_{reg} cell-directed IL-2–anti-IL-2 antibody complexes have been extensively studied in various animal models, including autoimmunity, allergy, infection, transplantation and cancer[5,164], as well as rheumatic diseases. In lupus-prone NZB/NZW F1 mice, an acquired deficiency of IL-2 and IL-2-producing CD4$^+$ T cells leads to hyperactivity of conventional T cells as well as an imbalance of T_{reg} cells and conventional T cells (that is, a low T_{reg} cell to conventional T cell ratio), which is associated with disease progression. T_{reg} cells in this model are functionally intact; treatment with low-dose IL-2 reverses this homeostatic T_{reg} cell dysregulation by increasing T_{reg} cell numbers in the blood, lymphoid organs and kidneys, leading to prolonged survival and reduced kidney injury[126,165]. Incidentally, this observation highlights that IL-2 treatment is not reserved solely for settings in which there is a T_{reg} cell deficiency. Improvement of lupus symptoms has been shown in two additional mouse models, NZB/W F1 mice and MRL/*lpr* mice[117,118,166–168].

Administration of IL-2–anti-IL-2 antibody complex suppressed collagen-induced arthritis, increased peripheral and synovial T_{reg} cell abundance and reduced the production of pro-inflammatory cytokines[169]. Infectious arthritis, which leads to rapid and severe destruction of the joint, is exacerbated by pathogenic CD4$^+$ T cells and T_H17 cells. In mice, IL-2-induced T_{reg} cells reduces septic arthritis severity and systemic inflammation, while preserving host immune defence[170,171]. In other mouse models of autoimmunity, including EAE[172], T1DM[173], sclerosing cholangitis[174], IL-2 or IL-2–anti-IL-2 antibody

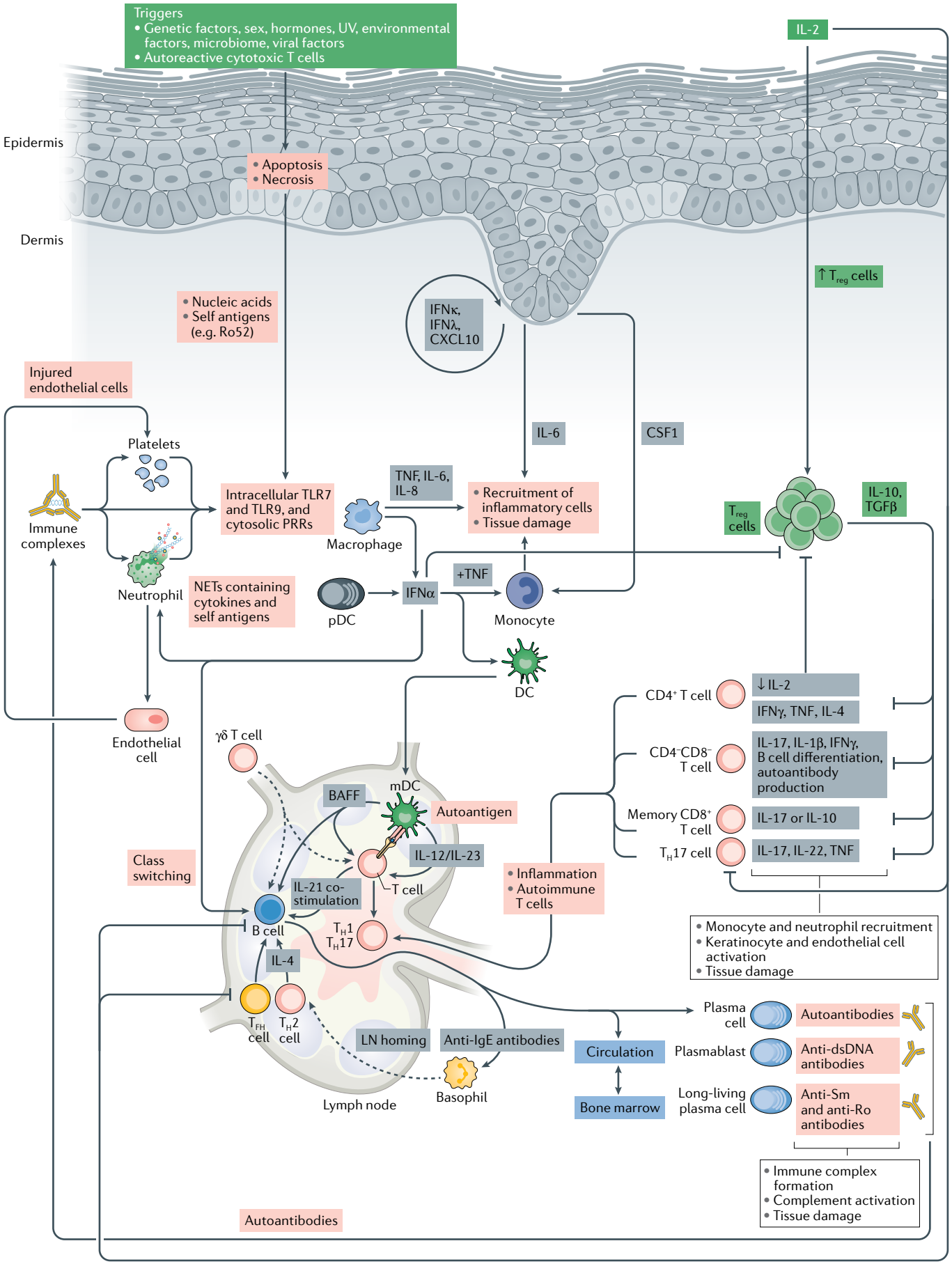

Triggers
• Genetic factors, sex, hormones, UV, environmental factors, microbiome, viral factors
• Autoreactive cytotoxic T cells
IL-2
Epidermis
Dermis
• Apoptosis
• Necrosis
↑ T_{reg} cells
• Nucleic acids
• Self antigens (e.g. Ro52)
IFNκ, IFNλ, CXCL10
Injured endothelial cells
Platelets
IL-6
CSF1
Immune complexes
Intracellular TLR7 and TLR9, and cytosolic PRRs
TNF, IL-6, IL-8
Macrophage
• Recruitment of inflammatory cells
• Tissue damage
T_{reg} cells
IL-10, TGFβ
Neutrophil
NETs containing cytokines and self antigens
pDC
IFNα
+TNF
Monocyte
DC
Endothelial cell
γδ T cell
BAFF
mDC
Autoantigen
IL-12/IL-23
T cell
Class switching
IL-21 co-stimulation
B cell
T_H1 T_H17
IL-4
T_{FH} cell
T_H2 cell
LN homing
Anti-IgE antibodies
Basophil
Lymph node
CD4⁺ T cell
↓ IL-2
IFNγ, TNF, IL-4
CD4⁻CD8⁻ T cell
IL-17, IL-1β, IFNγ, B cell differentiation, autoantibody production
Memory CD8⁺ T cell
IL-17 or IL-10
T_H17 cell
IL-17, IL-22, TNF
• Inflammation
• Autoimmune T cells
• Monocyte and neutrophil recruitment
• Keratinocyte and endothelial cell activation
• Tissue damage
Circulation
Bone marrow
Plasma cell
Plasmablast
Long-living plasma cell
Autoantibodies
Anti-dsDNA antibodies
Anti-Sm and anti-Ro antibodies
• Immune complex formation
• Complement activation
• Tissue damage
Autoantibodies

Fig. 3 | **Pathogenesis of SLE and implications for IL-2 therapy.** Certain triggers in genetically predisposed individuals lead to keratinocyte cell death, which results in the release of autoantigens that are recognized by pattern recognition receptors (PRRs) such as the membrane-bound Toll-like receptors (TLRs; in particular, intracellular TLR3, TLR7 and TLR9) or TLR-independent cytoplasmic nucleic-acid-detecting PRRs (such as NOD receptors, RIG1, MDA5 or cGAS–STING) in plasmacytoid dendritic cells (pDCs) or macrophages. In addition, immune complexes, formed by FcγRIIa-mediated endocytosis, and self-antigens and cytokines released during increased NETosis (a form of neutrophil cell death involving release of neutrophil extracellular traps (NETs)), can activate PRRs in pDCs, which are the main producers of interferon-α (IFNα). Macrophages and IFNα-activated monocytes are pivotal in organ damage. Activated autoantigen-presenting dendritic cells (DCs) induce T cell differentiation and autoantibody production, drive T follicular helper (T_{FH}) cell differentiation, promote B cell help and impair regulatory T (T_{reg}) cell function. Furthermore, direct IFNα or TLR stimulation, cytokines such as BAFF from DCs or IL-21 from T_{FH} cells or T cell costimulation, induce B cell maturation and expansion, plasma cell generation and production of high-affinity autoantibodies against IgG, IgA and IgE. Immune complex generation, consisting of self-antigens and the corresponding autoantibodies, as well as activation of basophils and pDCs by autoantibodies against IgE, B cells and neutrophils by nucleic acids and self-antigens, and keratinocytes by autocrine type I and type III interferons (especially IFNκ and IFNλ) or CXCL10-induced IL-6 production, act as feedback loops that further drive hyperactivation of innate immune pathways. T_{reg} cells are reduced in number or are functionally impaired in systemic lupus erythematosus but, in addition, they are affected by low IL-2 availability and direct inhibition by IFNα. Low-dose IL-2 treatment leads to T_{reg} cell expansion and direct inhibition of B cells, T_{FH} cells and T helper 17 (T_H17) cells, thereby leading to control of autoimmunity. dsDNA, double-stranded DNA; LN, lymph node; mDC, myeloid dendritic cell.

complex, administration leads to increased T_{reg} cell abundance, resulting in partial or complete responses.

IL-2 therapy in rheumatic diseases

The clinical use of IL-2 was first reported in 1984 as high-dose treatment with the aim of stimulating conventional T cell activity against tumour cells[175,176], which was followed by studies that led to FDA approval of high-dose IL-2 for treatment of metastatic renal cell cancer and later for metastatic melanoma. At high doses, IL-2 led to a durable complete response in 5–7% of patients, but also caused severe toxic adverse effects and a major increase in T_{reg} cell number, which most likely limited its therapeutic efficacy by a contra-productive regulatory activity of T_{reg} cells against the previously activated conventional T cells[5,176]. In the early 1990s, different *Il2* or *Il2R* knockout mice revealed severe autoimmunity rather than the expected immunodeficiency[177]. This 'IL-2 paradox' was explained by a T_{reg} cell deficiency[178]. The relevance of IL-2 in autoimmunity became clear but the therapeutic potential of this cytokine in autoimmunity was blunted by its pleiotropic effects and the risk of stimulating dangerous (tissue-damaging) effector T cells. More than 10 years later, in 2006, researchers reasoned that because T_{reg} cells constitutively express the high-affinity IL-2 receptor, using low-dose IL-2 in human autoimmune diseases might serve as means to stimulate T_{reg} cell activity preferentially without activating effector T cells[5,150]. This proved to be the case and it is now established that at doses of 300,000–3,000,000 IU daily in humans, IL-2 stimulates mainly T_{reg} cells and is well tolerated[5,179,180]. As T_{reg} cell insufficiency is central to the pathogenesis of most autoimmune diseases, low-dose IL-2 may have great therapeutic potential and broad clinical applicability. However, several factors can influence the efficacy of low-dose IL-2 in different autoimmune diseases, including defects in IL-2 signalling, inflammation and bystander activation of conventional T cells, NK cells, ILC2s and eosinophils, and inappropriate resistance of these immune cells to the suppressive effect of T_{reg} cells in the inflammatory context[179].

Safety. Low-dose IL-2 is well tolerated and adverse events are mostly dose-dependent, mild-to-moderate and transient. In several studies, the most frequently reported adverse events were injection site reactions (which are common to many injectable biotherapies), transient fever, influenza-like symptoms, myalgia and nausea[150,179,181,182]. At the highest end of the low-dose range (3 MIU daily), adverse events such as chills, influenza-like symptoms, headaches, dizziness, arthralgia and myalgia occur with increased frequency and severity[145,181]. A recent meta-analysis of low-dose IL-2 safety confirmed these results[180]. Of note, 1-year treatment with low-dose IL-2 was well tolerated in children (6–12 years of age) with recent-onset T1DM[183]. To date, no induction of new or aggravation of pre-existing autoimmune syndromes has been observed in people treated with low-dose IL-2. Thyroiditis, which has been reported in people treated with high-dose IL-2, is rarely observed with low-dose IL-2 and is reversible after treatment cessation[184]. A transient increase in eosinophils, fibrinogen and D-dimer has been reported with low-dose IL-2, without relevant alterations in blood counts, liver enzymes, renal parameters, coagulation, plasma proteins and immunoglobulins[179,181]. While anti-IL-2 antibodies have been described in patients with cancer who receive high-dose IL-2, bona fide anti-IL-2 antibodies have not been described after low-dose IL-2 treatment[112,179,185].

In contrast to current immunosuppressive treatments for autoimmune diseases, low-dose IL-2 treatment is not associated with serious infections. In fact, in SLE, lower infection rates were recorded in the IL-2-treated group compared with the placebo group (6.9% versus 20%)[182]. Moreover, there was no increase in HCV viral loads in HCV-induced vasculitis[150,182]. These lower infection rates might be explained by improved virus-specific $CD8^+$ T cell responses[186] and/or enhanced NK cell activity with increased expression of IFNγ, NKp46 and NKG2D[182]. Of note, long-term administration of low-dose IL-2 in mice does not inhibit immune responses to vaccination and subsequent infection, nor does it affect cancer occurrence and growth[163,187].

Clinical efficacy. The first trial of low-dose IL-2 in people with an autoimmune disease showed clinical improvement in eight of ten patients with HCV-induced vasculitis[150], whereas in three of three patients arthralgia disappeared. In the TRANSREG study in 46 patients with any of 11 different autoimmune diseases, low-dose IL-2 triggered a remarkably similar and specific increase in T_{reg} cell number[179]. The clinical outcome was evaluated using a Clinical Global Impression tool[188], which recorded a significant improvement ($P<0.001$) during the 6 months of treatment and at 2 months of follow-up. Disease-specific scores, such as the Bath Ankylosing Spondylitis Disease Activity Index for ankylosing

spondylitis, the SLE Disease Activity Index (SLEDAI) for SLE and the Psoriasis Area Severity Index for psoriasis, improved, along with patient-reported arthralgia and fatigue[179]. Similar overall results have now been obtained in a further 80 patients in the TRANSREG study (D.K., unpublished observations).

Low-dose IL-2 has been extensively investigated in patients with SLE and in multiple studies has produced clinical benefit[110,181,182,189]. First, low-dose IL-2 was remarkably efficient in the treatment of a single patient with severe refractory SLE[190]. Subsequently, four open label studies[110,181,189,191], one single-site randomized controlled trial (RCT)[182] and a multicentre RCT (D.K., unpublished observations) were conducted. Most studies (aside from one study[110], in which no clinical assessment was performed) showed a clinically significant improvement, as shown by SLEDAI[181,189], as well as a reduction in concomitant prednisone use by ≥50% in 44% of patients receiving prednisone alone versus 67% of patients treated with low-dose IL-2 (REFS[182,189]). In the single-site RCT, SLE Responder Index 4 (SRI-4) at 12 weeks was not significantly different between treatment and placebo groups ($P = 0.052$), which means that the primary end point was not met, but showed significance ($P < 0.05$) at 6, 8, 10, 16 and 24 weeks[182]. In this trial, in patients with lupus nephritis (13/30 in the low-dose IL-2 group and 12/30 in the placebo group), complete remission was achieved in 53.85% of patients (7/13) versus 16.67% of patients (2/12), respectively ($P = 0.036$). In addition, in patients with lupus nephritis, increased serum IL-2 levels at week 10 correlated with significantly higher remission rates (61% versus 9%, $P = 0.041$) and a significant reduction in 24-h urine protein ($P = 0.002$)[192]. In this study, the decrease in renal involvement in the low-dose IL-2 and conventional treatment groups was significant (52% and 23%, respectively, $P = 0.025$)[192]. Rash, alopecia and arthralgia responded well to IL-2, decreasing significantly[182,189] (TABLE 1; an expanded version of this table is available in Supplementary Table 1).

Biological efficacy. In the TRANSREG study, an increase in T_{reg} cell abundance was observed in all patients and did not differ according to disease; after daily injections of 1 MIU IL-2 for 5 days, the mean increase was about twofold on day 8 and conventional T cells were not activated[179]. T_{reg} cells from patients with SLE, which showed reduced or undetectable IL-2Rα expression before IL-2 treatment, showed a marked upregulation of IL-2Rα T_{reg}[181].

Aside from T_{reg} cells, in all studies, NK cells and eosinophils were the main cell populations that responded to low-dose IL-2 (that is, increased). The subpopulation of NK cells that responded was $CD56^{hi}$ regulatory NK cells (that is, NK cells that produce cytokines but are not cytotoxic).

Interestingly, B cells also respond to IL-2. In the first study of low-dose IL-2 treatment of patients with HCV vasculitis, decreased numbers of B cells were found[150]. The $IgD^{+}CD27^{+}$ marginal-zone B cell subset was particularly affected, as has also been found in patients with SLE[181]. Similarly, in patients with T1DM treated with low-dose IL-2, the increase in T_{reg} cell number correlated with decreased B cell number[112]. These observations remain unexplained but are possibly favourable towards the achievement of clinical response in autoimmune diseases.

IL-2 therapy in other autoimmune diseases. In HCV-induced cryoglobulinaemic vasculitis, a 5-day course of 1.5 MIU IL-2 daily was followed by a 5-day course of 3 MIU daily at weeks 3, 6 and 9. Improvement in vasculitis occurred in eight of ten patients and disappearance of purpura in seven of seven patients and arthralgia in three of three patients. Vasculitis flares or increased HCV viraemia were not noted[150].

In other diseases (TABLE 1; Supplementary Table 1), such as pSS[193], ankylosing spondylitis[194], psoriatic arthritis[25] and polymyositis or dermatomyositis[195,196], single-course low-dose IL-2 regimens resulted in an increase in T_{reg} cell number. A single 5-day course of low-dose IL-2 daily in patients with pSS transiently restored the T_H17 cell–T_{reg} cell balance but did not result in clinical improvement[193]. By contrast, in psoriatic arthritis, a similar treatment led to clinical improvement in most activity measures, including tender or swollen joint counts and pain visual analogue scale scores[25]. Despite promising clinical results in a pilot study in alopecia areata[197], low-dose IL-2-dependent increases in T_{reg} cell abundance did not correlate with clinical efficacy in a larger randomized, placebo-controlled trial[198]. Two case reports showed efficacy of low-dose IL-2 in two of three patients with immune thrombocytopenia[199] and autoimmune hepatitis[200]. In two studies in patients with polymyositis or dermatomyositis, a single course of low-dose IL-2 in addition to conventional treatment induced reduction of muscle enzymes[195,196].

IL-2 in combination therapies

T_{reg} cell activation now seems to be a promising novel target for the treatment of autoimmune diseases. The first demonstration that low-dose IL-2 specifically activates T_{reg} cells kindled interest in the use of IL-2 in combination with drugs that target inflammatory pathways implicated in autoimmune disease pathogenesis. For example, biologics that block pro-inflammatory cytokines such as TNF, IL-1, IL-6, IL-17 and IL-23 can have synergistic effects by inhibiting inflammation and concomitantly improving T_{reg} cell immunosuppressive activity[5]. Time-controlled combination therapy of low-dose IL-2 with effector T cell-depleting treatments or immunosuppressive drugs also has clear potential as effector T cells get reduced while T_{reg} cells are boosted.

IL-2 and TNF blockers. T_{reg} cells have a decreased ability to suppress production of IFNγ and TNF in RA, and treatment with anti-TNF biologics increases their immunosuppressive function[201]. By binding to surface TNF on monocytes, the anti-TNF biologic adalimumab causes an upregulation of monocyte surface TNF, which signals through TNFR2 on T_{reg} cells[202–205], enhancing IL-2 responsiveness and T_{reg} cell stability and resulting in increased expression of IL-2Rα and FOXP3 in human T_{reg} cells[206,207] by *FOXP3* promoter hypomethylation[208,209]. Both IL-2 and TNF–TNFR2 signalling are necessary for

Table 1 | **Trials of low-dose IL-2 treatment in autoimmune diseases**

Study	Disease (*n*)	Type of study	IL-2 regimen	Clinical outcome	Biological outcome
Saadoun et al. (2011)[150]	HCV-induced cryoglobulinaemic vasculitis (10)	Prospective open-label, phase I/IIa	5-day cycle of 1.5 MIU daily, followed by 5-day cycle of 3 MIU daily at weeks 3, 6 and 9	Improvement in vasculitis with disappearance of purpura and arthralgia; partly improved kidney dysfunction; two patients with neuropathy only did not improve; no vasculitis flare or increased HCV viraemia	Clinical improvements coincided with increased T_{reg} cells; about threefold increase in T_{reg} cells, with potent suppressive activity; no activation of $CD4^+$ or $CD8^+$ conventional T cells
Rosenzwajg et al. (2019) (TRANSREG study)[179]	Mild-to-moderate RA (4), AS (10), SLE (6), psoriasis (5), Behçet disease (2), GPA (1), Takayasu arteritis (1), Crohn's disease (7), ulcerative colitis (4), autoimmune hepatitis (2) or sclerosing cholangitis (4)	Prospective, open-label, phase I/IIa	1 MIU daily for 5 days followed by two single injections weekly for 6 months	No disease flares; CGI improved significantly ($P<0.001$) at 3 and 6 months	About twofold increase in T_{reg} cells ($P<0.0001$) on day 8 without activation of conventional T cells
von Spee-Mayer et al. (2016)[110]	SLE (5)	Prospective, open-label	One 5-day cycle of 1.5 MIU IL-2 daily	Not analysed	Increased T_{reg} cells; slight increase in $CD25^+$ T cells, $CD8^+$ T cells, NKT cells and NK cells
He et al. (2016)[189]	SLE (40)	Prospective, open-label	Three cycles of 1 MIU IL-2 every other day for 2 weeks followed by a 2-week break in treatment	38 patients completed the study; SRI-4 increased 2.8-fold from week 2 to week 12; SLEDAI significantly reduced ($P<0.001$); 25 patients (67.6%) reduced prednisone use by ≥50% at week 12	Significant increase in T_{reg} cell number and suppressive function; significant decrease in T_{FH} cells, T_H17 cells, $CD4^-CD8^-$ αβ T cells and, in high-responders, T_{FH} cell to T_{reg} cell and T_H17 cell to T_{reg} cell ratios
Humrich et al. (2019) (PRO-IMMUN study)[181]	SLE (12)	Phase I/IIa	5-day cycles of 1.5 MIU daily followed by 3 or 4.5 MIU daily according to T_{reg} cell increase at weeks 2, 5 and 8	Significant decrease in SELENA-SLEDAI before second treatment cycle in responders ($P=0.03$); SELENA-SLEDAI changes significantly correlated with $CD25^{hi}$ T_{reg} cell increase	Primary end point[a] met in 92% of patients (11/12), which correlated with the cumulative IL-2 dose
Zhao et al. (2019)[191]	SLE (120)	Prospective, multicourse, open-label	100 'WIU' IL-2 s.c. for 3–5 days each month and 0.5 mg oral rapamycin every other day	Significant improvement in SLEDAI from 6 to 24 weeks ($P<0.001$)	T_{reg} cells significantly increased ($P<0.001$)
Shao et al. (2019)[192]	SLE (80)	Open label	Three cycles of a 2-week course with 1 MIU IL-2 every other day, followed by a 2-week treatment break (50/80 patients received IL-2)	In patients with IL-2 treatment, reduced serum IL-2 concentrations correlated with LN less remission and higher severity measured by 24-h urine protein	T_{reg} cells increased (only statistically significant in patients with LN)
He et al. (2020)[182]	SLE (60)	Randomized, single-centre, double-blind, placebo-controlled	Three cycles of 1 MIU IL-2 every other day for 2 weeks followed by a 2-week treatment break	Primary end point (significant increase in SRI-4) met at weeks 6, 8, 10, 16 and 24 ($P<0.05$) but not at week 12 ($P=0.052$); complete remission of LN in 54% (7/13) receiving IL-2 and 17% (2/12) receiving placebo ($P=0.036$)	Significant increase in T_{reg} cells ($P<0.05$); anti-dsDNA antibody titres decreased significantly; 24-h proteinuria decreased about threefold in the IL-2 group (baseline versus week 24)
Miao et al. (2018)[193]	Primary Sjögren syndrome (190)	Single-course, open-label	0.5 MIU IL-2 daily for 5 days versus continuing immunosuppression	No clinical improvement in the short term	T_{reg} cells and conventional $CD4^+$ T cells (including T_H17 cells) increased; T_H17 cell to T_{reg} cell ratio normalized after IL-2 treatment
Wang et al. (2020)[25]	PsA (117)	Single-course, randomized, open-label	0.5 MIU IL-2 (s.c.) daily for 5 days and conventional treatment (22/117) or conventional treatment only (95/117)	Significant improvements ($P<0.05$) in most disease activity measures (e.g. tender or swollen joint count)	Significant increase in absolute number of T_{reg} cells and T_H17 cell counts ($P<0.001$)

Table 1 (cont.) | **Trials of low-dose IL-2 treatment in autoimmune diseases**

Study	Disease (*n*)	Type of study	IL-2 regimen	Clinical outcome	Biological outcome
An et al. (2019)[194]	AS (48)	Single-course, open-label	50 'WIU' IL-2 (s.c.) daily for 5 days	Not assessed	Increase in $CD4^+$ T_{reg} cells, $CD8^+$ T_{reg} cells and T_H17 cells
Castela et al. (2014)[197]	Alopecia areata (5)	Single-course single-centre, prospective open-label	5-day course of 1.5 MIU IL-2 daily followed by 3 MIU IL-2 daily at weeks 3, 6 and 9	Partial hair regrowth (4/5), maintained and even improved after 6 months	T_{reg} cell recruitment into lesional skin and slight increase in peripheral blood
Le Duff et al. (2021)[198]	Alopecia areata (43)	Prospective, multicentre, randomized, placebo-controlled	5-day course of 1.5 MIU IL-2 daily followed by 5-day course of 3 MIU IL-2 daily at weeks 3, 6 and 9	No significant improvement in body hair or nails	Increase in total, especially naive, T_{reg} cells and NK cells
Zhang et al. (2018)[199]	Immune thrombocytopenia (3)	Case report	5-day cycle of 1.0 MIU IL-2 daily per week or 2 or 4 weeks	Thrombocyte increase in patients 1 and 2 but not in patient 3 (although T_{reg} cells increased)	T_{reg} cells increased twofold to threefold in patients 2 and 3; clinical response in patient 1, although T_{reg} numbers unchanged
Lim et al. (2018)[200]	Autoimmune hepatitis (2)	Case report	Monthly cycles of 5 days of 1 MIU IL-2 daily for 6 months	AST and IgG levels normalized in patient 2 and remained elevated in patient 1	Significant increase in T_{reg} cells at day 9 of each cycle ($P < 0.005$)
Feng et al. (2019)[195]	PM (10) or DM (61)	Open label	Conventional treatment and 0.5 MIU IL-2 daily for 5 days (7/10 PM; 35/61 DM) versus conventional treatment alone (3/10 PM; 26/61 DM)	CK, LDH and HBDH decreased significantly with both treatment regimens	Conventional T cells and T_{reg} cells significantly reduced with conventional treatment versus significantly increased with IL-2 combination therapy
Zhang et al. (2019)[196]	PM (39) or DM (108)	Open label	Conventional treatment and 0.5 MIU IL-2 daily for 5 days (31/147) versus conventional treatment alone (116/147)	Greater decrease in VAS, ESR, CK, CK-MB, LDH and HBDH with IL-2 than with conventional treatment alone	Fourfold increase in T_{reg} cells and modest increase in total T cells and subsets, and B cells

AS, ankylosing spondylitis; AST, aspartate aminotransferase; CGI, clinical global impression; CK, creatine kinase; CK-MB, CK myocardial band; DM, dermatomyositis; ESR, erythrocyte sedimentation rate; dsDNA, double-stranded DNA; GPA, granulomatosis with polyangiitis; HBDH, hydroxybutyrate dehydrogenase; HCV, hepatitis C virus; LDH, lactate dehydrogenase; LN, lupus nephritis; MIU, million international units; NK cell, natural killer cell; NKT cell, natural killer T cell; PM, polymyositis; PsA, psoriatic arthritis; RA, rheumatoid arthritis; s.c., subcutaneously; SELENA, Safety of Estrogens in Lupus Erythematosus National Assessment; SLE, systemic lupus erythematosus; SLEDAI, Systemic Lupus Erythematosus Disease Activity Index; SRI-4, SLE Responder Index 4; T_{FH} cell, T follicular helper cell; T_H17 cell, T helper 17 cell; T_{reg} cell, regulatory T cell; VAS, visual analogue scale. [a]Defined as a 100% increase in T_{reg} cell levels from baseline to the treatment end point at day 62.

enhanced T_{reg} cell function[64] but combination therapy with TNF blockers and IL-2 has not yet been assessed in clinical trials.

IL-2 and tocilizumab. IL-6, together with TGFβ, shifts the balance between T_H17 cells and T_{reg} cells towards T_H17 cell differentiation by abrogating the FOXP3-associated inhibition of *RORA* and *RORC* expression[210,211] and through STAT3-mediated downregulation of *FOXP3* expression[100]. Treatment of patients with RA with tocilizumab, an anti-IL-6 receptor antibody, leads to an expansion of the T_{reg} cell population[62,212–214] that is associated with increased Helios expression[215]. In a study in 50 patients with RA who were randomly assigned to a combination of IL-2 (at a dosage referred to as '50 WIU' for 5 days), tocilizumab (160 mg at day 1 and 3) and standard of care (glucocorticoids and DMARDs), or of IL-2 and standard of care, or standard of care alone, the triple combination of IL-2 and tocilizumab resulted in expansion of T_{reg} cells without affecting T_H17 cells, whereas IL-2 plus standard of care led to expansion of both cell populations. Compared with standard of care, the triple combination therapy significantly reduced the numbers of tender joints ($P < 0.01$) and swollen joints ($P < 0.05$)[216]. Of note, IL-6 promotes resistance of conventional T cells to T_{reg} cell-mediated suppression, providing an additional rationale for combining IL-6R blockade and low-dose IL-2 therapy[217,218].

T_{reg} cells and ustekinumab. Aside from its importance for T_H17 cell maintenance and IL-17 production, IL-23 decreases IL-2 production by impairing NF-κB^{p65} in mice[219] and induces differentiation of T_{reg} cells into pro-inflammatory $RoR\gamma t^+IL\text{-}17A^+$ T_{reg} cells in mice[220]. Furthermore, IL-23 inhibits T_{reg} cell induction[221] and differentiation of induced T_{reg} cells by regulating their responsiveness to IL-33 and, in thymic T_{reg} cells, ST2 signal transduction and expression of GATA3-regulated genes[222]. In psoriasis, treatment with the anti-IL-12–IL-23 monoclonal antibody ustekinumab leads to a T_{reg} cell increase in humans and mice[223]. In a recent double-blind, randomized, controlled phase II trial in SLE, ustekinumab showed superior efficacy compared with placebo and improved serum levels of anti-dsDNA antibodies

and C3 (REF.[224]), although the follow-up phase III trial was interrupted following futility analysis. Combination of IL-2 and ustekinumab therapy have not yet been tested but not only would inhibition of the IL-23 pathway restore T_{reg} cells but also might positively support T_{reg} cell activation with the addition of low-dose IL-2.

IL-2 and rapamycin. Combination therapy with low-dose IL-2 and rapamycin, a mTOR inhibitor that inhibits T_H17 cell differentiation while not affecting T_{reg} cell proliferation and differentiation, resulted in significant clinical responses in patients with therapy-refractory SLE[191] (TABLE 1; Supplementary Table 1). In autoimmune diseases, overactivation of conventional T cells and diminished T_{reg} cell functions both foster a pro-inflammatory environment for disease pathogenesis; even with normal activity, T_{reg} cells would have difficulty in controlling an overactivated immune system. The combination of specific inhibition of effector T cells and specific induction of T_{reg} cells has great potential for controlling autoimmunity and restoring immune tolerance.

Novel IL-2 therapies

Low-dose IL-2 has a half-life of ~10 min after intravenous injection and 4–6 h after subcutaneous injection in humans but its biological effect on T_{reg} cells lasts for days or weeks (depending on the dose)[112]. Improvement of the IL-2 half-life should increase the interval between injections[225], although limiting the ease of treatment interruption in the event of adverse effects. Dose optimization trials have shown a dose-dependent activation of conventional T cells at doses of IL-2 >3 MIU daily[5]. Although no study so far has found low-dose IL-2-dependent serious adverse events that could be related to stimulation of conventional T cells, improving the selectivity of IL-2 could allow the use of higher doses to trigger a greater effect on T_{reg} cells. Similarly, many academic and pharmaceutical industry researchers have expressed interest in generating novel IL-2 formulations with improved half-life and selectivity. For the treatment of cancer, the goal is to stimulate conventional T cells without stimulating T_{reg} cells, mostly by reducing the affinity of IL-2Rα for IL-2. Conversely, for the treatment of autoimmunity and inflammation, improving selectivity for T_{reg} cells is mostly based on reducing the affinity of IL-2Rβ for IL-2 (TABLE 2).

To improve selectivity and reduce toxic adverse effects — especially of high-dose IL-2 for treatment of tumours — so-called IL-2 muteins (IL-2 proteins with an altered amino acid sequence) were created, by, for example, alterations that increase IL-2 affinity for IL-2Rβ (as in IL-2 superkines, thereby forgoing the requirement for binding to IL-1Rα first) or PEGylation (as in NKTR-214, in which IL-2 PEGylation blocks its interaction with IL-2Rα)[39,226].

The characteristics of IL-2–IL-2 antibody complexes have also been modulated by using IL-2 monoclonal antibodies that bind to different sites in IL-2, and this has identified complexes that show improved half-life and selectivity and can be directed to either memory $CD8^+$ T cells and NK cells (using mouse S4B6 or the human MAB602 monoclonal antibodies) or T_{reg} cells (using mouse JES6-1 or human 5344 monoclonal antibodies)[39]. Mouse T_{reg} cell-directed complexes have been used in several models of autoimmune, inflammatory and metabolic diseases and transplantation. In addition, humanized complexes such as IL-2–F5111.2 are effective in mouse models of T1DM, EAE and GVHD[227].

Fusion proteins containing IL-2 (or muteins)[228,229] bound to an antibody against a cytokine or antigen[230], so-called immunocytokines, can direct IL-2 to a specific cell type or locally to a tissue site, thereby enhancing local immunity[231]. T_{reg} cell-directed approaches have so far targeted TNFR2 (REF.[232]) or have involved a fusion of an IgG1 molecule to two IL-2 N88D muteins[233]. In addition, SLAMF3 on T_{reg} cells would be a promising candidate in SLE[234]. SLAMF3 is a co-regulatory molecule in T cell activation and differentiation that enhances IL-2 sensitivity in T cells by IL-2Rα upregulation, especially promoting T_{reg} cell differentiation. Therefore, SLAMF3 co-stimulation in addition to low-dose IL-2 treatment could reverse the T_{reg} cell deficiency in SLE, and thus SLAMF3 represents a promising therapeutic target. IL-2 targeting to tissues seems to be the next promising approach for the treatment of autoimmune diseases.

Alternatively, viral vector-mediated gene transfer can be used to ensure continuous IL-2 production, either systemically (preventing T1DM in NOD mice)[163]. However, this approach does not permit treatment interruption or cessation in the event of unwanted adverse effects. Aerosolized IL-2 seems a feasible delivery mechanism[235] and nanoparticles coated with anti-CD2 or anti-CD4 antibodies have been used to deliver encapsulated IL-2 and TGFβ to T cells in mice with lupus[236].

Conclusions

Expanding and activating T_{reg} cells is now unanimously viewed as an important novel approach to developing treatments for autoimmune diseases. IL-2 is the first-in-class molecule for this approach. IL-2 in its natural form and used at low doses has already achieved some important development milestones. Numerous studies have shown that it is possible to efficiently expand and activate T_{reg} cells using low-dose IL-2 without activation of conventional T cells, and this treatment has a good safety profile, and clinical efficacy results are encouraging. However, these results now need to be confirmed in phase III trials, which, if successful, could introduce IL-2 as a first-line, second-line or add-on biologic in the treatment of many autoimmune diseases.

The use of IL-2 should be welcome given that the adverse effects are limited and mild and the expected clinical benefit appears soon after commencing treatment. Furthermore, IL-2 might perform well in combination with established anti-inflammatory drugs and biologics to enhance the established therapeutic benefit. It is quite possible that IL-2 might serve as the first-line biologic of choice for use in people with mild or moderate autoimmune disease and may obviate the use of steroids or cytotoxic drugs.

In people with severe autoimmune disease, the use of IL-2 in conjunction with other anti-inflammatory therapeutics may be required because, as discussed extensively, T_{reg} cells do not fare well in severe inflammation

and IL-2 may be of limited or even no benefit in these conditions. It is possible that therapeutic schemes will involve first the use of an established anti-inflammatory drug followed by IL-2 administration. In addition, some drugs, such as methotrexate or TNF inhibitors, even induce T_{reg} cells. The use of IL-2 should result in

Table 2 | **IL-2 formulations for the treatment of rheumatic diseases**

Formulation	Description	Effect on T_{reg} cells	Effect on effector immune cells	Adverse effects	Disease or mouse model
Low-dose IL-2 (REFS[25,110,150,179,181,182,189,194])	Trimeric IL-2R with high affinity for IL-2	Strong activation and expansion	Marked increase in $CD56^{hi}$ NK cells	Dose-dependent, mild-to-moderate and transient: injection site reactions (≤31%), influenza-like symptoms (≤10%), fever (≤14%), infections (≤12%), myalgia (≤9%) and nausea (≤4%)	Human autoimmune diseases, including SLE, RA, AS, psoriasis, Behçet disease, GPA, Takayasu arteritis, Crohn's disease, ulcerative colitis, autoimmune hepatitis, sclerosing cholangitis, vasculitis and ITP
NKTR-358 (REF.[237])	PEGylated IL-2 mutein with reduced affinity for IL-2Rβ	Strong activation	No activation of $CD4^+$ T cells or $CD8^+$ T cells; less than fourfold activation of NK cells	No dose-limiting toxicities, mostly mild injection site reactions	Human SLE
IL-2Rα-directed IL-2–anti-IL-2 antibody complexes					
IL-2–JES6-1 (REF.[238])	Anti-mouse IL-2 mAb JES6-1 (blocks IL-2Rβ/γ binding site on IL-2)	Strong activation	Efficient inhibition of naive and memory $CD4^+$ T cells and $CD8^+$ T cells, and NK cells	Not reported	Mouse models of autoimmunity, including T1DM in NOD mice, collagen-induced arthritis, T cell-mediated asthma, EAE, experimental myasthenia, pancreatic islet graft transplantation, and mouse skin allograft[238]
IL-2–F5111.2 (REF.[227])	Anti-human IL-2 mAb F5111.2 (blocks IL-2Rγ binding site on IL-2)	Strong activation	Efficient inhibition of naive and memory $CD4^+$ T cells and $CD8^+$ T cells, and NK cells	Not reported	T1DM in NOD mice, EAE and GVHD[227]
IL-2–UFKA1 (REF.[239])	Anti-human IL-2 mAb UFKA1 (blocks IL-2Rγ binding site on IL-2)	Strong activation	Efficient inhibition of naive and memory $CD4^+$ T cells and $CD8^+$ T cells, and NK cells	Not tested in vivo	Not tested in vivo
Fusion proteins					
IL-2–EHD2–sc-mTNFR2 (REFS[232,240])	TNFR2-specific TNF–IL-2 fusion	Strong activation of $CD4^+$ T_{reg} cells and $CD8^+$ T_{reg} cells	No activation of $CD4^+$ T cells, $CD8^+$ T cells, B cells, DCs, monocytes or neutrophils; NK cells not assessed	Not reported	Experimental collagen-induced arthritis
$IgG1–(IL-2N88D)_2$ (REF.[233])	Two IL-2 muteins fused to IgG1	Strong activation of $CD4^+$ T_{reg} cells and $CD8^+$ T_{reg} cells	No activation of $CD4^+$ T cells, $CD8^+$ T cells, B cells, DCs, monocytes or neutrophils; NK cells not assessed	Not reported	Experimental collagen-induced arthritis
Selectikine[241]	IL-2 fused to anti-DNA mAb NHS76	T_{reg} cell expansion lower than that of $CD4^+$ T cells	Strong activation of $CD4^+$ T cells and $CD8^+$ T cells and weak activation of NK cells	Minor (rash, lymphopenia hypotension); no major toxicity (such as VLS)	Human metastatic or locally advanced tumours
Alternative mechanisms of IL-2 delivery					
Nanoparticles[236]	Anti-CD2 and anti-CD4 antibody-coated, T cell-targeted nanoparticles loaded with IL-2 and TGFβ	Expansion of $CD4^+$ T_{reg} cells (twofold) and $CD8^+$ T_{reg} cells (fourfold)	No activation of $CD4^+$ T cells, $CD8^+$ T cells, monocytes, DCs or granulocytes; NK cells not assessed	Reduction in circulating anti-dsDNA antibodies and proteinuria	BDF1 mouse lupus model
Viral vector[163]	Viral vector-mediated gene transfer for systemic IL-2 production	Increased T_{reg} cell abundance	No increase in $CD8^+$ T cells, NK cells or B cells in peripheral blood; increase in T_{eff} cells in spleen	None	T1DM in NOD mice

AS, ankylosing spondylitis; DC, dendritic cell; EAE, experimental autoimmune encephalomyelitis; GPA, granulomatosis with polyangiitis; GVHD, graft-versus-host disease; IL-2R, IL-2 receptor; ITP, immune thrombocytopenia; mAb, monoclonal antibody; NK cell, natural killer cell; NOD, non-obese diabetic; PEG, polyethylene glycol; RA, rheumatoid arthritis; SLE, systemic lupus erythematosus; T1DM, type 1 diabetes mellitus; TGFβ, transforming growth factor-β; T_{eff} cells, effector T cells; T_{reg} cell, regulatory T cell; VLS, vascular leak syndrome.

a decrease in the total usage of steroids and minimize their known adverse effects in patients with SLE and the total dose of methotrexate in patients with RA. Similar assumptions can be made for other autoimmune diseases that require continuous use of steroids or other immunosuppressive drugs.

It is hoped that IL-2 muteins or biologics in which IL-2 is fused to the Fc portion of IgG or PEGylated will be more effective than the natural form of IL-2. IL-2 muteins that are designed to avoid activating cytotoxic T cells and NK cells have reduced affinity for IL-2Rβ. However, this alteration results in an intrinsic loss of bioactivity towards all T cells, which requires using higher doses of the modified IL-2 that might favour the development of antibodies against this biologic. Also, a mild stimulatory effect of cytotoxic cells might even be desirable, as the abundance of these cells is decreased in SLE and other autoimmune diseases but they are needed to fight infection and control autoimmunity[128]. IL-2–IgG Fc fragment fusions could enable week-long or month-long intervals between injections.

Another consideration is that T_{reg} cells do not respond to IL-2 as expected when they are in the presence of pro-inflammatory cytokines or they might not respond at all because of an inherent signalling defect[116]. Drug developers might consider fusion of IL-2 with molecules that, by co-engaging other surface molecules, might enhance IL-2's T_{reg} cell-potentiating effect. For example, IL-2 could be fused with antibodies or other ligands to direct this cytokine to inflamed tissues, where it could suppress inflammation and contribute to the tissue repair process. However, the possible development of neutralizing antibodies that are not routinely elicited when natural IL-2 is used is a potential drawback.

While there is substantial interest in the development of IL-2-based biologics, there are a number of important considerations in the planning of clinical trials. First, for many of the diseases, it is difficult to demonstrate improvement over established, commonly used immunosuppressive regimens that have a placebo effect. Second, it might be necessary to stratify potential trial participants according to potential biomarkers such as T_{reg} cell number or activation markers. Third, the definition of clinical outcome does not always reflect biological or disease response. For example, several studies used T_{reg} cell expansion as the primary outcome, which does not necessarily correspond to immunosuppressive capacity or clinical improvement. However, this is attributable to the difficulties in comparing fine differences in inflammation in various organs in autoimmune diseases such as SLE, which not strictly overlap in all patients. Therefore, it is important that an organ-specific manifestation or an autoimmune disease with easily defined clinical measures should be chosen to test the efficacy of IL-2-based therapies in clinical trials.

We expect that the natural form or engineered derivatives of IL-2 will offer a distinct clinical benefit in multiple autoimmune diseases and, in view of the lack of serious adverse effects, their use in the clinic should rapidly gain support.

Dysregulation of T_{reg} cells by reduced abundance or functional impairment has been shown to occur in many human autoimmune diseases. Treatment with low-dose IL-2 has shown good potential to overcome these deficiencies, and promising results, including clinically significant improvements, have been obtained with T_{reg} cell induction in autoimmune diseases. Further development of novel, improved IL-2 formulations, by affinity for specific IL-2R subunits, reducing adverse effects or targeting IL-2 to the right tissue using many different approaches, are now in clinical development.

These T_{reg} cell-directed therapies have the potential to regulate and restore immune balance in autoimmune diseases. It is important that we discover the setting in which curbing a pro-inflammatory milieu by combination therapy with immunosuppressive drugs is most appropriate. Immunosuppressive drugs that promote T_{reg} cell differentiation should be used preferentially in autoimmune diseases. In the future, efforts to restore immune balance will include overcoming the challenge of inducing antigen-specific T_{reg} cells by specific stimulation instead of the current approach of general, non-specific T_{reg} cell activation. In the future, the use of IL-2 to induce antigen-specific T_{reg} cells could be the ultimate approach to reaching long-term remission or even cure.

Published online 2 November 2021

1. Campbell, C. & Rudensky, A. Roles of regulatory T cells in tissue pathophysiology and metabolism. *Cell Metab.* **31**, 18–25 (2020).
2. Morgan, D. A., Ruscetti, F. W. & Gallo, R. Selective in vitro growth of T lymphocytes from normal human bone marrows. *Science* **193**, 1007–1008 (1976).
3. Taniguchi, T. et al. Structure and expression of a cloned cDNA for human interleukin-2. *Nature* **302**, 305–310 (1983).
4. Malek, T. R. & Castro, I. Interleukin-2 receptor signaling: at the interface between tolerance and immunity. *Immunity* **33**, 153–165 (2010).
5. Klatzmann, D. & Abbas, A. K. The promise of low-dose interleukin-2 therapy for autoimmune and inflammatory diseases. *Nat. Rev. Immunol.* **15**, 283–294 (2015).
6. Abbas, A. K., Trotta, E., Simeonov, D. R., Marson, A. & Bluestone, J. A. Revisiting IL-2: biology and therapeutic prospects. *Sci. Immunol.* **3**, eaat1482 (2018).
7. Fontenot, J. D., Rasmussen, J. P., Gavin, M. A. & Rudensky, A. Y. A function for interleukin 2 in Foxp3-expressing regulatory T cells. *Nat. Immunol.* **6**, 1142–1151 (2005).
8. Knoechel, B., Lohr, J., Kahn, E., Bluestone, J. A. & Abbas, A. K. Sequential development of interleukin 2-dependent effector and regulatory T cells in response to endogenous systemic antigen. *J. Exp. Med.* **202**, 1375–1386 (2005).
9. Tang, Q. et al. Central role of defective interleukin-2 production in the triggering of islet autoimmune destruction. *Immunity* **28**, 687–697 (2008).
10. Sakaguchi, S., Sakaguchi, N., Asano, M., Itoh, M. & Toda, M. Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *J. Immunol.* **155**, 1151–1164 (1995).
11. Fuchs, A. et al. Minimum information about T regulatory cells: a step toward reproducibility and standardization. *Front. Immunol.* **8**, 1844 (2017).
12. Jonuleit, H. et al. Identification and functional characterization of human CD4(+)CD25(+) T cells with regulatory properties isolated from peripheral blood. *J. Exp. Med.* **193**, 1285–1294 (2001).
13. Piccirillo, C. A. & Shevach, E. M. Naturally-occurring CD4+CD25+ immunoregulatory T cells: central players in the arena of peripheral tolerance. *Semin. Immunol.* **16**, 81–88 (2004).
14. Liu, W. et al. CD127 expression inversely correlates with FoxP3 and suppressive function of human CD4+ T reg cells. *J. Exp. Med.* **203**, 1701–1711 (2006).
15. Li, W., Deng, C., Yang, H. & Wang, G. The regulatory T cell in active systemic lupus erythematosus patients: a systemic review and meta-analysis. *Front. Immunol.* **10**, 159 (2019).
16. Yang, H. X. et al. Are CD4+CD25-Foxp3+ cells in untreated new-onset lupus patients regulatory T cells? *Arthritis Res. Ther.* **11**, R153 (2009).
17. Sharabi, A. et al. Regulatory T cells in the treatment of disease. *Nat. Rev. Drug Discov.* **17**, 823–844 (2018).
18. Christoffersson, G. & von Herrath, M. Regulatory immune mechanisms beyond regulatory T cells. *Trends Immunol.* **40**, 482–491 (2019).
19. Churlaud, G. et al. Human and mouse CD8(+)CD25(+) FOXP3(+) regulatory T cells at steady state and during interleukin-2 therapy. *Front. Immunol.* **6**, 171 (2015).
20. Chaput, N. et al. Identification of CD8+CD25+Foxp3+ suppressive T cells in colorectal cancer tissue. *Gut* **58**, 520–529 (2009).

21. Roncarolo, M. G., Gregori, S., Bacchetta, R., Battaglia, M. & Gagliani, N. The biology of T regulatory type 1 cells and their therapeutic application in immune-mediated diseases. *Immunity* **49**, 1004–1019 (2018).
22. Ferreira, L. M. R., Muller, Y. D., Bluestone, J. A. & Tang, Q. Next-generation regulatory T cell therapy. *Nat. Rev. Drug Discov.* **18**, 749–769 (2019).
23. Spolski, R., Li, P. & Leonard, W. J. Biology and regulation of IL-2: from molecular mechanisms to human therapy. *Nat. Rev. Immunol.* **18**, 648–659 (2018).
24. Guo, H. et al. Functional defects in CD4(+) CD25(high) FoxP3(+) regulatory cells in ankylosing spondylitis. *Sci. Rep.* **6**, 37559 (2016).
25. Wang, J. et al. The numbers of peripheral regulatory T cells are reduced in patients with psoriatic arthritis and are restored by low-dose interleukin-2. *Ther. Adv. Chronic Dis.* **11**, 2040622320916014 (2020).
26. Haufe, S. et al. Impaired suppression of synovial fluid CD4+CD25- T cells from patients with juvenile idiopathic arthritis by CD4+CD25+ Treg cells. *Arthritis Rheum.* **63**, 3153–3162 (2011).
27. Frantz, C., Auffray, C., Avouac, J. & Allanore, Y. Regulatory T cells in systemic sclerosis. *Front. Immunol.* **9**, 2356 (2018).
28. Slobodin, G. & Rimar, D. Regulatory T cells in systemic sclerosis: a comprehensive review. *Clin. Rev. Allergy Immunol.* **52**, 194–201 (2017).
29. Grunewald, J. et al. Sarcoidosis. *Nat. Rev. Dis. Prim.* **5**, 45 (2019).
30. Powell, J. D., Ragheb, J. A., Kitagawa-Sakakida, S. & Schwartz, R. H. Molecular regulation of interleukin-2 expression by CD28 co-stimulation and anergy. *Immunol. Rev.* **165**, 287–300 (1998).
31. Smigiel, K. S. et al. CCR7 provides localized access to IL-2 and defines homeostatically distinct regulatory T cell subsets. *J. Exp. Med.* **211**, 121–136 (2014).
32. O'Gorman, W. E. et al. The initial phase of an immune response functions to activate regulatory T cells. *J. Immunol.* **183**, 332–339 (2009).
33. Liu, Z. et al. Immune homeostasis enforced by co-localized effector and regulatory T cells. *Nature* **528**, 225–230 (2015).
34. Yu, A. et al. Selective IL-2 responsiveness of regulatory T cells through multiple intrinsic mechanisms supports the use of low-dose IL-2 therapy in type 1 diabetes. *Diabetes* **64**, 2172–2183 (2015).
35. Yu, A., Zhu, L., Altman, N. H. & Malek, T. R. A low interleukin-2 receptor signaling threshold supports the development and homeostasis of T regulatory cells. *Immunity* **30**, 204–217 (2009).
36. Hori, S., Nomura, T. & Sakaguchi, S. Control of regulatory T cell development by the transcription factor Foxp3. *Science* **299**, 1057–1061 (2003).
37. Nelson, B. H. & Willerford, D. M. Biology of the interleukin-2 receptor. *Adv. Immunol.* **70**, 1–81 (1998).
38. Liao, W., Lin, J. X. & Leonard, W. J. Interleukin-2 at the crossroads of effector responses, tolerance, and immunotherapy. *Immunity* **38**, 13–25 (2013).
39. Arenas-Ramirez, N., Woytschak, J. & Boyman, O. Interleukin-2: biology, design and application. *Trends Immunol.* **36**, 763–777 (2015).
40. Stauber, D. J., Debler, E. W., Horton, P. A., Smith, K. A. & Wilson, I. A. Crystal structure of the IL-2 signaling complex: paradigm for a heterotrimeric cytokine receptor. *Proc. Natl Acad. Sci. USA* **103**, 2788–2793 (2006).
41. Boyman, O. & Sprent, J. The role of interleukin-2 during homeostasis and activation of the immune system. *Nat. Rev. Immunol.* **12**, 180–190 (2012).
42. Hoxhaj, G. & Manning, B. D. The PI3K-AKT network at the interface of oncogenic signalling and cancer metabolism. *Nat. Rev. Cancer* **20**, 74–88 (2020).
43. Schwartz, D. M. et al. JAK inhibition as a therapeutic strategy for immune and inflammatory diseases. *Nat. Rev. Drug Discov.* **16**, 843–862 (2017).
44. Hemar, A. et al. Endocytosis of interleukin 2 receptors in human T lymphocytes: distinct intracellular localization and fate of the receptor alpha, beta, and gamma chains. *J. Cell Biol.* **129**, 55–64 (1995).
45. Muller, M. R. & Rao, A. NFAT, immunity and cancer: a transcription factor comes of age. *Nat. Rev. Immunol.* **10**, 645–656 (2010).
46. Burzyn, D., Benoist, C. & Mathis, D. Regulatory T cells in nonlymphoid tissues. *Nat. Immunol.* **14**, 1007–1013 (2013).
47. Michalek, R. D. et al. Cutting edge: distinct glycolytic and lipid oxidative metabolic programs are essential for effector and regulatory CD4+ T cell subsets. *J. Immunol.* **186**, 3299–3303 (2011).
48. Kim, H. P., Imbert, J. & Leonard, W. J. Both integrated and differential regulation of components of the IL-2/IL-2 receptor system. *Cytokine Growth Factor. Rev.* **17**, 349–366 (2006).
49. Walsh, P. T. et al. PTEN inhibits IL-2 receptor-mediated expansion of CD4+ CD25+ Tregs. *J. Clin. Invest.* **116**, 2521–2531 (2006).
50. Delgoffe, G. M. et al. Stability and function of regulatory T cells is maintained by a neuropilin-1-semaphorin-4a axis. *Nature* **501**, 252–256 (2013).
51. Wing, K. et al. CTLA-4 control over Foxp3+ regulatory T cell function. *Science* **322**, 271–275 (2008).
52. Qureshi, O. S. et al. Trans-endocytosis of CD80 and CD86: a molecular basis for the cell-extrinsic function of CTLA-4. *Science* **332**, 600–603 (2011).
53. Deaglio, S. et al. Adenosine generation catalyzed by CD39 and CD73 expressed on regulatory T cells mediates immune suppression. *J. Exp. Med.* **204**, 1257–1265 (2007).
54. Chinen, T. et al. An essential role for the IL-2 receptor in Treg cell function. *Nat. Immunol.* **17**, 1322–1333 (2016).
55. Busse, D. et al. Competing feedback loops shape IL-2 signaling between helper and regulatory T lymphocytes in cellular microenvironments. *Proc. Natl Acad. Sci. USA* **107**, 3058–3063 (2010).
56. Gondek, D. C., Lu, L. F., Quezada, S. A., Sakaguchi, S. & Noelle, R. J. Cutting edge: contact-mediated suppression by CD4+CD25+ regulatory cells involves a granzyme B-dependent, perforin-independent mechanism. *J. Immunol.* **174**, 1783–1786 (2005).
57. Cao, X. et al. Granzyme B and perforin are important for regulatory T cell-mediated suppression of tumor clearance. *Immunity* **27**, 635–646 (2007).
58. Barron, L. et al. Cutting edge: mechanisms of IL-2-dependent maintenance of functional regulatory T cells. *J. Immunol.* **185**, 6426–6430 (2010).
59. Asano, T. et al. PD-1 modulates regulatory T-cell homeostasis during low-dose interleukin-2 therapy. *Blood* **129**, 2186–2197 (2017).
60. Park, H. J. et al. Tumor-infiltrating regulatory T cells delineated by upregulation of PD-1 and inhibitory receptors. *Cell Immunol.* **278**, 76–83 (2012).
61. Konopacki, C., Pritykin, Y., Rubtsov, Y., Leslie, C. S. & Rudensky, A. Y. Transcription factor Foxp1 regulates Foxp3 chromatin binding and coordinates regulatory T cell function. *Nat. Immunol.* **20**, 232–242 (2019).
62. Scheinecker, C., Goschl, L. & Bonelli, M. Treg cells in health and autoimmune diseases: new insights from single cell analysis. *J. Autoimmun.* **110**, 102376 (2020).
63. Chen, X. & Oppenheim, J. J. The phenotypic and functional consequences of tumour necrosis factor receptor type 2 expression on CD4(+) FoxP3(+) regulatory T cells. *Immunology* **133**, 426–433 (2011).
64. Zaragoza, B. et al. Suppressive activity of human regulatory T cells is maintained in the presence of TNF. *Nat. Med.* **22**, 16–17 (2016).
65. Patton, D. T., Wilson, M. D., Rowan, W. C., Soond, D. R. & Okkenhaug, K. The PI3K p110δ regulates expression of CD38 on regulatory T cells. *PLoS ONE* **6**, e17359 (2011).
66. Krejcik, J. et al. Daratumumab depletes CD38+ immune regulatory cells, promotes T-cell expansion, and skews T-cell repertoire in multiple myeloma. *Blood* **128**, 384–394 (2016).
67. Burkett, P. R., Meyer zu Horste, G. & Kuchroo, V. K. Pouring fuel on the fire: Th17 cells, the environment, and autoimmunity. *J. Clin. Invest.* **125**, 2211–2219 (2015).
68. Zielinski, C. E. et al. Pathogen-induced human TH17 cells produce IFN-ɣ or IL-10 and are regulated by IL-1β. *Nature* **484**, 514–518 (2012).
69. Stockinger, B. & Omenetti, S. The dichotomous nature of T helper 17 cells. *Nat. Rev. Immunol.* **17**, 535–544 (2017).
70. Liao, W., Lin, J. X., Wang, L., Li, P. & Leonard, W. J. Modulation of cytokine receptors by IL-2 broadly regulates differentiation into helper T cell lineages. *Nat. Immunol.* **12**, 551–559 (2011).
71. Laurence, A. et al. Interleukin-2 signaling via STAT5 constrains T helper 17 cell generation. *Immunity* **26**, 371–381 (2007).
72. Yang, X. P. et al. Opposing regulation of the locus encoding IL-17 through direct, reciprocal actions of STAT3 and STAT5. *Nat. Immunol.* **12**, 247–254 (2011).
73. Ballesteros-Tato, A. et al. Interleukin-2 inhibits germinal center formation by limiting T follicular helper cell differentiation. *Immunity* **36**, 847–856 (2012).
74. Li, J., Lu, E., Yi, T. & Cyster, J. G. EBI2 augments Tfh cell fate by promoting interaction with IL-2-quenching dendritic cells. *Nature* **533**, 110–114 (2016).
75. Oestreich, K. J., Mohn, S. E. & Weinmann, A. S. Molecular mechanisms that control the expression and activity of Bcl-6 in TH1 cells to regulate flexibility with a TFH-like gene profile. *Nat. Immunol.* **13**, 405–411 (2012).
76. Oestreich, K. J. et al. Bcl-6 directly represses the gene program of the glycolysis pathway. *Nat. Immunol.* **15**, 957–964 (2014).
77. Wing, J. B., Tekguc, M. & Sakaguchi, S. Control of germinal center responses by T-follicular regulatory cells. *Front. Immunol.* **9**, 1910 (2018).
78. Ritvo, P. G. et al. Tfr cells lack IL-2Rα but express decoy IL-1R2 and IL-1Ra and suppress the IL-1-dependent activation of Tfh cells. *Sci. Immunol.* **2**, eaan0368 (2017).
79. Deng, J., Wei, Y., Fonseca, V. R., Graca, L. & Yu, D. T follicular helper cells and T follicular regulatory cells in rheumatic diseases. *Nat. Rev. Rheumatol.* **15**, 475–490 (2019).
80. Panduro, M., Benoist, C. & Mathis, D. Tissue Tregs. *Annu. Rev. Immunol.* **34**, 609–633 (2016).
81. Hovhannisyan, Z., Treatman, J., Littman, D. R. & Mayer, L. Characterization of interleukin-17-producing regulatory T cells in inflamed intestinal mucosa from patients with inflammatory bowel diseases. *Gastroenterology* **140**, 957–965 (2011).
82. Malhotra, N. et al. RORα-expressing T regulatory cells restrain allergic skin inflammation. *Sci. Immunol.* **3**, eaao6923 (2018).
83. Delacher, M. et al. Genome-wide DNA-methylation landscape defines specialization of regulatory T cells in tissues. *Nat. Immunol.* **18**, 1160–1172 (2017).
84. Feuerer, M., Hill, J. A., Mathis, D. & Benoist, C. Foxp3+ regulatory T cells: differentiation, specification, subphenotypes. *Nat. Immunol.* **10**, 689–695 (2009).
85. Ali, N. et al. Regulatory T cells in skin facilitate epithelial stem cell differentiation. *Cell* **169**, 1119–1129.e11 (2017).
86. Hirata, Y. et al. CD150(high) bone marrow Tregs maintain hematopoietic stem cell quiescence and immune privilege via adenosine. *Cell Stem Cell* **22**, 445–453.e5 (2018).
87. Fujisaki, J. et al. In vivo imaging of Treg cells providing immune privilege to the haematopoietic stem-cell niche. *Nature* **474**, 216–219 (2011).
88. Li, J., Tan, J., Martino, M. M. & Lui, K. O. Regulatory T-cells: potential regulator of tissue repair and regeneration. *Front. Immunol.* **9**, 585 (2018).
89. Burzyn, D. et al. A special population of regulatory T cells potentiates muscle repair. *Cell* **155**, 1282–1295 (2013).
90. Villalta, S. A. et al. Regulatory T cells suppress muscle inflammation and injury in muscular dystrophy. *Sci. Transl. Med.* **6**, 258ra142 (2014).
91. Sharma, A. & Rudra, D. Emerging functions of regulatory T cells in tissue homeostasis. *Front. Immunol.* **9**, 883 (2018).
92. Arpaia, N. et al. A distinct function of regulatory T cells in tissue protection. *Cell* **162**, 1078–1089 (2015).
93. Zaiss, D. M. W., Gause, W. C., Osborne, L. C. & Artis, D. Emerging functions of amphiregulin in orchestrating immunity, inflammation, and tissue repair. *Immunity* **42**, 216–226 (2015).
94. Boothby, M. Signaling in T cells–is anything the m(a) TOR with the picture(s)? *F1000Res.* **5**, 191 (2016).
95. Sharabi, A. & Tsokos, G. C. T cell metabolism: new insights in systemic lupus erythematosus pathogenesis and therapy. *Nat. Rev. Rheumatol.* **16**, 100–112 (2020).
96. Angelin, A. et al. Foxp3 reprograms T cell metabolism to function in low-glucose, high-lactate environments. *Cell Metab.* **25**, 1282–1293.e7 (2017).
97. Zeng, H. & Chi, H. Metabolic control of regulatory T cell development and function. *Trends Immunol.* **36**, 3–12 (2015).
98. De Rosa, V. et al. Glycolysis controls the induction of human regulatory T cells by modulating the expression of FOXP3 exon 2 splicing variants. *Nat. Immunol.* **16**, 1174–1184 (2015).
99. Smith, P. M. et al. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341**, 569–573 (2013).
100. Yang, X. O. et al. Molecular antagonism and plasticity of regulatory and inflammatory T cell programs. *Immunity* **29**, 44–56 (2008).
101. Takeuchi, Y., Hirota, K. & Sakaguchi, S. Synovial tissue inflammation mediated by autoimmune T cells. *Front. Immunol.* **10**, 1989 (2019).
102. Alexander, T. et al. Foxp3+ Helios+ regulatory T cells are expanded in active systemic lupus erythematosus. *Ann. Rheum. Dis.* **72**, 1549–1558 (2013).
103. Golding, A., Hasni, S., Illei, G. & Shevach, E. M. The percentage of FoxP3+Helios+ Treg cells correlates

positively with disease activity in systemic lupus erythematosus. *Arthritis Rheum.* **65**, 2898–2906 (2013).
104. Smolen, J. S. et al. Rheumatoid arthritis. *Nat. Rev. Dis. Prim.* **4**, 18001 (2018).
105. Malmstrom, V., Catrina, A. I. & Klareskog, L. The immunopathogenesis of seropositive rheumatoid arthritis: from triggering to targeting. *Nat. Rev. Immunol.* **17**, 60–75 (2017).
106. McInnes, I. B. & Schett, G. The pathogenesis of rheumatoid arthritis. *N. Engl. J. Med.* **365**, 2205–2219 (2011).
107. Humby, F. et al. Ectopic lymphoid structures support ongoing production of class-switched autoantibodies in rheumatoid synovium. *PLoS Med.* **6**, e1 (2009).
108. Ziff, M. Relation of cellular infiltration of rheumatoid synovial membrane to its immune response. *Arthritis Rheum.* **17**, 313–319 (1974).
109. Kitas, G. D., Salmon, M., Farr, M., Gaston, J. S. & Bacon, P. A. Deficient interleukin 2 production in rheumatoid arthritis: association with active disease and systemic complications. *Clin. Exp. Immunol.* **73**, 242–249 (1988).
110. von Spee-Mayer, C. et al. Low-dose interleukin-2 selectively corrects regulatory T cell defects in patients with systemic lupus erythematosus. *Ann. Rheum. Dis.* **75**, 1407–1415 (2016).
111. Lieberman, L. A. & Tsokos, G. C. The IL-2 defect in systemic lupus erythematosus disease has an expansive effect on host immunity. *J. Biomed. Biotechnol.* **2010**, 740619 (2010).
112. Rosenzwajg, M. et al. Low-dose interleukin-2 fosters a dose-dependent regulatory T cell tuned milieu in T1D patients. *J. Autoimmun.* **58**, 48–58 (2015).
113. Morita, T. et al. The proportion of regulatory T cells in patients with rheumatoid arthritis: a meta-analysis. *PLoS ONE* **11**, e0162306 (2016).
114. Cammarata, I. et al. Counter-regulation of regulatory T cells by autoreactive CD8(+) T cells in rheumatoid arthritis. *J. Autoimmun.* **99**, 81–97 (2019).
115. Tsokos, G. C. Systemic lupus erythematosus. *N. Engl. J. Med.* **365**, 2110–2121 (2011).
116. Tsokos, G. C. Autoimmunity and organ damage in systemic lupus erythematosus. *Nat. Immunol.* **21**, 605–614 (2020).
117. Humrich, J. Y. & Riemekasten, G. Restoring regulation–IL-2 therapy in systemic lupus erythematosus. *Expert Rev. Clin. Immunol.* **12**, 1153–1160 (2016).
118. Mizui, M. & Tsokos, G. C. Targeting regulatory T cells to treat patients with systemic lupus erythematosus. *Front. Immunol.* **9**, 786 (2018).
119. Ohl, K. & Tenbrock, K. Regulatory T cells in systemic lupus erythematosus. *Eur. J. Immunol.* **45**, 344–355 (2015).
120. Tsokos, G. C., Lo, M. S., Costa Reis, P. & Sullivan, K. E. New insights into the immunopathogenesis of systemic lupus erythematosus. *Nat. Rev. Rheumatol.* **12**, 716–730 (2016).
121. Katsuyama, T., Tsokos, G. C. & Moulton, V. R. Aberrant T cell signaling and subsets in systemic lupus erythematosus. *Front. Immunol.* **9**, 1088 (2018).
122. Comte, D. et al. Brief report: CD4+ T cells from patients with systemic lupus erythematosus respond poorly to exogenous interleukin-2. *Arthritis Rheumatol.* **69**, 808–813 (2017).
123. Linker-Israeli, M. et al. Defective production of interleukin 1 and interleukin 2 in patients with systemic lupus erythematosus (SLE). *J. Immunol.* **130**, 2651–2655 (1983).
124. Alcocer-Varela, J. & Alarcon-Segovia, D. Decreased production of and response to interleukin-2 by cultured lymphocytes from patients with systemic lupus erythematosus. *J. Clin. Invest.* **69**, 1388–1392 (1982).
125. Dauphinee, M. J., Kipper, S. B., Wofsy, D. & Talal, N. Interleukin 2 deficiency is a common feature of autoimmune mice. *J. Immunol.* **127**, 2483–2487 (1981).
126. Humrich, J. Y. et al. Homeostatic imbalance of regulatory and effector T cells due to IL-2 deprivation amplifies murine lupus. *Proc. Natl Acad. Sci. USA* **107**, 204–209 (2010).
127. Koga, T. et al. CaMK4-dependent activation of AKT/mTOR and CREM-α underlies autoimmunity-associated Th17 imbalance. *J. Clin. Invest.* **124**, 2234–2245 (2014).
128. Moulton, V. R., Grammatikos, A. P., Fitzgerald, L. M. & Tsokos, G. C. Splicing factor SF2/ASF rescues IL-2 production in T cells from systemic lupus erythematosus patients by activating IL-2 transcription. *Proc. Natl Acad. Sci. USA* **110**, 1845–1850 (2013).
129. Kyttaris, V. C., Juang, Y. T., Tenbrock, K., Weinstein, A. & Tsokos, G. C. Cyclic adenosine 5′-monophosphate response element modulator is responsible for the decreased expression of c-fos and activator protein-1 binding in T cells from patients with systemic lupus erythematosus. *J. Immunol.* **173**, 3557–3563 (2004).
130. Costa, N. et al. Two separate effects contribute to regulatory T cell defect in systemic lupus erythematosus patients and their unaffected relatives. *Clin. Exp. Immunol.* **189**, 318–330 (2017).
131. Li, S. et al. Downregulation of IL-10 secretion by Treg cells in osteoarthritis is associated with a reduction in Tim-3 expression. *Biomed. Pharmacother.* **79**, 159–165 (2016).
132. Moradi, B. et al. CD4(+)CD25(+)/highCD127low/(-) regulatory T cells are enriched in rheumatoid arthritis and osteoarthritis joints–analysis of frequency and phenotype in synovial membrane, synovial fluid and peripheral blood. *Arthritis Res. Ther.* **16**, R97 (2014).
133. Ranganathan, V., Gracey, E., Brown, M. A., Inman, R. D. & Haroon, N. Pathogenesis of ankylosing spondylitis – recent advances and future directions. *Nat. Rev. Rheumatol.* **13**, 359–367 (2017).
134. Bravo, A. & Kavanaugh, A. Bedside to bench: defining the immunopathogenesis of psoriatic arthritis. *Nat. Rev. Rheumatol.* **15**, 645–656 (2019).
135. Appel, H. et al. Synovial and peripheral blood CD4+FoxP3+ T cells in spondyloarthritis. *J. Rheumatol.* **38**, 2445–2451 (2011).
136. Copland, A. & Bending, D. Foxp3 molecular dynamics in Treg in juvenile idiopathic arthritis. *Front. Immunol.* **9**, 2273 (2018).
137. Brito-Zerón, P. et al. Sjögren syndrome. *Nat. Rev. Dis. Prim.* **2**, 16047 (2016).
138. Li, X. et al. T regulatory cells are markedly diminished in diseased salivary glands of patients with primary Sjögren's syndrome. *J. Rheumatol.* **34**, 2438–2445 (2007).
139. Alunno, A. et al. T regulatory and T helper 17 cells in primary Sjögren's syndrome: facts and perspectives. *Med. Inflamm.* **2015**, 243723 (2015).
140. Luo, J. et al. IL-2 inhibition of Th17 generation rather than induction of treg cells is impaired in primary Sjögren's syndrome patients. *Front. Immunol.* **9**, 1755 (2018).
141. Liu, C., Guan, Z., Zhao, L., Song, Y. & Wang, H. Elevated level of circulating CD4(+)Helios(+)FoxP3(+) cells in primary Sjögren's syndrome patients. *Mod. Rheumatol.* **27**, 630–637 (2017).
142. Alunno, A. et al. Characterization of a new regulatory CD4+ T cell subset in primary Sjögren's syndrome. *Rheumatology* **52**, 1387–1396 (2013).
143. Allanore, Y. et al. Systemic sclerosis. *Nat. Rev. Dis. Prim.* **1**, 15002 (2015).
144. Koreth, J. et al. Efficacy, durability, and response predictors of low-dose interleukin-2 therapy for chronic graft-versus-host disease. *Blood* **128**, 130–137 (2016).
145. Koreth, J. et al. Interleukin-2 and regulatory T cells in graft-versus-host disease. *N. Engl. J. Med.* **365**, 2055–2066 (2011).
146. Morgan, M. D. et al. Patients with Wegener's granulomatosis demonstrate a relative deficiency and functional impairment of T-regulatory cells. *Immunology* **130**, 64–73 (2010).
147. Tsurikisawa, N., Saito, H., Oshikata, C., Tsuburai, T. & Akiyama, K. Decreases in the numbers of peripheral blood regulatory T cells, and increases in the levels of memory and activated B cells, in patients with active eosinophilic granulomatosis and polyangiitis. *J. Clin. Immunol.* **33**, 965–976 (2013).
148. Tsurikisawa, N., Saito, H., Oshikata, C., Tsuburai, T. & Akiyama, K. High-dose intravenous immunoglobulin therapy for eosinophilic granulomatosis with polyangiitis. *Clin. Transl. Allergy* **4**, 38 (2014).
149. Boyer, O. et al. CD4+CD25+ regulatory T-cell deficiency in patients with hepatitis C-mixed cryoglobulinemia vasculitis. *Blood* **103**, 3428–3430 (2004).
150. Saadoun, D. et al. Regulatory T-cell responses to low-dose interleukin-2 in HCV-induced vasculitis. *N. Engl. J. Med.* **365**, 2067–2077 (2011).
151. Shimojima, Y., Ishii, W., Kishida, D., Fukushima, K. & Ikeda, S. I. Imbalanced expression of dysfunctional regulatory T cells and T-helper cells relates to immunopathogenesis in polyarteritis nodosa. *Mod. Rheumatol.* **27**, 102–109 (2017).
152. Sugita, S., Yamada, Y., Kaneko, S., Horie, S. & Mochizuki, M. Induction of regulatory T cells by infliximab in Behçet's disease. *Invest. Ophthalmol. Vis. Sci.* **52**, 476–484 (2011).
153. Olivito, B. et al. Defective FOXP3 expression in patients with acute Kawasaki disease and restoration by intravenous immunoglobulin therapy. *Clin. Exp. Rheumatol.* **28**, 93–97 (2010).
154. Kong, X. et al. The critical role of IL-6 in the pathogenesis of Takayasu arteritis. *Clin. Exp. Rheumatol.* **34**, S21–S27 (2016).
155. Samson, M. et al. Th1 and Th17 lymphocytes expressing CD161 are implicated in giant cell arteritis and polymyalgia rheumatica pathogenesis. *Arthritis Rheum.* **64**, 3788–3798 (2012).
156. Miyara, M. et al. The immune paradox of sarcoidosis and regulatory T cells. *J. Exp. Med.* **203**, 359–370 (2006).
157. Broos, C. E. et al. Impaired survival of regulatory T cells in pulmonary sarcoidosis. *Respir. Res.* **16**, 108 (2015).
158. Verwoerd, A. et al. Infliximab therapy balances regulatory T cells, tumour necrosis factor receptor 2 (TNFR2) expression and soluble TNFR2 in sarcoidosis. *Clin. Exp. Immunol.* **185**, 263–270 (2016).
159. So, A. K. & Martinon, F. Inflammation in gout: mechanisms and therapeutic targets. *Nat. Rev. Rheumatol.* **13**, 639–647 (2017).
160. Dai, X. J. et al. Changes of Treg/Th17 ratio in spleen of acute gouty arthritis rat induced by MSU crystals. *Inflammation* **41**, 1955–1964 (2018).
161. Bonnet, B. et al. Low-dose IL-2 induces regulatory T cell-mediated control of experimental food allergy. *J. Immunol.* **197**, 188–198 (2016).
162. Churlaud, G. et al. Pharmacodynamics of regulatory T cells in mice and humans treated with low-dose IL-2. *J. Allergy Clin. Immunol.* **142**, 1344–1346.e3 (2018).
163. Churlaud, G. et al. Sustained stimulation and expansion of Tregs by IL2 control autoimmunity without impairing immune responses to infection, vaccination and cancer. *Clin. Immunol.* **151**, 114–126 (2014).
164. Tahvildari, M. & Dana, R. Low-dose IL-2 therapy in transplantation, autoimmunity, and inflammatory diseases. *J. Immunol.* **203**, 2749–2755 (2019).
165. Rose, A. et al. IL-2 therapy diminishes renal inflammation and the activity of kidney-infiltrating CD4+ T cells in murine lupus nephritis. *Cells* https://doi.org/10.3390/cells8101234 (2019).
166. Yan, J. J. et al. IL-2/anti-IL-2 complexes ameliorate lupus nephritis by expansion of CD4(+)CD25(+)Foxp3(+) regulatory T cells. *Kidney Int.* **91**, 603–615 (2017).
167. Mizui, M. et al. IL-2 protects lupus-prone mice from multiple end-organ damage by limiting CD4-CD8- IL-17-producing T cells. *J. Immunol.* **193**, 2168–2177 (2014).
168. Crispin, J. C. et al. Expanded double negative T cells in patients with systemic lupus erythematosus produce IL-17 and infiltrate the kidneys. *J. Immunol.* **181**, 8761–8766 (2008).
169. Yokoyama, Y. et al. IL-2-anti-IL-2 monoclonal antibody immune complexes inhibit collagen-induced arthritis by augmenting regulatory T cell functions. *J. Immunol.* **201**, 1899–1906 (2018).
170. Dey, I. & Bishayi, B. Impact of simultaneous neutralization of IL-17A and treatment with recombinant IL-2 on Th17-Treg cell population in S.aureus induced septic arthritis. *Microb. Pathog.* **139**, 103903 (2020).
171. Bergmann, B. et al. Pre-treatment with IL2 gene therapy alleviates Staphylococcus aureus arthritis in mice. *BMC Infect. Dis.* **20**, 185 (2020).
172. Webster, K. E. et al. In vivo expansion of T reg cells with IL-2-mAb complexes: induction of resistance to EAE and long-term acceptance of islet allografts without immunosuppression. *J. Exp. Med.* **206**, 751–760 (2009).
173. Izquierdo, C. et al. Treatment of T1D via optimized expansion of antigen-specific Tregs induced by IL-2/anti-IL-2 monoclonal antibody complexes and peptide/MHC tetramers. *Sci. Rep.* **8**, 8106 (2018).
174. Taylor, A. E. et al. Interleukin 2 promotes hepatic regulatory T cell responses and protects from biliary fibrosis in murine sclerosing cholangitis. *Hepatology* **68**, 1905–1921 (2018).
175. Rosenberg, S. A. IL-2: the first effective immunotherapy for human cancer. *J. Immunol.* **192**, 5451–5458 (2014).
176. Rosenberg, S. A., Mule, J. J., Spiess, P. J., Reichert, C. M. & Schwarz, S. L. Regression of established pulmonary metastases and subcutaneous tumor mediated by the systemic administration of high-dose recombinant interleukin 2. *J. Exp. Med.* **161**, 1169–1188 (1985).
177. Sadlack, B. et al. Ulcerative colitis-like disease in mice with a disrupted interleukin-2 gene. *Cell* **75**, 253–261 (1993).
178. Malek, T. R., Yu, A., Vincek, V., Scibelli, P. & Kong, L. CD4 regulatory T cells prevent lethal autoimmunity in IL-2Rβ-deficient mice. Implications for the

nonredundant function of IL-2. *Immunity* **17**, 167–178 (2002).
179. Rosenzwajg, M. et al. Immunological and clinical effects of low-dose interleukin-2 across 11 autoimmune diseases in a single, open clinical trial. *Ann. Rheum. Dis.* **78**, 209–217 (2019).
180. Mahmoudpour, S. H. et al. Safety of low-dose subcutaneous recombinant interleukin-2: systematic review and meta-analysis of randomized controlled trials. *Sci. Rep.* **9**, 7145 (2019).
181. Humrich, J. Y. et al. Low-dose interleukin-2 therapy in refractory systemic lupus erythematosus: an investigator-initiated, single-centre phase 1 and 2a clinical trial. *Lancet Rheumatol.* **1**, e44–e54 (2019).
182. He, J. et al. Efficacy and safety of low-dose IL-2 in the treatment of systemic lupus erythematosus: a randomised, double-blind, placebo-controlled trial. *Ann. Rheum. Dis.* **79**, 141–149 (2020).
183. Rosenzwajg, M. et al. Low-dose IL-2 in children with recently diagnosed type 1 diabetes: a phase I/II randomised, double-blind, placebo-controlled, dose-finding study. *Diabetologia* **63**, 1808–1821 (2020).
184. Krouse, R. S. et al. Thyroid dysfunction in 281 patients with metastatic melanoma or renal carcinoma treated with interleukin-2 alone. *J. Immunother. Emphas. Tumor Immunol.* **18**, 272–278 (1995).
185. Churlaud, G. et al. IL-2 antibodies in type 1 diabetes and during IL-2 therapy. *Diabetologia* **61**, 2066–2068 (2018).
186. Blattman, J. N. et al. Therapeutic use of IL-2 to enhance antiviral T-cell responses in vivo. *Nat. Med.* **9**, 540–547 (2003).
187. Hervier, B. et al. Phenotype and function of natural killer cells in systemic lupus erythematosus: excess interferon-γ production in patients with active disease. *Arthritis Rheum.* **63**, 1698–1706 (2011).
188. Busner, J. & Targum, S. D. The Clinical Global Impressions Scale: applying a research tool in clinical practice. *Psychiatry* **4**, 28–37 (2007).
189. He, J. et al. Low-dose interleukin-2 treatment selectively modulates CD4(+) T cell subsets in patients with systemic lupus erythematosus. *Nat. Med.* **22**, 991–993 (2016).
190. Humrich, J. Y. et al. Rapid induction of clinical remission by low-dose interleukin-2 in a patient with refractory SLE. *Ann. Rheum. Dis.* **74**, 791–792 (2015).
191. Zhao, C. et al. Low dose of IL-2 combined with rapamycin restores and maintains the long-term balance of Th17/Treg cells in refractory SLE patients. *BMC Immunol.* **20**, 32 (2019).
192. Shao, M. et al. Interleukin-2 deficiency associated with renal impairment in systemic lupus erythematosus. *J. Interferon Cytokine Res.* **39**, 117–124 (2019).
193. Miao, M. et al. Short-term and low-dose IL-2 therapy restores the Th17/Treg balance in the peripheral blood of patients with primary Sjögren's syndrome. *Ann. Rheum. Dis.* **77**, 1838–1840 (2018).
194. An, H. et al. The absolute counts of peripheral T lymphocyte subsets in patient with ankylosing spondylitis and the effect of low-dose interleukin-2. *Medicine* **98**, e15094 (2019).
195. Feng, M. et al. Absolute reduction of regulatory T cells and regulatory effect of short-term and low-dose IL-2 in polymyositis or dermatomyositis. *Int. Immunopharmacol.* **77**, 105912 (2019).
196. Zhang, S. X. et al. Circulating regulatory T cells were absolutely decreased in dermatomyositis/polymyositis patients and restored by low-dose IL-2. *Ann. Rheum. Dis.* https://doi.org/10.1136/annrheumdis-2019-216246 (2019).
197. Castela, E. et al. Effects of low-dose recombinant interleukin 2 to promote T-regulatory cells in alopecia areata. *JAMA Dermatol.* **150**, 748–751 (2014).
198. Le Duff, F. et al. Low-dose IL-2 for treating moderate to severe alopecia areata: a 52-week multicenter prospective placebo-controlled study assessing its impact on T regulatory cell and NK cell populations. *J. Invest. Dermatol.* **141**, 933–936 (2021).
199. Zhang, J. et al. Therapeutic potential of low-dose IL-2 in immune thrombocytopenia: an analysis of 3 cases. *Cytom. B Clin. Cytom.* **94**, 428–433 (2018).
200. Lim, T. Y. et al. Low-dose interleukin-2 for refractory autoimmune hepatitis. *Hepatology* **68**, 1649–1652 (2018).
201. Ehrenstein, M. R. et al. Compromised function of regulatory T cells in rheumatoid arthritis and reversal by anti-TNFα therapy. *J. Exp. Med.* **200**, 277–285 (2004).
202. Nguyen, D. X. & Ehrenstein, M. R. Anti-TNF drives regulatory T cell expansion by paradoxically promoting membrane TNF-TNF-RII binding in rheumatoid arthritis. *J. Exp. Med.* **213**, 1241–1253 (2016).
203. He, X. et al. A TNFR2-agonist facilitates high purity expansion of human low purity treg cells. *PLoS ONE* **11**, e0156311 (2016).
204. Okubo, Y., Mera, T., Wang, L. & Faustman, D. L. Homogeneous expansion of human T-regulatory cells via tumor necrosis factor receptor 2. *Sci. Rep.* **3**, 3153 (2013).
205. Chopra, M. et al. Exogenous TNFR2 activation protects from acute GvHD via host T reg cell expansion. *J. Exp. Med.* **213**, 1881–1900 (2016).
206. Chen, Z. et al. FOXP3 and RORγt: transcriptional regulation of Treg and Th17. *Int. Immunopharmacol.* **11**, 536–542 (2011).
207. Valencia, X. et al. TNF downmodulates the function of human CD4+CD25hi T-regulatory cells. *Blood* **108**, 253–261 (2006).
208. Tseng, W. Y. et al. TNF receptor 2 signaling prevents DNA methylation at the Foxp3 promoter and prevents pathogenic conversion of regulatory T cells. *Proc. Natl Acad. Sci. USA* **116**, 21666–21672 (2019).
209. Santinon, F. et al. Involvement of tumor necrosis factor receptor type II in FoxP3 stability and as a marker of Treg cells specifically expanded by anti-tumor necrosis factor treatments in rheumatoid arthritis. *Arthritis Rheumatol.* **72**, 576–587 (2020).
210. Zhang, C., Zhang, X. & Chen, X. H. Inhibition of the interleukin-6 signaling pathway: a strategy to induce immune tolerance. *Clin. Rev. Allergy Immunol.* **47**, 163–173 (2014).
211. Schinnerling, K., Aguillon, J. C., Catalan, D. & Soto, L. The role of interleukin-6 signalling and its therapeutic blockage in skewing the T cell balance in rheumatoid arthritis. *Clin. Exp. Immunol.* **189**, 12–20 (2017).
212. Thiolat, A. et al. Interleukin-6 receptor blockade enhances CD39+ regulatory T cell development in rheumatoid arthritis and in experimental arthritis. *Arthritis Rheumatol.* **66**, 273–283 (2014).
213. Samson, M. et al. Brief report: inhibition of interleukin-6 function corrects Th17/Treg cell imbalance in patients with rheumatoid arthritis. *Arthritis Rheum.* **64**, 2499–2503 (2012).
214. Kikuchi, J. et al. Peripheral blood CD4(+)CD25(+) CD127(low) regulatory T cells are significantly increased by tocilizumab treatment in patients with rheumatoid arthritis: increase in regulatory T cells correlates with clinical response. *Arthritis Res. Ther.* **17**, 10 (2015).
215. Takatori, H. et al. Helios enhances Treg cell function in cooperation with FoxP3. *Arthritis Rheumatol.* **67**, 1491–1502 (2015).
216. Sheng-Xiao, Z. et al. SAT0181 Low dose interleukin-2 combined with tocilizumab selectively increases regulatory T cells helping refractory rheumatoid arthritis patients achieve remission more rapidly. *Ann. Rheum. Dis.* **76**, 839 (2017).
217. Pasare, C. & Medzhitov, R. Toll pathway-dependent blockade of CD4+CD25+ T cell-mediated suppression by dendritic cells. *Science* **299**, 1033–1036 (2003).
218. Wan, S., Xia, C. & Morel, L. IL-6 produced by dendritic cells from lupus-prone mice inhibits CD4+CD25+ T cell regulatory functions. *J. Immunol.* **178**, 271–279 (2007).
219. Dai, H., He, F., Tsokos, G. C. & Kyttaris, V. C. IL-23 limits the production of IL-2 and promotes autoimmunity in lupus. *J. Immunol.* **199**, 903–910 (2017).
220. Kannan, A. K. et al. IL-23 induces regulatory T cell plasticity with implications for inflammatory skin diseases. *Sci. Rep.* **9**, 17675 (2019).
221. Izcue, A. et al. Interleukin-23 restrains regulatory T cell activity to drive T cell-dependent colitis. *Immunity* **28**, 559–570 (2008).
222. Schiering, C. et al. The alarmin IL-33 promotes regulatory T-cell function in the intestine. *Nature* **513**, 564–568 (2014).
223. Nussbaum, L., Chen, Y. L. & Ogg, G. S. Role of regulatory T cells in psoriasis pathogenesis and treatment. *Br. J. Dermatol.* **184**, 14–24 (2021).
224. van Vollenhoven, R. F. et al. Maintenance of efficacy and safety of ustekinumab through one year in a phase II multicenter, prospective, randomized, double-blind, placebo-controlled crossover trial of patients with active systemic lupus erythematosus. *Arthritis Rheumatol.* **72**, 761–768 (2020).
225. Konrad, M. W. et al. Pharmacokinetics of recombinant interleukin 2 in humans. *Cancer Res.* **50**, 2009–2017 (1990).
226. Charych, D. H. et al. NKTR-214, an engineered cytokine with biased IL2 receptor binding, increased tumor exposure, and marked efficacy in mouse tumor models. *Clin. Cancer Res.* **22**, 680–690 (2016).
227. Trotta, E. et al. A human anti-IL-2 antibody that potentiates regulatory T cells by a structure-based mechanism. *Nat. Med.* **24**, 1005–1014 (2018).
228. Klein, C. et al. Cergutuzumab amunaleukin (CEA-IL2v), a CEA-targeted IL-2 variant-based immunocytokine for combination cancer immunotherapy: overcoming limitations of aldesleukin and conventional IL-2-based immunocytokines. *Oncoimmunology* **6**, e1277306 (2017).
229. van Brummelen, E. M. J. et al. (89)Zr-labeled CEA-targeted IL-2 variant immunocytokine in patients with solid tumors: CEA-mediated tumor accumulation and role of IL-2 receptor-binding. *Oncotarget* **9**, 24737–24749 (2018).
230. Weide, B. et al. A phase II study of the L19IL2 immunocytokine in combination with dacarbazine in advanced metastatic melanoma patients. *Cancer Immunol. Immunother.* **68**, 1547–1559 (2019).
231. Neri, D. & Sondel, P. M. Immunocytokines for cancer treatment: past, present and future. *Curr. Opin. Immunol.* **40**, 96–102 (2016).
232. Padutsch, T. et al. Superior Treg-expanding properties of a novel dual-acting cytokine fusion protein. *Front. Pharmacol.* **10**, 1490 (2019).
233. Peterson, L. B. et al. A long-lived IL-2 mutein that selectively activates and expands regulatory T cells as a therapy for autoimmune disease. *J. Autoimmun.* **95**, 1–14 (2018).
234. Comte, D. et al. Engagement of SLAMF3 enhances CD4+ T-cell sensitivity to IL-2 and favors regulatory T-cell polarization in systemic lupus erythematosus. *Proc. Natl Acad. Sci. USA* **113**, 9321–9326 (2016).
235. Guma, S. R. et al. Natural killer cell therapy and aerosol interleukin-2 for the treatment of osteosarcoma lung metastasis. *Pediatr. Blood Cancer* **61**, 618–626 (2014).
236. Horwitz, D. A., Bickerton, S., Koss, M., Fahmy, T. M. & La Cava, A. Suppression of murine lupus by CD4+ and CD8+ Treg cells induced by T cell-targeted nanoparticles loaded with interleukin-2 and transforming growth factor β. *Arthritis Rheumatol.* **71**, 632–640 (2019).
237. Siddhanti, S. et al. NKTR-358, a novel IL-2 conjugate, stimulates high levels of regulatory T cells in patients with systemic lupus erythematosus [abstract THU0054]. *Ann. Rheum. Dis.* **79**, 238–239 (2020).
238. Pilat, N. et al. Treg-mediated prolonged survival of skin allografts without immunosuppression. *Proc. Natl Acad. Sci. USA* **116**, 13508–13516 (2019).
239. Karakus, U. et al. Receptor-gated IL-2 delivery by an anti-human IL-2 antibody activates regulatory T cells in three different species. *Sci. Transl. Med.* **12**, eabb9283 (2020).
240. Fischer, R. et al. Selective activation of tumor necrosis factor receptor II induces antiinflammatory responses and alleviates experimental arthritis. *Arthritis Rheumatol.* **70**, 722–735 (2018).
241. Laurent, J. et al. T-cell activation by treatment of cancer patients with EMD 521873 (Selectikine), an IL-2/anti-DNA fusion protein. *J. Transl. Med.* **11**, 5 (2013).

Author contributions

The authors contributed equally to all aspects of the article.

Competing interests

D.K. is an inventor of patents belonging to Sorbonne Université that cover the use of IL-2 in autoimmune diseases; he holds shares in ILTOO Pharma. A.G.A.K. and G.C.T. declare no competing interests.

Peer review information

Nature Reviews Rheumatology thanks Z.-G. Li, J. Humrich, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

The online version contains supplementary material available at https://doi.org/10.1038/s41584-021-00707-x.