

PATTERNS, PREDICTIONS, PROGNOSIS
Statistical models in paediatric rheumatologic diseases

Pieter van Dijkhuizen

ISBN: 978-90-393-6939-5

© Pieter van Dijkhuizen, 2018

The copyright of the articles that have been published or accepted for publication has been transferred to the respective journals.

All rights reserved. No part of this thesis may be reproduced, stored in a retrieval system or transmitted in any form or by any means without prior written permission of the author.

Cover: Piet Mondrian, Composition with red, blue and yellow, 1930. Artwork in the public domain.

Cover design, layout and printing: Off Page, Amsterdam, www.offpage.nl

Printing of this thesis was financially supported by the Dutch Arthritis Foundation, Pfizer and Sobi.

PATTERNS, PREDICTIONS, PROGNOSIS
Statistical models in paediatric rheumatologic diseases

PATRONEN, PREDICTIES, PROGNOSE
Statistische modellen voor kinderreumatologische aandoeningen
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van
de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het
college voor promoties in het openbaar te verdedigen op dinsdag 27 februari 2018
des middags te 2.30 uur

door

Evert Hendrik Pieter van Dijkhuizen
geboren op 9 januari 1988 te Amersfoort

Promotoren: Prof.dr. N.M. Wulffraat
Prof.dr. A. Martini

Copromotor: Dr. C. Malattia

Voor mijn ouders

TABLE OF CONTENTS

Chapter 1	General introduction	9
Part I	Juvenile idiopathic arthritis	25
Chapter 2	Early predictors of prognosis in juvenile idiopathic arthritis: A systematic literature review	27
Chapter 3	Prediction of inactive disease in juvenile idiopathic arthritis: A multicenter observational cohort study	53
Chapter 4	Microbiome analytics of the gut microbiota in juvenile idiopathic arthritis patients: An observational, longitudinal cohort study	77
Part II	Juvenile dermatomyositis	119
Chapter 5	Modelling disease activity in juvenile dermatomyositis: A Bayesian approach	121
Chapter 6	Clinical signs and symptoms in a joint model of four disease activity parameters in juvenile dermatomyositis: A prospective, longitudinal, multicenter cohort study	145
Part III	Methotrexate	161
Chapter 7	Prediction of methotrexate efficacy and adverse events in patients with juvenile idiopathic arthritis: A systematic literature review	163
Chapter 8	High prevalence of methotrexate intolerance in juvenile idiopathic arthritis: Development and validation of a methotrexate intolerance severity score	191
Chapter 9	Methotrexate intolerance in oral and subcutaneous administration in patients with juvenile idiopathic arthritis: A cross-sectional, observational study	207
Chapter 10	Prediction of methotrexate intolerance in juvenile idiopathic arthritis: A prospective, observational cohort study	223
Chapter 11	General discussion	243
Chapter 12	Summary	267
	Nederlandse samenvatting	273
	Dankwoord	278
	Curriculum vitae	282
	List of publications	283
Appendix A	Prediction of clinical non-response to methotrexate treatment in juvenile idiopathic arthritis	287

Chapter

GENERAL INTRODUCTION

1

Essentially, all models are wrong, but some are useful
George E.P. Box

Reality is complex. Multiple causes, often seemingly insignificant and easily overlooked, operate in concert to produce an effect. Usually, none of these causes on its own is sufficient to occasion the effect, nor does the combination of multiple causes, even though sufficient, need to be necessary for the effect. Alternative combinations of causes might also be sufficient. For instance, smoking is a cause of lung cancer, however, in itself it is not sufficient to cause a carcinoma.¹ This can be easily verified, since many smokers do not develop lung cancer. Additionally, tobacco smoking is not *necessary* to cause lung cancer either, i.e., lung cancer can develop in non-smokers. Other conditions, such as exposure to environmental pollutants or inflammatory lung processes, are causes as well.¹

These observations led J.L. Mackie to coin the term INUS-condition, an *insufficient* but *non-redundant* part of an *unnecessary* but *sufficient* condition.^{2,3} According to his framework, causes are often complexes of cooperating parts that bring about an effect. Thus, each part is in itself insufficient, yet it is non-redundant in the sense that in its absence the effect would not have occurred. Each complex of cooperating parts is sufficient for the effect, but unnecessary: other complexes of cooperating parts could have produced the effect as well. In our example, smoking is such an INUS-condition.

Yet, in a specific patient with lung cancer, his smoking behaviour is often cited as the cause of lung cancer. What is meant by such a statement? In light of the discussion above, it is not entirely true. Smoking *on its own* did not cause the carcinoma, yet, the remark seeks to communicate that smoking, in the case at hand, was the most important part (and perhaps also the most easily identifiable) of the combination of conditions that led to the development of cancer. In other words, the complexity of reality is reduced to highlight one, or a few, aspects deemed most important in the current case.

The report about lung cancer can be viewed as a model of reality. It simplifies reality in order to highlight one or a few important aspects of it. As such, according to the aphorism quoted at the start of this chapter, it is wrong but useful.⁴ Wrong, because it is a simplification of reality and mentions as the cause something which is only an INUS-condition. Useful, because it focuses attention on key aspects. It is important to note, however, that it is precisely the 'wrongness' of the model that makes it useful. It is the simplification of reality that lets us understand and target one aspect of it, while temporarily setting aside other aspects.

Of all areas of reality, human physiology and pathophysiology is probably one of the most complex. Numerous are the inter- and intra-cellular pathways and the connections between them and many of these have not yet been elucidated. Therefore, disease models must necessarily be a simplification of the complexity of reality and are, thus, wrong in this respect. Yet, they are helpful. These models help us focus on key points, which can form the basis of interventions. The identification of INUS-conditions involved

in the pathogenesis of disease is the first step in the process of developing treatments against it, or of stratifying patients based on the presence or absence of risk factors. To be sure, knowledge of complete disease pathways (i.e., sufficient complexes of INUS-conditions) would be of great help, but we need not despair in the absence of such knowledge. Knowledge of only one or a few INUS-conditions may be sufficient to institute treatment or even prevention. For example, quitting smoking greatly reduces the risk of lung cancer and cardio-vascular disease,^{5;6} even though smoking is 'just' an INUS-condition for these diseases. The reason why this is so can be easily seen using Mackie's framework.^{2;3} Elimination of an INUS-condition leads to blockade of at least one sufficient complex of parts, thus interrupting the pathway from cause to effect.

All of the above is true in particular in the case of auto-immune diseases. This is a very heterogeneous group of diseases, all of which are characterised by a disruption of the immune balance towards decreased tolerance of self-antigens. The aetiology of auto-immune disorders is still unknown, but various contributing factors have been studied and demonstrated to play a role, among which genetic predisposition and environmental factors.⁷⁻¹² The causes of auto-immune disorders interfere with the immune system in such a way that the balance between pro-inflammatory and anti-inflammatory pathways is disrupted, thus leading to inflammation.^{7;13-16} It is, however, unclear how these general mechanisms lead to the specific manifestations characteristic of the various auto-immune diseases.

Against this backdrop, the development of disease models for auto-immune diseases is of great importance, as they may aid in the identification of predisposing factors or in the classification of patients into subgroups of disease. Thus, they could provide opportunities to identify people at risk to develop disease or to develop new therapeutic strategies. The models may also elucidate factors associated with a worse prognosis, which could guide clinical decisions regarding starting and stopping treatment.

Such models will be useful, even though they might be wrong. Wrong in the abovementioned sense that they will inevitably simplify reality. Moreover, there is an additional way in which the models may be incorrect: The models may point towards statistical associations, which are merely that, without underlying causal relationship.^{17;18} In other words, to find an association between a clinical parameter and an auto-immune disorder (or its prognosis) does not mean that the disorder is caused by this factor, not even in the weak sense of that factor being an INUS-condition for the disease. It might be that both the clinical factor and the onset of the auto-immune disorder are being caused by a third condition. Nonetheless, even though this observation is important when talking about causes of disease, it is less significant with respect to prognosis, as long as the observed relationship is maintained across populations and is not the immediate target of treatment.¹⁹

This thesis focuses on statistical models in the context of paediatric rheumatologic diseases. The next sections present an introduction to these diseases as well as an outline of the thesis.

JUVENILE IDIOPATHIC ARTHRITIS

Juvenile idiopathic arthritis (JIA) is a group of auto-immune disorders characterised by chronic arthritis. It is defined as arthritis of unknown cause, lasting for more than six weeks, with an onset before 16 years of age.^{7,20} All known causes of arthritis, such as septic arthritis or reactive arthritis, but also haemato-oncological disorders have to be excluded before making a diagnosis of JIA. It is the most common paediatric rheumatologic disorder with a prevalence of about 16-150 per 100,000 children.²⁰ Clinically, JIA is subdivided in seven categories, based on the number of joints involved, rheumatoid factor positivity and the presence of extra-articular symptoms, such as enthesitis, psoriasis or systemic features.²¹

As discussed above, like any other auto-immune disorder, the aetiology of JIA is unknown. A multifactorial cause is hypothesised, in which genetic predisposition as well as environmental factors play a role.^{7,8} The best-known genetic association with respect to JIA is the one between the category of enthesitis-related arthritis and human leucocyte antigen (HLA) B27, but other associations have been found as well, most notably between the oligoarticular category and HLA A2, DR5 and DR8. Rheumatoid factor negative polyarticular JIA is associated with HLA DRB1*08 and DPB1*03 and rheumatoid factor positive polyarticular JIA is associated with HLA DRB1*04, DQA1*03 and DQB1*03.⁸⁻¹² Next to HLA associations, various associations with non-HLA loci have been found.¹⁰⁻¹²

However, given that the concordance between monozygotic twins is in the 25-40% range,⁸ environmental factors must play a role in the aetiology of JIA as well. Potential contributors include pathological viruses and bacteria as well as commensal bacteria.^{7,15,16} This idea is reinforced by the finding that antibiotic administration in children was associated with a higher risk of developing JIA. A potential explanation for this finding is that some strains of commensal micro-organisms that are protective against the development of autoimmunity are eradicated upon antibiotic administration. However, much is unknown yet about the role of environmental factors in the development of auto-immune diseases. A possible approach to address this knowledge gap is by demonstrating, first, that the exposition to these environmental factors (e.g., gut microbiota) differs between patients and healthy subjects, and, secondly, that these differences play a role in the pathogenesis of autoimmune diseases.

As a consequence of the unknown cause of JIA, factors defining various subgroups with a milder or more severe prognosis are unknown as well. Thus, some patients with oligoarthritis may enter disease remission after treatment with relatively mild drugs, whereas others experience disease extension beyond four joints and develop extended oligoarthritis, a disease category mimicking rheumatoid factor negative polyarthritis.²¹ More in general, even though the JIA categories reflect some differences in burden of disease, they do not adequately distinguish between patients with a differing prognosis, having distinct treatment requirements.

This lack of information is clinically highly relevant, since it impedes the swift and adequate determination of the right therapy for the right patient. Nowadays, a range of potent medications is available for JIA patients, ranging from local-acting corticosteroid joint injections to biological agents, targeting specific inflammatory molecules, such as tumour necrosis factor (TNF)- α , interleukin (IL)-1 and IL-6. Current treatment protocols follow a step-up approach, in which drugs are substituted by the next line of therapy in case of treatment failure.²² However, this may lead to an important delay in the administration of the right therapy and exposes the child to the side effects of various ineffective medications. This is especially important given the window-of-opportunity, which is, roughly, the idea that treatment administered early in the course of disease is more efficacious than later treatment. This effect has been observed in rheumatoid arthritis (RA)²³⁻²⁸ and may exist in JIA as well.²⁹⁻³¹ However, immediate prescription of biological agents without prior failure of at least one synthetic disease-modifying anti-rheumatic drug (DMARD) such as methotrexate (MTX) is precluded by legislation in many countries and might not be cost-effective, either.³² Moreover, it is unnecessary for a substantial group of children with JIA who respond to injections with or without MTX.³³⁻³⁵

In light of the above, there is a clinical need of factors that are able to classify patients into subgroups with different prognosis. Treatment plans can then be adapted accordingly and tailored to the patient's needs. Some patients could be treated with and expected to respond to mild therapy, whereas others could be managed with more aggressive treatment right from the onset of disease, leading to better outcomes and an overall reduction in the exposure to unnecessary and ineffective drugs and, hence, a reduction in healthcare related costs.

JUVENILE DERMATOMYOSITIS

Juvenile dermatomyositis (JDM) is a rare inflammatory disorder with an incidence of about 1.9-4.1 per million children per year.³⁶ It is part of a larger spectrum of diseases, called the juvenile idiopathic inflammatory myopathies (IIM), of which it is the most common.^{37,38} JDM is a systemic vasculopathy, clinically characterised by symmetric proximal muscle weakness, as well as skin involvement, including Gottron's papules, photosensitive heliotrope and malar rash, rashes over the large and small joints and other, less common manifestations. Capillaries in the nail fold can be examined and appear abnormal. Severe vasculopathy manifests itself by gastrointestinal or pulmonary involvement, potentially leading to ulceration and perforation. Various overlap forms with other auto-immune disorders have been described, including JIA, systemic lupus erythematosus, systemic sclerosis and localised scleroderma. Long-term effects of JDM include muscle and skin atrophy, calcinosis and lipoatrophy.³⁷

The prognosis of children with JDM is variable. About 24-40% of patients experience a monocyclic course and recover within 2 years with adequate therapy,

whereas the remainder experiences a chronic course. The latter group is more at risk to develop the more severe, systemic manifestations of the disease. The mortality rate is about 2-3%.³⁷

Taken together, despite its rarity, JDM represents a large burden of disease, due to its chronicity and invalidating manifestations.

Similar to the case made for JIA, classification of JDM patients into subsets with shared clinical phenotype and prognosis is important. Over the last couple of years, it was demonstrated that the presence or absence of various autoantibodies distinguishes between clinical subgroups of juvenile IIM patients.^{39,40} The most frequent of the myositis-specific antibodies is anti-transcription intermediary factor 1 γ (anti-TIF-1 γ , also known as anti-p155/140), followed by anti-nuclear matrix protein 2 (anti-NXP-2, also known as anti-MJ).³⁹⁻⁴² Patients with these antibodies are mostly white, relatively young at onset and present with mild to moderate disease. They also present cutaneous symptoms significantly more often, compared to patients with other antibodies. These antibodies are therefore associated with JDM.⁴⁰ Of the myositis-associated antibodies, anti-Ro is the most frequent in the juvenile IIM. Approximately 20-30% of juvenile IIM patients are autoantibody negative.

Less is known about the prognosis of JDM patients and potential dependencies of the prognosis on the antibody pattern. Like for the overall group of juvenile IIM patients, anti-TIF-1 γ and anti-NXP-2 are the most frequent myositis-specific antibodies, anti-Ro is the most frequent myositis-associated antibody and about 20-30% of patients is autoantibody negative.^{39,40} JDM patients with the anti-melanoma differentiation-associated gene 5 (anti-MDA5) antibody were found to have milder muscular disease activity, but more severe skin involvement, as well as arthritis.⁴¹ In another study, the subgroup of anti-Mi-2 positive patients, despite having worse findings in the muscle biopsy, had a higher probability of being off medication, taken as a surrogate marker for lower disease activity.⁴² Not surprisingly, worse findings in the muscle biopsy, indicated by higher global muscle pathology scores or total biopsy scores predicted a lower probability of being off treatment at 5 years after diagnosis. Taken together, these findings might point to the fact that patients with anti-Mi-2 antibodies, despite having higher disease activity at baseline, are more responsive to therapy.

Thus, the juvenile IIM can be classified into distinct subgroups based on their antibody patterns. One of these subgroups corresponds by and large to the clinical diagnosis of JDM. Nevertheless, even the prognosis of JDM patients is variable and it is currently impossible to predict this prognosis accurately, hampering the prescription of medication tailored to the patient, like in the case of JIA.

METHOTREXATE

Methotrexate (MTX) is a cornerstone in the treatment of both JIA and JDM.^{22,43,44} Developed as a cytostatic drug, it was discovered that in dosages of 10-15 mg/m²/week it is effective as a potent anti-inflammatory agent.^{33,34} It is usually administered orally or subcutaneously. The bioavailability of MTX administered orally is lower in comparison with subcutaneous administration and some think this leads to a lower efficacy.⁴⁵⁻⁴⁸ However, this has not been demonstrated clinically.^{49,50}

The exact mechanism of action of MTX as an anti-inflammatory agent is not fully understood. What is known is that MTX is transported into the cell by reduced folate carrier (RFC), also known as solute carrier (SLC) 19A1.⁵¹⁻⁵³ Once inside the cell, a glutamate tail of various length is added by folylpolyglutamate synthetase (FPGS), contributing to the cellular retention of MTX.⁵⁴⁻⁵⁸ Polyglutamated MTX exerts various actions inside the cell by inhibiting dihydrofolate reductase (DHFR) and thymidylate synthase (TS), enzymes involved in folate metabolism, leading to a decrease in purine and pyrimidine synthesis and a down regulation of cytokine release pathways. It was once thought that these actions led to a decreased proliferation of pro-inflammatory cells, thus explaining the anti-inflammatory action of MTX.⁵⁹ However, no depletion of cytotoxic T lymphocytes or indeed any other pro-inflammatory cell line was observed in patients treated with MTX. It is now thought that the anti-inflammatory actions of MTX are mediated by an increase of adenosine, itself a potent anti-inflammatory agent. Adenosine inhibits release of toxic oxygen metabolites and of pro-inflammatory cytokines such as TNF- α , while promoting the release of anti-inflammatory cytokines such as IL-10.⁵⁹

Polyglutamation of MTX is reversed by γ -glutamyl hydrolase (GGH).⁵⁸ Thus, the intracellular action of MTX is mediated by a balance between polyglutamation by FPGS and depolyglutamation by GGH. Depolyglutamated MTX is excreted from the cell by members of the adenosine triphosphate binding cassette (ABC).⁵¹

The discussion above makes clear that resistance to the actions of MTX can occur at multiple levels, such as by way of impaired uptake, decreased polyglutamation, increased depolyglutamation and increased efflux, as well as resistance of MTX target genes to its actions.⁵⁸ Indeed, about one third of JIA patients is non-responsive to the drug, necessitating them to switch to more costly biological agents.⁴⁴ Parallel to the discussion described earlier, regarding the need to start effective treatment as soon as possible in the course of disease, it is key to know in an early stage, and ideally even before starting the drug, the probability of response to the drug. Thus, two sets of predictors are needed: one to estimate the severity of disease and consequent need to start aggressive treatment or not; the other to estimate the probability of response to treatment, independent of the severity of disease. A patient with mild (predicted) disease severity might need more aggressive therapy, because of an individually determined resistance to MTX.

Despite being a well-tolerated drug, MTX treatment is associated with side effects. The best-known adverse effect is a temporary increase of the liver transaminases, which

usually normalises after suspension of treatment.⁶⁰ Moreover, addition of a low dose of folic acid, the natural antagonist to MTX, 24 hours after MTX administration, reduces the frequency of this side effect.^{61;62} Other adverse effects are related to the gastrointestinal tract and include nausea, abdominal pain and vomiting. Even though not serious, these gastrointestinal adverse effects potentially have a large impact on the quality of life, sometimes leading to reluctance to take the drug.⁶³ Furthermore, some patients develop an anticipatory or associative component, thus experiencing these side effects even before MTX administration or when being reminded of MTX, for example by its yellow colour.^{64;65} This complex of symptoms is termed MTX intolerance. Occurrence of MTX intolerance frequently leads to discontinuation of the drug, notwithstanding its efficacy, and substitution by a biological agent, rendering it an important clinical problem. Here too, prior prediction of the risk of MTX intolerance is helpful in its management, for example by starting anti-emetic agents or offering psychological support to deal with the anticipatory and associative component of MTX intolerance.^{66;67}

OUTLINE OF THE THESIS

This thesis attempts to provide answers to some of the questions highlighted in the introduction. More specifically, it concerns models of JIA and JDM disease activity, as well as MTX adverse effects. Thus, the thesis has been divided into three parts. **Part I** deals with JIA. In **Chapter 2** the literature is systematically reviewed regarding existing knowledge about predictors of the prognosis of JIA and conclusions are drawn regarding the feasibility of predicting JIA prognosis using knowledge available in literature. Next, we add to this knowledge by attempting to create a prediction model for the prognosis of JIA in a large, well-defined, international cohort of JIA patients, using clinical, imaging, microbiota and immunological data as potential predictors. The results of this study are discussed in **Chapter 3**.

Data from the same study are also used to contribute to the problem of JIA aetiology. As discussed above, potential environmental factors responsible for the development of JIA have not yet been elucidated. It has been hypothesised that commensal microorganisms, so-called microbiota might play a role. To demonstrate that, a first step is to show that the microbiota composition is different or behaves differently in patients compared to healthy people. Therefore, in **Chapter 4** the gut microbiota composition of Italian and Dutch JIA patients is assessed and compared to the microbiota composition in healthy controls.

In **Part II**, the attention is shifted towards JDM. **Chapters 5 and 6** discuss the development and results of a disease activity model in this chronic disease. Using a large cohort of JDM patients, followed prospectively, clinical determinants of disease activity are assessed. **Chapter 5** focuses on the statistical intricacies of model

development, whereas **Chapter 6** concerns the results of the model and places them in a clinical context.

Finally, **Part III** deals with MTX. Known predictors of MTX efficacy and adverse effects are reviewed in **Chapter 7**. The other chapters in this part of the thesis concern MTX adverse effects, specifically MTX intolerance. The outline of the problem and its associations with the route of administration are sketched in **Chapter 8** for a Dutch cohort of patients and in **Chapter 9** for a German cohort. **Chapter 10** discusses the development of a prediction model for the occurrence of MTX intolerance.

Finally, in **Chapter 11** the results of this thesis are critically discussed and interpreted. Furthermore, implications for patient care and future research are considered.

REFERENCE LIST

1. McIntyre A, Ganti AK. Lung cancer-A global perspective. *J Surg Oncol* 2017; 115(5):550-554.
2. Mackie JL. Causes and Conditions. In: Loux MJ, editor. *Metaphysics: Contemporary readings*. 2nd ed. New York: Routledge; 2008. 301-325.
3. Loux MJ. *Metaphysics: A contemporary introduction*. 3rd ed. New York: Routledge; 2006.
4. Box GEP, Draper NR. *Empirical model-building and response surfaces*. New York: John Wiley & Sons; 1987.
5. Jha P, Ramasundarahettige C, Landsman V, Rostron B, Thun M, Anderson RN et al. 21st-century hazards of smoking and benefits of cessation in the United States. *N Engl J Med* 2013; 368(4):341-350.
6. Anthonisen NR, Skeans MA, Wise RA, Manfreda J, Kanner RE, Connett JE. The effects of a smoking cessation intervention on 14.5-year mortality: a randomized clinical trial. *Ann Intern Med* 2005; 142(4):233-239.
7. Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. *Lancet* 2011; 377(9783):2138-2149.
8. Rigante D, Bosco A, Esposito S. The Etiology of Juvenile Idiopathic Arthritis. *Clin Rev Allergy Immunol* 2015; 49(2):253-261.
9. Hinks A, Bowes J, Cobb J, Ainsworth HC, Marion MC, Comeau ME et al. Fine-mapping the MHC locus in juvenile idiopathic arthritis (JIA) reveals genetic heterogeneity corresponding to distinct adult inflammatory arthritic diseases. *Ann Rheum Dis* 2016.
10. Hinks A, Cobb J, Marion MC, Prahalad S, Sudman M, Bowes J et al. Dense genotyping of immune-related disease regions identifies 14 new susceptibility loci for juvenile idiopathic arthritis. *Nat Genet* 2013; 45(6):664-669.
11. Hinks A, Martin P, Thompson SD, Sudman M, Stock CJ, Thomson W et al. Autoinflammatory gene polymorphisms and susceptibility to UK juvenile idiopathic arthritis. *Pediatr Rheumatol Online J* 2013; 11(1):14.
12. Hinks A, Cobb J, Sudman M, Eyre S, Martin P, Flynn E et al. Investigation of rheumatoid arthritis susceptibility loci in juvenile idiopathic arthritis confirms high degree of overlap. *Ann Rheum Dis* 2012; 71(7):1117-1121.
13. 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* 1995; 155(3):1151-1164.
14. Hori S, Takahashi T, Sakaguchi S. Control of autoimmunity by naturally arising regulatory CD4+ T cells. *Adv Immunol* 2003; 81:331-371.
15. Christen U, von Herrath MG. Initiation of autoimmunity. *Curr Opin Immunol* 2004; 16(6):759-767.
16. Lin YT, Wang CT, Gershwin ME, Chiang BL. The pathogenesis of oligoarticular/polyarticular vs systemic juvenile idiopathic arthritis. *Autoimmun Rev* 2011; 10(8):482-489.
17. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer; 2009.
18. Steyerberg EW, Harrell Jr FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JD. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology* 2001; 54(8):774-781.
19. Grobbee DE, Hoes AW. *Clinical epidemiology: Principles, methods, and applications for clinical research*. 2nd ed. Burlington: Jones & Bartlett Learning; 2015.
20. Ravelli A, Martini A. Juvenile idiopathic arthritis. *Lancet* 2007; 369(9563):767-778.
21. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004; 31(2):390-392.

22. Beukelman T, Patkar NM, Saag KG, Tolleson-Rinehart S, Cron RQ, DeWitt EM et al. 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: initiation and safety monitoring of therapeutic agents for the treatment of arthritis and systemic features. *Arthritis Care Res (Hoboken)* 2011; 63(4):465-482.
23. Anderson JJ, Wells G, Verhoeven AC, Felson DT. Factors predicting response to treatment in rheumatoid arthritis: the importance of disease duration. *Arthritis Rheum* 2000; 43(1):22-29.
24. Boers M. Understanding the window of opportunity concept in early rheumatoid arthritis. *Arthritis Rheum* 2003; 48(7):1771-1774.
25. Moreland LW, Bridges SL, Jr. Early rheumatoid arthritis: a medical emergency? *Am J Med* 2001; 111(6):498-500.
26. Mottonen T, Hannonen P, Korpela M, Nissila M, Kautiainen H, Ilonen J et al. Delay to institution of therapy and induction of remission using single-drug or combination-disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. *Arthritis Rheum* 2002; 46(4):894-898.
27. O'Dell JR. Treating rheumatoid arthritis early: a window of opportunity? *Arthritis Rheum* 2002; 46(2):283-285.
28. Quinn MA, Emery P. Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. *Clin Exp Rheumatol* 2003; 21(5 Suppl 31):S154-S157.
29. Albers HM, Wessels JA, van der Straaten RJ, Brinkman DM, Suijlekom-Smit LW, Kamphuis SS et al. Time to treatment as an important factor for the response to methotrexate in juvenile idiopathic arthritis. *Arthritis Rheum* 2009; 61(1):46-51.
30. Tynjala P, Vahasalo P, Tarkiainen M, Kroger L, Aalto K, Malin M et al. Aggressive combination drug therapy in very early polyarticular juvenile idiopathic arthritis (ACUTE-JIA): a multicentre randomised open-label clinical trial. *Ann Rheum Dis* 2011; 70(9):1605-1612.
31. Wallace CA, Giannini EH, Spalding SJ, Hashkes PJ, O'Neil KM, Zeft AS et al. Trial of early aggressive therapy in polyarticular juvenile idiopathic arthritis. *Arthritis Rheum* 2012; 64(6):2012-2021.
32. Luca NJ, Burnett HF, Ungar WJ, Moretti ME, Beukelman T, Feldman BM et al. Cost-Effectiveness Analysis of First-Line Treatment With Biologic Agents in Polyarticular Juvenile Idiopathic Arthritis. *Arthritis Care Res (Hoboken)* 2016; 68(12):1803-1811.
33. Ruperto N, Murray KJ, Gerloni V, Wulffraat N, deOliveira SK, Falcini F et al. A randomized trial of parenteral methotrexate comparing an intermediate dose with a higher dose in children with juvenile idiopathic arthritis who failed to respond to standard doses of methotrexate. *Arthritis Rheum* 2004; 50(7):2191-2201.
34. Giannini EH, Brewer EJ, Kuzmina N, Shaikov A, Maximov A, Vorontsov I et al. Methotrexate in resistant juvenile rheumatoid arthritis. Results of the U.S.A.-U.S.S.R. double-blind, placebo-controlled trial. The Pediatric Rheumatology Collaborative Study Group and The Cooperative Children's Study Group. *N Engl J Med* 1992; 326(16):1043-1049.
35. Ravelli A, Davi S, Bracciolini G, Pistorio A, Consolaro A, van Dijkhuizen EH et al. Intra-articular corticosteroids versus intra-articular corticosteroids plus methotrexate in oligoarticular juvenile idiopathic arthritis: a multicentre, prospective, randomised, open-label trial. *Lancet* 2017; 389(10072):909-916.
36. Ravelli A, Trail L, Ferrari C, Ruperto N, Pistorio A, Pilkington C et al. Long-term outcome and prognostic factors of juvenile dermatomyositis: a multinational, multicenter study of 490 patients. *Arthritis Care Res (Hoboken)* 2010; 62(1):63-72.
37. Rider LG, Katz JD, Jones OY. Developments in the classification and treatment of the juvenile idiopathic inflammatory myopathies. *Rheum Dis Clin North Am* 2013; 39(4):877-904.
38. Nistala K, Wedderburn LR. Update in juvenile myositis. *Curr Opin Rheumatol* 2013; 25(6):742-746.

39. Shah M, Mamyrova G, Targoff IN, Huber AM, Malley JD, Rice MM et al. The clinical phenotypes of the juvenile idiopathic inflammatory myopathies. *Medicine (Baltimore)* 2013; 92(1):25-41.
40. Rider LG, Shah M, Mamyrova G, Huber AM, Rice MM, Targoff IN et al. The myositis autoantibody phenotypes of the juvenile idiopathic inflammatory myopathies. *Medicine (Baltimore)* 2013; 92(4):223-243.
41. Tansley SL, Betteridge ZE, Gunawardena H, Jacques TS, Owens CM, Pilkington C et al. Anti-MDA5 autoantibodies in juvenile dermatomyositis identify a distinct clinical phenotype: a prospective cohort study. *Arthritis Res Ther* 2014; 16(4):R138.
42. Deakin CT, Yasin SA, Simou S, Arnold KA, Tansley SL, Betteridge ZE et al. Muscle Biopsy Findings in Combination With Myositis-Specific Autoantibodies Aid Prediction of Outcomes in Juvenile Dermatomyositis. *Arthritis Rheumatol* 2016; 68(11):2806-2816.
43. Ruperto N, Pistorio A, Oliveira S, Zulian F, Cuttica R, Ravelli A et al. Prednisone versus prednisone plus ciclosporin versus prednisone plus methotrexate in new-onset juvenile dermatomyositis: a randomised trial. *Lancet* 2016; 387(10019):671-678.
44. Bulatovic Calasan M, Wulffraat NM. Methotrexate in juvenile idiopathic arthritis: towards tailor-made treatment. *Expert Rev Clin Immunol* 2014; 10(7):843-854.
45. Islam MS, Haq SA, Islam MN, Azad AK, Islam MA, Barua R et al. Comparative efficacy of subcutaneous versus oral methotrexate in active rheumatoid arthritis. *Mymensingh Med J* 2013; 22(3):483-488.
46. Teresi ME, Crom WR, Choi KE, Mirro J, Evans WE. Methotrexate bioavailability after oral and intramuscular administration in children. *J Pediatr* 1987; 110(5):788-792.
47. Tukova J, Chladek J, Nemcova D, Chladkova J, Dolezalova P. Methotrexate bioavailability after oral and subcutaneous administration in children with juvenile idiopathic arthritis. *Clin Exp Rheumatol* 2009; 27(6):1047-1053.
48. Alsufyani K, Ortiz-Alvarez O, Cabral DA, Tucker LB, Petty RE, Malleson PN. The role of subcutaneous administration of methotrexate in children with juvenile idiopathic arthritis who have failed oral methotrexate. *J Rheumatol* 2004; 31(1):179-182.
49. Klein A, Kaul I, Foeldvari I, Ganser G, Urban A, Horneff G. Efficacy and safety of oral and parenteral methotrexate therapy in children with juvenile idiopathic arthritis: an observational study with patients from the German Methotrexate Registry. *Arthritis Care Res (Hoboken)* 2012; 64(9):1349-1356.
50. Moncrieffe H, Bennett MF, Tsoras M, Luyrink LK, Johnson AL, Xu H et al. Transcriptional profiles of JIA patient blood with subsequent poor response to methotrexate. *Rheumatology (Oxford)* 2017.
51. de Rotte MC, Bulatovic M, Heijstek MW, Jansen G, Heil SG, van Schaik RH et al. ABCB1 and ABCC3 gene polymorphisms are associated with first-year response to methotrexate in juvenile idiopathic arthritis. *J Rheumatol* 2012; 39(10):2032-2040.
52. van der Heijden JW, Dijkmans BA, Scheper RJ, Jansen G. Drug Insight: resistance to methotrexate and other disease-modifying antirheumatic drugs--from bench to bedside. *Nat Clin Pract Rheumatol* 2007; 3(1):26-34.
53. Malik F, Ranganathan P. Methotrexate pharmacogenetics in rheumatoid arthritis: a status report. *Pharmacogenomics* 2013; 14(3):305-314.
54. Becker ML, van HL, Gaedigk R, Lasky A, Hoeltzel M, Stobaugh J et al. Analysis of intracellular methotrexate polyglutamates in patients with juvenile idiopathic arthritis: effect of route of administration on variability in intracellular methotrexate polyglutamate concentrations. *Arthritis Rheum* 2010; 62(6):1803-1812.
55. Becker ML, Gaedigk R, Van Haandel L, Thomas B, Lasky A, Hoeltzel M et al. The effect of genotype on methotrexate polyglutamate variability in juvenile idiopathic arthritis and association with drug response. *Arthritis Rheum* 2011; 63(1):276-285.
56. Bulatovic Calasan M, Den Boer E, de Rotte MC, Vastert SJ, Kamphuis S, De Jonge R et al.

- Methotrexate polyglutamates in erythrocytes are associated with lower disease activity in juvenile idiopathic arthritis patients. *Ann Rheum Dis* 2015; 74(2):402-407.
57. de Rotte MC, Den Boer E, de Jong PH, Pluijm SM, Bulatovic Calasan M, Weel AE et al. Methotrexate polyglutamates in erythrocytes are associated with lower disease activity in patients with rheumatoid arthritis. *Ann Rheum Dis* 2015; 74(2):408-414.
 58. Kremer JM. Toward a better understanding of methotrexate. *Arthritis Rheum* 2004; 50(5):1370-1382.
 59. Cronstein BN. Molecular therapeutics. Methotrexate and its mechanism of action. *Arthritis Rheum* 1996; 39(12):1951-1960.
 60. Ortiz-Alvarez O, Morishita K, Avery G, Green J, Petty RE, Tucker LB et al. Guidelines for blood test monitoring of methotrexate toxicity in juvenile idiopathic arthritis. *J Rheumatol* 2004; 31(12):2501-2506.
 61. Alarcon GS, Morgan SL. Folinic acid to prevent side effects of methotrexate in juvenile rheumatoid arthritis. *J Rheumatol* 1996; 23(12):2184-2185.
 62. Ravelli A, Migliavacca D, Viola S, Ruperto N, Pistorio A, Martini A. Efficacy of folinic acid in reducing methotrexate toxicity in juvenile idiopathic arthritis. *Clin Exp Rheumatol* 1999; 17(5):625-627.
 63. Brunner HI, Johnson AL, Barron AC, Passo MH, Griffin TA, Graham TB et al. Gastrointestinal symptoms and their association with health-related quality of life of children with juvenile rheumatoid arthritis: validation of a gastrointestinal symptom questionnaire. *J Clin Rheumatol* 2005; 11(4):194-204.
 64. van der Meer A, Wulffraat NM, Prakken BJ, Gijsbers B, Rademaker CM, Sinnema G. Psychological side effects of MTX treatment in juvenile idiopathic arthritis: a pilot study. *Clin Exp Rheumatol* 2007; 25(3):480-485.
 65. Murray KJ, Lovell DJ. Advanced therapy for juvenile arthritis. *Best Pract Res Clin Rheumatol* 2002; 16(3):361-378.
 66. Duff AJ. Incorporating psychological approaches into routine paediatric venepuncture. *Arch Dis Child* 2003; 88(10):931-937.
 67. Figueroa-Moseley C, Jean-Pierre P, Roscoe JA, Ryan JL, Kohli S, Palesh OG et al. Behavioral interventions in treating anticipatory nausea and vomiting. *J Natl Compr Canc Netw* 2007; 5(1):44-50.

Part

JUVENILE IDIOPATHIC ARTHRITIS



Chapter

2

EARLY PREDICTORS OF PROGNOSIS IN JUVENILE IDIOPATHIC ARTHRITIS: A SYSTEMATIC LITERATURE REVIEW

E. H. Pieter van Dijkhuizen and Nico M. Wulffraat

ABSTRACT

Objectives

Juvenile idiopathic arthritis (JIA) is subdivided in seven categories. Even within these categories, the prognosis varies markedly. To start appropriate treatment in patients with JIA and to inform patients and their parents correctly, it is essential to know the individual prognosis, preferably at the time of diagnosis. The aim of this study was to identify variables, which predict disease activity, joint damage, functional ability and quality of life (QoL) early in the disease course.

Methods

A systematic literature review was performed and 3,679 articles were identified. The results were screened and critically appraised, using predefined criteria. Articles that described validated outcomes, such as the Wallace criteria, the childhood health-assessment questionnaire (CHAQ) and the juvenile arthritis damage index (JADI) and that determined predictors in the first 6 months of disease were selected.

Results

Forty mostly retrospective articles were selected. Polyarticular onset predicted a worse prognosis for all outcomes, except QoL. A diagnostic delay and the systemic category predicted continuation of active disease. Notably, antinuclear antibodies (ANA) did not predict disease activity. Symmetric involvement and rheumatoid factor positivity predicted more damage. More disease activity was mainly associated with worse functional outcome. However, most predictors were not validated.

Conclusions

Few predictors for the selected outcomes were found. Prospective, longitudinal studies using standardised outcome measurements, and evaluating a broader range of predictors, such as genetics, immunologic and imaging data, should be performed. For the outcomes joint assessment and quality of life, standardised and validated outcomes should be developed.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is characterised by joint inflammation of unknown aetiology, lasting for at least 6 weeks with an onset before 16 years of age.¹ It is classified in seven categories, based on the number of joints involved in the first 6 months and the presence of extra-articular symptoms, such as fever, rash, enthesitis or psoriasis.^{1,2} JIA is heterogeneous, because of differing symptoms and a varying prognosis even within the same category, ranging from a short and modest inflammation of a single joint to long-lasting, debilitating inflammation of multiple joints, leading to joint destruction and life-long incapacity.³⁻⁶

Nowadays, a range of effective medication stands at the physician's disposal, from non-steroidal anti-inflammatory drugs (NSAIDs) and steroidal intra-articular injections, to disease-modifying anti-rheumatic drugs (DMARDs) like methotrexate, and biologicals such as etanercept and adalimumab.⁷⁻¹³ In adult rheumatology, early and effective treatment in the so-called window of opportunity improves the outcome.¹⁴⁻¹⁹ In paediatric rheumatology, the time to initiation of methotrexate therapy was shown to be an important factor for the response.²⁰ In a recent study of early aggressive therapy in polyarticular JIA, fairly high response rates were achieved,²¹ while early treatment of systemic JIA with anti-IL1 therapy was very successful and reduced the use of steroids markedly.²² It seems therefore likely that starting early and aggressive therapy in children with JIA can prevent long-term sequelae. On the other hand, it is desirable to spare milder cases aggressive treatment, for both safety and economic reasons. Consequently, it is essential to know the prognosis for the individual patient early in the course of disease and preferentially at the time of diagnosis, in order to start the right treatment immediately. Furthermore, patients and their parents not only want to know what kind of disease JIA is in general, but especially how it will affect their personal lives and prospects, for which, too, it is essential to know the individual prognosis.

Much effort has already been done to elucidate predictors of the prognosis. These studies have been reviewed by Adib *et al.* in 2005.^{3,4} Major finding at that time was the variability among studies owing to a lack of standardised criteria. Since then, however, the Wallace criteria for disease remission have been developed and validated,²³⁻²⁶ potentially leading to a reduced variability among studies and the ability to draw consistent conclusions.

Therefore, the aim of this study was to systematically review the literature for studies that determine predictors for the prognosis within 6 months after the diagnosis of JIA. We focused on the outcome measurements disease activity, joint damage, functional ability and quality of life.

METHODS

On 3 October 2013, a systematic literature search was performed in PubMed, Embase, The Cochrane Library and PsycInfo. Using the algorithm in table 1, 3,679 articles were identified (figure 1). All titles and abstracts were screened for applicability to the research subject, using predefined criteria. Studies that evaluated systemic JIA only were excluded, because this category is thought to have a different pathogenesis, warranting a separate review.^{1;27-31} Likewise, studies only assessing predictors for the development of uveitis were excluded.

Articles that seemed relevant were selected for in-depth full-text screening. If a full-text could not be retrieved, the author was contacted. The references of selected articles were screened as well. If a relevant article had been missed, the title and abstract of that article were screened to identify the relevant key words. Those were then added to the search algorithm and the search was repeated, until no more relevant articles were found. Doing so, the articles were screened several times. Table 1 shows the final search algorithm.

Selected articles were critically appraised using pre-defined criteria (see supplementary table 1). In order to reduce the expected variability among the articles,⁴ the outcome definition and the time of determination of the predictors were emphasised. Outcome measurement regarding disease activity was focused on the validated Wallace criteria for clinical remission.²³⁻²⁶ Studies in which remission was defined much like the Wallace

Table 1. Search strategy^a

	Search algorithm	PubMed ^b	Embase ^b	Cochrane ^c	PsycInfo ^b
#1	"juvenile idiopathic arthritis" OR "juvenile chronic arthritis" OR "juvenile rheumatoid arthritis" OR "juvenile rheumatic arthritis" OR "childhood arthritis" OR "juvenile arthritis" OR JIA OR JCA OR JRA	7,505	8,630	250	381
#2	outcome OR outcomes OR prognosis OR "inactive disease" OR "active disease" OR "disease activity" OR "persistent disease" OR remission OR nonremission OR "disease course" OR "disease courses" OR "disease severity" OR Wallace OR function OR "functional impairment" OR "childhood health assessment questionnaire" OR chaq OR "bone erosion" OR damage OR JADI OR "quality of life" OR qol OR hrqol	2,737,564	2,810,638	226,805	513,983
#3	#1 AND #2	2,438	3,245	117	134

^a Search performed on 3 October 2013. No limitations were applied; ^b In PubMed, Embase and PsycInfo terms were searched in title and abstract only; ^c In The Cochrane Library, terms were searched in title, abstract and keywords only.

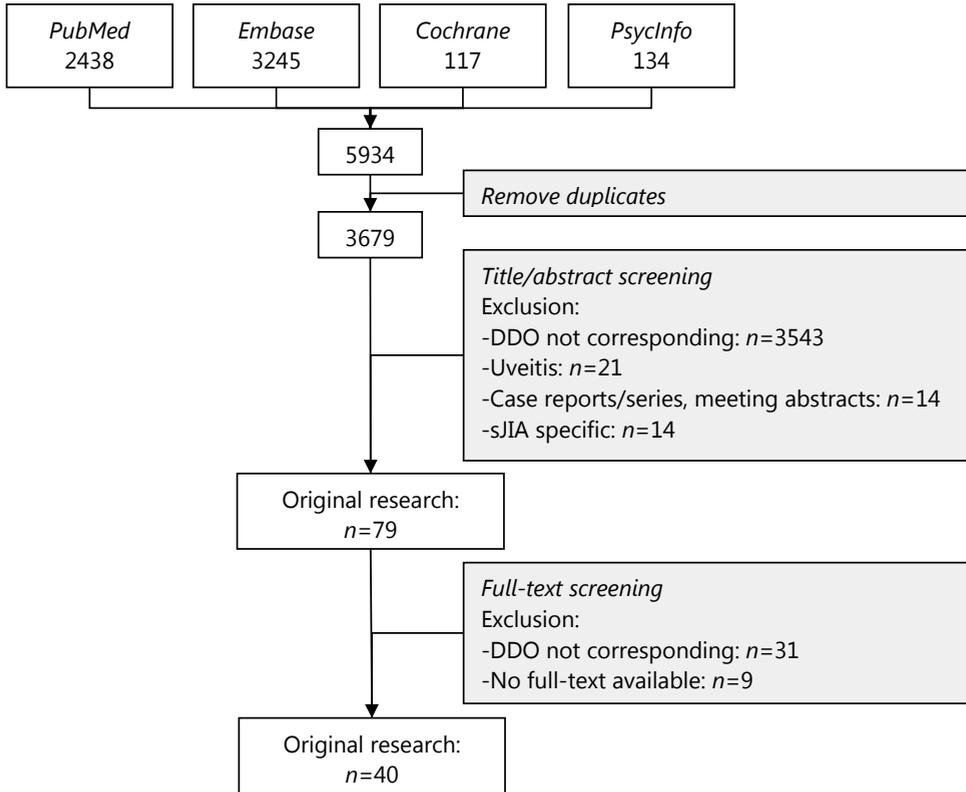


Figure 1. Flow chart. **Abbreviations:** DDO, domain, determinant and outcome; sJIA, systemic juvenile idiopathic arthritis.

criteria were included as well, but other definitions were excluded. Regarding the other outcomes, validated measurements were emphasised, such as the juvenile arthritis damage index (JADI)³² and the childhood health assessment questionnaire (CHAQ)³³, although we were less strict regarding the outcomes joint damage and quality of life, since no single, broadly recognised evaluation is available in these fields. As an additional criterion, we selected articles in which the predictors were determined within 6 months after the diagnosis, to enable timely prediction of the prognosis.

Finally, the two authors discussed which articles would be included for analysis, based on the critical appraisal.

RESULTS

Forty articles with original research and 7 reviews were identified after full-text screening (figure 1), of which the original research articles were critically appraised (supplementary

table 1). Of those, 27 assessed disease activity, 14 joint damage, 22 functional ability and 4 quality of life (multiple outcomes possible). Almost none of the articles described the process of blinding and none described whether the group with the unfavourable outcome was eligible to receive the same care as the group with a favourable outcome. Studies using the validated outcomes and determining predictors within 6 months after diagnosis were selected.

Disease activity

Ten studies assessed disease activity (table 2). Due to our selection, all studies used Wallace criteria or very similar criteria as outcome measurement. Nonetheless, there was still some heterogeneity, since these criteria were applied in different ways. For example, some studies used the percentage of patients achieving either clinical inactive disease (CID), clinical remission on medication (CRM) or clinical remission off all medications (CR),^{24,34-39} whereas others assessed the percentage or cumulative amount of time spent in CID, CRM or CR.^{24,40,41} Other study parameters differed as well, most notably the duration of follow up which varied from 6 months to a median of 15.3 years. Only two studies were prospective studies.^{34,38} One cross-sectional study was included, which assessed only a genetic predictor.⁴¹ Two studies assessed only a single category of JIA.^{37,39}

Table 3 shows the results the studies obtained, demonstrating that most of the variables were clinical and routine laboratory predictors. Only few were genetic, there were hardly any imaging-related and no immunological (i.e., cytokine, cellular characteristics) predictors. Only demographics, JIA category, antinuclear antibodies (ANA) and few others were evaluated more than once. The results show that demographics were not able to predict disease activity, with the notable exception of a longer duration between disease onset and diagnosis, which predicted a worse prognosis in a multivariate analysis in two studies.^{35,38} The results of JIA category reflected the known worse prognosis in the systemic and polyarticular categories. Because of our selection criteria, only onset categories could be evaluated making it impossible to note differences between the oligoarticular persistent and extended categories. Disease activity parameters showed a doubtful ability to predict a worse prognosis. Finally, ANA positivity did not predict disease activity.

Joint damage

Seven studies evaluated joint damage, none of them using the JADI (table 2).³² Instead, other outcome measures were used, such as the Dale or Steinbrocker classification.^{42,43} Other studies assessed the presence of joint space narrowing, erosions or ankylosis, without specifying a classification or scoring system.⁴⁴⁻⁴⁶ Follow up ranged from 5 years to a median of 14.9 years, but was unclear in 1 study.⁴⁴ Three studies were prospective cohort studies.^{44,47,48} Two studies assessed oligoarticular onset JIA patients only.^{45,49}

Table 2. Characteristics of included studies

Reference	Design	Country of origin	N	Inclusion criteria	Outcome ^a	Follow up
24	Retrospective	Italy, USA	437	JIA, minimum 4 y FU	1b, 1c, 1e, 1f	>4 y
34	Prospective	Denmark, Finland, Norway, Sweden	440	JIA, population based	1b	8 y
41	Cross-sectional	Taiwan	107	JRA, minimum 1 y FU	1d	? ^b
38	Prospective	Canada	356	JIA, maximum 6 mo FU	1g, 3a, 4b	6 mo
37	Retrospective	Norway	55	JIA, ERA only	1b, 3g	Median 15.3 y (1b), median 23.0 y (3g)
39	Retrospective	USA	104	JIA, RF+ and RF - poly only, minimum 6 mo FU	1a, 1b	>6 – >36 mo
40	Retrospective	Belgium, Germany, Switzerland, The Netherlands	146	JIA, FU >5 y	1f	5 y
35	Retrospective	Italy	683	JCA, minimum 1 y FU	1h	Mean 10 y
36	Retrospective	Norway	268	JRA	1i, 2a, 3f	Median 14.9 y
67	Retrospective	Germany	171	JCA, JSpA, population based	1i	Mean 7.4 y
44	Prospective	USA	2,704	JIA	2c, 3a	?
48	Prospective	Norway	84	JRA and JSpA	2a, 3b	Mean 9.7 y
47c	Prospective	Sweden	132	JCA, population based	2b, 3d	5 y
49	Retrospective	France	207	JIA, oligo-onset only	2a	6 y
45	Retrospective	Canada	205	JRA, oligo-onset only, minimum 5 y FU	2c, 3c	Median 10.8 y (CHAQ 0-3 y earlier)
46	Retrospective	Canada	181	JRA, minimum 5 y FU, at least 8 y old	2c, 3e	Median >9.2 y (2c), median >11.6 y (3e)
51	Prospective	UK	740	JIA, minimum 1 y FU	3e	>1 y
50c	Prospective	Sweden	132	JCA, population based	3b, 4a	17 y
52	Cross-sectional	Portugal	114	JIA	3a	Median 3 y ^b

Table 2. (continued)

Reference	Design	Country of origin	N	Inclusion criteria	Outcome ^a	Follow up
68	Retrospective	Canada	393	JRA, minimum 5 y FU, at least 8 y old	3a, 3e	Median >11.6 y
69	Retrospective	Italy, USA	227	JRA, maximum 6 mo disease duration, minimum 5 y FU	3e, 4d	>5 y
70	Retrospective	Italy	75	JIA, oligo only	3e	>5 y, median >7 y
71	Retrospective	Turkey	67	JRA, minimum 3 y FU	3e	Mean 4.8 y
72	Retrospective	Canada	235	JIA, minimum 6 mo FU	4c	Mean >3.1 y

Abbreviations: ACR, American College of Rheumatology; AD, active disease; (C)HAQ, (childhood) health assessment questionnaire; CID, clinical inactive disease; CR, clinical remission off all medications; CRM, clinical remission on medication; FU, follow up; JAQQ, juvenile arthritis quality of life questionnaire; JCA, juvenile chronic arthritis; JIA, juvenile idiopathic arthritis; JRA, juvenile rheumatoid arthritis; JSpA, juvenile spondyloarthropathy; MCS, mental component score; mo, months; oligo, oligoarticular; PCS, physical component score; poly, polyarticular; QOLS, quality of life scales; RF, rheumatoid factor; SF-36, Short Form 36 items; y, years.

^a **1a:** Wallace: achievement of CID in first year; **1b:** Wallace: achievement of CRM and/or CR during FU; **1c:** Wallace: duration first period CID; **1d:** Wallace: duration first period CRM; **1e:** Wallace: duration of AD before CID; **1f:** Wallace: percentage of time spent in AD, CID, CRM and/or CR; **1g:** Quasi Wallace: achievement of CID; **1h:** Quasi Wallace: achievement of CR during FU; **1i:** ACR: percentage of patients in remission at last FU visit; **2a:** Dale radiographic classification; **2b:** Cassidy radiographic classification; **2c:** joint space narrowing, erosions and/or ankylosis on imaging; **3a:** (C)HAQ continuous; **3b:** (C)HAQ dichotomized >0; **3c:** (C)HAQ dichotomized >0.12; **3d:** (C)HAQ dichotomized >0.5; **3e:** (C)HAQ dichotomized >0.75; **3f:** SF-36 continuous; **3g:** SF-36 PCS continuous; **4a:** SF-36 MCS; **4b:** JAQQ continuous; **4c:** JAQQ, psychosocial function continuous; **4d:** QOLS dichotomized >80; ^b Time after diagnosis; ^c These describe the same cohort, but different predictors, outcomes and follow up.

Table 4 shows the results. Again, it is notable that almost all predictors were clinical and laboratory values. Only a few genetic and no imaging or immunological predictors were studied. Due to the paucity of studies, many predictors were assessed only once. Nonetheless, the results revealed a worse prognosis in the polyarticular categories. Of note is the negative predictive value of symmetric joint involvement^{36,45} and of rheumatoid factor (RF) positivity^{36,47} in two studies. This may be due to correlations with the polyarticular categories, because the independent effect of these predictors was not studied in multivariate analysis together with JIA category. Other negative predictors included disease activity parameters, which, however, were mostly not validated.

Table 3. Results for outcome disease activity^a

Baseline predictors (within 6 months after diagnosis)	Number of studies					References
	--	< ^b	NS	> ^b	+	
Demographics						
Gender: female		1/1	4			24;35-38;40
Higher age at onset			5	1/0	1	24;35-38;40;67
Higher age at diagnosis		1/1				38;39
Longer time from onset to diagnosis	2					35;38
Family history of AS in first degree relative	1					37
JIA category^c						
Oligoarticular persistent				1		40
Oligoarticular extended				0/1		67
Systemic		0/1		0/1		24;67
Poly RF-	1	0/1		1/1		24;36;38;67
Poly RF+	1	0/2	1			24;35;36;38
Psoriatic			0/3	2		24;35;36;38;39
ERA				1		38
Undifferentiated		0/1	1			34;38
Disease activity						
Higher active joint count	1	1/0	1			36;38;67
Higher limited joint count			2			38;67
Higher PGA		1/0				38
Higher parent/patient GA		1/0				38
Higher pain VAS		1/0				38
Higher CHAQ score		1/0				38
Higher JAQQ score	1					38
Hip involvement	1					34
Ankle involvement	1					37
Sacroiliitis present	1					34
Enthesitis present	1					34
Uveitis		1/0				36
Laboratory						
ANA positive		1/0	3			24;36;38;40
RF positive	1	1/0				36;38
Higher ESR			2			38;67
Higher number of months with elevated ESR (in first 6 months)	1					36
Higher CRP			2			38;67
α ₂ - or γ-globulins			1			67
Genetics						
HLA-B27 positive	1	1/1				34;36;67
HLA-DRB1*08	2					36;37
RANTES -28G allele		0/1				41
RANTES -403A allele			1			41

Table 3. (continued)

Baseline predictors (within 6 months after diagnosis)	Number of studies					References
	--	< ^b	NS	> ^b	+	
Imaging						
Erosions or joint space narrowing on images				0/1		39

Abbreviations: ANA, antinuclear antibodies; AS, ankylosing spondylitis; CHAQ, childhood health assessment questionnaire; CRP, C-reactive protein; ERA, enthesitis-related arthritis; ESR, erythrocyte sedimentation rate; GA, global assessment; HLA, human leukocyte antigen; IFN, interferon; IL, interleukin; JAQQ, juvenile arthritis quality of life questionnaire; JIA, juvenile idiopathic arthritis; MIF, macrophage migratory inhibition factor; PIP, proximal interphalangeal joint; PGA, physician's global assessment; RF, rheumatoid factor; TGF, transforming growth factor; TNF, tumour necrosis factor; VAS, visual analog scale.

^a Values are the number of studies which found that particular result for the respective predictors; ^b The first number in this column refers to the number of studies which performed a multivariate analysis *disconfirming* the univariate finding (studies in which the multivariate analysis *confirmed* the univariate finding, are shown in the columns labelled -- and +, respectively). The second number in this column refers to the number of studies which *did not perform* a multivariate analysis; ^c In studies that evaluated onset categories oligoarticular, polyarticular and systemic, results have been duplicated for oligoarthritis persistent and extended, and polyarthritis RF positive and negative respectively.

Symbols used:

-- lower chance of favourable outcome in multivariate analysis ($p \leq 0.05$)

< lower chance of favourable outcome in univariate analysis ($p \leq 0.05$)

NS: not significant ($p > 0.05$)

S: significant in univariate analysis. Direction of the effect not shown ($p \leq 0.05$)

> higher chance of favourable outcome in univariate analysis ($p \leq 0.05$)

+ higher chance of favourable outcome in multivariate analysis ($p \leq 0.05$)

Functional ability

Fifteen studies were selected, assessing functional ability (table 2). All but two used the CHAQ³³ or the HAQ when appropriate as outcome measurement. However, in some studies the score was dichotomised, mostly using a score of 0.75 as the cut-off point, whereas in other studies it was assessed as a continuous score. Two studies used the 36-item Short-Form health survey (SF-36).^{36;37} Follow up ranged from 6 months to as much as 23 years. Six out of 15 studies were prospective cohort studies^{38;44;47;48;50;51} and one was a cross-sectional study⁵², which assessed genetic predictors only.

Table 5 shows the results. Demographic variables showed varying results. JIA category showed the worse prognosis for polyarticular JIA again. Among disease activity parameters, quite some significantly associated predictors were identified. In general, more disease activity at diagnosis was predictive for a worse prognosis. The two most notable predictors were a higher CHAQ score at baseline and, again, the presence of symmetric joint involvement. The laboratory parameters showed some mixed results, making it difficult to draw an overall conclusion. Some genetic factors, being human leucocyte antigens (HLA) and single nucleotide polymorphisms (SNPs) in cytokine genes were analysed. However, almost all of these were analysed only once and were insignificant.

Table 4. Results for outcome absence of joint damage^a

Baseline predictors (within 6 months after diagnosis)	Number of studies					References
	--	< ^b	NS	S ^b	> ^b +	
Demographics						
Gender: female			2			36;48
Ethnicity: Hispanic		0/1				44
Ethnicity: North-American native	1					46
Higher age at onset		0/1	2	1/0		36;46-48
Longer time from onset to diagnosis	1					48
Diagnosis at first visit			1			48
JIA category^c						
Oligoarticular persistent						
Oligoarticular extended						
Systemic			1	1/0		36;46
Poly RF-		2/1		1/0		36;46-48
Poly RF+		2/1		1/0		36;46-48
Psoriatic						
ERA						
Undifferentiated						
Disease activity						
Higher active joint count	1	1/0				36;48
Higher PGA		1/0				48
Cervical involvement		1/0				36
Wrist involvement		1/0				36
Finger joint involvement		1/0				36
Hip involvement		1/0				36
Ankle and/or wrist involvement	1					36;45
Symmetric involvement	2					36;45
Laboratory						
ANA positive			1			48
RF positive	1	0/1				36;47
Higher ESR		2/1				36;48;49
Higher number of months with elevated ESR (in first 6 months)	1					36
Higher CRP		1/0				48
Higher number of platelets		1/0				48
Genetics						
HLA-B27 positive			1			48
HLA-DPB1 *02				1/0		36
TGF-β1 codon 25G/G (wild type)					1	46
IL-10 -1082G→A genotypes				1/0		46

For declaration of the symbols, footnotes and abbreviations, see table 3.

Table 5. Results for outcome functional ability^a

Baseline predictors (within 6 months after diagnosis)	Number of studies					References
	--	< ^b	NS	S ^b	> ^b +	
Demographics						
Gender: female	2		5			36-38;48;50;51;68
Ethnicity: Hispanic	1					44
Ethnicity: Native North American				1/0		68
Higher age at onset		2/1	5			36-38;47;48;50;68;69
Higher age at diagnosis			1			51
Longer time from onset to diagnosis		2/0	2			38;48;51;68
Diagnosis at first visit			1			48
Family history of AS in first degree relative	1					37
Residence on a reserve	1					68
Rural residence			1			68
JIA category^c			2			48;50
Oligoarticular persistent						
Oligoarticular extended						
Systemic		1/0	3			36;38;68;69
Poly RF-	1	2/0	1			36;38;68;69
Poly RF+	1	2/0	1			36;38;68;69
Psoriatic			1			38
ERA			1			38
Disease activity						
Higher active joint count		4/0	2			36;38;48;50;51;69
Higher limited joint count			0/1			38
Higher involved joint count	1		1			37;50
Higher articular severity score	1					69
Higher PGA		1/1	1			38;48;51
Higher parent/patient GA		2/0				38;51
Higher pain VAS		2/0				38;51
Higher CHAQ score	2					38;51
Higher JAQQ score	1					38
Hand involvement		2/0				36;69
Hip involvement	1		1			36;37
Ankle involvement			1			37
Symmetric involvement	2	1/0				36;45;69
Laboratory						
ANA positive			2		1	50;68;69
RF positive	1	1/0	1			36;50;69
Higher ESR		2/0	4			37;38;48;50;51;69
Higher number of months with elevated ESR (in first 6 months)	1					36
Higher CRP			2			38;48
Higher number of platelets			1			48
Higher albumin level				1/0		69

Table 5. (continued)

Baseline predictors (within 6 months after diagnosis)	Number of studies					References	
	--	< ^b	NS	S ^b	> ^b		+
Genetics							
HLA-B5 (51,52)			1				69
HLA-B8			1				69
HLA-B27			1				37
HLA-B35			1				69
HLA-C3 (9,10)				1/0			69
HLA-DR1 (15,16)			1				69
HLA-DR3 (17,18)			1				69
HLA-DR5 (11,12)			1				69
HLA-DRB1*08		1/0					37
HLA-DRB1*01			1				37
HLA-DRB1*04			1				37
HLA-DPB1*02			1				37
HLA-DPB1*03			1				37
HLA-DQ3			1				69
TNF- α -308G→A			2				46;52
TGF- β codon 10T→C			1				46
TGF- β codon 25G→C			1				46
IFN- γ +874T→A			1				46
IL-6 -174G→C			1				46
IL-10 -1082G→A			1				46
IL-10 -819C→T			1				46
IL-10 -592C→A			1				46
MIF -173 C allele		0/1	1				70;71
MIF -794 CATT repeats			1				70

For declaration of the symbols, footnotes and abbreviations, see table 3.

Quality of life

Only four studies assessing quality of life were retrieved (table 2). Different outcome measurements were used, namely the SF-36 mental component summary scale, the Juvenile Arthritis Quality of Life Questionnaire (JAQQ)⁵³ and the Quality of Life Scales (QOLS).^{54;55} Follow up ranged from 6 months to 17 years. Two studies were prospective cohort studies.^{38;50}

Table 6 shows the results. A comparison is difficult to make, due to the paucity of studies and the difference in outcome measurements. Most of the predictors were studied only once. Generally, only disease activity parameters and potentially a longer doctor's delay showed some prospects as predictors of quality of life.

Table 6. Results for outcome quality of life^a

Baseline predictors (within 6 months after diagnosis)	Number of studies					References
	--	< ^b	NS	S ^b	> ^b +	
Demographics						
Gender: female			3			38;50;72
Higher age at onset			3			38;50;69
Higher age at diagnosis	1					72
Longer time from onset to diagnosis	1					38
JIA category^c			2	1/0		38;50;69
Disease activity						
Higher active joint count	1	1/0	1			38;50;69
Higher limited joint count		0/1	1			38;50
Higher articular severity score		1/0				69
Higher PGA		0/1				38
Higher parent/patient GA		1/0				38
Higher pain VAS	1					38
Higher CHAQ score		1/0				38
Higher JAQQ score	1					38
Hand involvement			1			69
Symmetric involvement			1			69
Laboratory						
ANA positive			2			50;69
RF positive		1/0	1			50;69
Higher ESR	1		2			38;50;69
Higher CRP			1			38
Higher albumin level			1			69
Genetics						
HLA-B5 (51,52)			1			69
HLA-B8			1			69
HLA-B35			1			69
HLA-C3 (9,10)			1			69
HLA-DR1 (15,16)			1			69
HLA-DR3 (17,18)			1			69
HLA-DR5 (11,12)			1			69
HLA-DQ3			1			69

For declaration of the symbols, footnotes and abbreviations, see table 3.

DISCUSSION

The aim of our study was to identify early predictors of disease activity, joint damage, functional ability and quality of life in JIA. The most consistent predictor was polyarticular onset of JIA, which was identified in multiple studies for three of the outcomes assessed. This reflects the known worse prognosis for the polyarticular categories of JIA.^{1;4;5}

However, as was already noted by Adib *et al.* in their review in 2005, there is considerable variance in the prognosis also within categories and it is hence useful to identify other factors which can predict that variance.⁴

Many of the assessed variables did not show that ability. For example, demographic and laboratory data showed too variable results to be valuable predictors, with the notable exception of a diagnostic delay, predicting continuation of disease activity. Other valuable results were shown by disease activity parameters, such as the number of active joints at baseline, the physician's global assessment, the parent or patient global assessment, the CHAQ score and especially symmetric joint involvement. However, most of these parameters were assessed in only a few studies, leaving the question of reproducibility open. Moreover, the issue of confounding arises, since the number of affected joints defines the distinction between oligoarticular and polyarticular onset JIA, and other disease activity parameters may correlate with this predictor. To what extent disease activity parameters show individual predictive capabilities, should be evaluated in multivariate analysis. Lastly, it remains to be elucidated if the involvement of specific joints has any relationship with the outcome. Here too, confounding might obscure the relationships, since, for example, the finger joints are more frequently affected in the polyarticular categories.^{1,6}

Overall, demographic, clinical and laboratory values were insufficient to predict the individual prognosis timely. It is noteworthy that hardly any other potential predictors were taken into account, such as cytokine levels, cell characteristics, results of imaging obtained early in the disease course or genetic evaluations, such as HLA and SNPs in genes with a known function in the immune system. Future studies should address this issue.

There were some concerns about the validity of the studies. Many of the identified studies were retrospective. Since children with a worse disease course tend to visit their physician more often and for a longer period, these retrospective studies are prone to a selection bias. Moreover, many of the studies did not perform a multivariate analysis in order to elucidate individual predictors and try to rule out confounding. Furthermore, almost none of the studies described the process of blinding, a lack of which could lead to information bias, especially in retrospective studies, where the outcome is known at the time of predictor determination. Finally, the selected studies were performed in different periods and made use of different treatment regimens, which potentially confounded our results.

In 2011, the American College of Rheumatology (ACR) issued their recommendations for the treatment of JIA, which included some features of poor prognosis.¹² It was difficult to compare these features with our results, since the paper did not mention their outcome criteria and these features were not all determined in the first 6 months after diagnosis, as in our review. Furthermore, they included predictors that associated only in univariate analysis, but not in multivariate analysis, that were only based on the authors' clinical experience or that were shown to correlate with the outcome instead of predicting

the prognosis. Consequently, the evidence for some of these prognostic features was wanting.

In 2005, Adib *et al.* noted the variability among studies and recommended the development of standardised outcome criteria. Since then, the Wallace criteria have been developed.²³⁻²⁶ In short, each visit is scored as clinical inactive disease (CID) if there are no joints with active arthritis, no systemic features, no uveitis, ESR and/or CRP are normal, physician's global assessment is the best possible on the scale used and duration of morning stiffness is less than 15 minutes. A child is regarded to be in clinical remission on medication (CRM) when these criteria are met during 6 consecutive months and she is regarded to be in clinical remission off medications (CR) when these criteria are met during 12 consecutive months while using no anti-arthritis or anti-uveitis medication.²³⁻²⁶ In our results, Wallace criteria were used in different ways. For example, outcome was defined as the percentage of patients in CID, CRM or CR at the last visit, or as the cumulative amount of time patients spent in CID, CRM or CR during follow up. To avoid variability, the outcome definition should be standardised. Since JIA has a fluctuating disease course,^{24;35;56} the longitudinal aspect of disease activity should be taken into account. Hence, the best outcome measurement seems to be the cumulative amount of time in CID, CRM or CR. A core set of data should routinely be collected in daily clinical practice allowing the evaluation of JIA according to these standardised outcomes such as Wallace criteria or the juvenile arthritis disease activity score (JADAS)⁵⁷, thus facilitating clinical studies.

Regarding joint damage, none of the identified studies used the JADI, which is a validated assessment of both articular and extra-articular damage.³² Instead, they used the visual assessment of radiographic images for joint space narrowing, erosions and ankylosis, which is regarded as the gold standard.⁵⁸ However, its reliability, validity and reproducibility are limited due to the lack of validated scoring systems in paediatric rheumatology.^{32;58} Those that exist focus on the knee and wrist joints, whereas damage can occur in any joint.^{32;43;59-61} Consequently, such validated scoring systems should be developed for other joints as well and longitudinal outcome studies should be performed with either the JADI or these scoring systems.

The CHAQ is a standardised and validated questionnaire for functional ability.³³ It measures the abilities and disabilities in various day to day activities, such as dressing, grooming, eating and grip on a 0-3 point scale. It is still a widely used questionnaire, although there are concerns about its ceiling effect and sensitivity to change.⁶²⁻⁶⁵ Since most of the analysed studies used the CHAQ, comparability among them was good.

Finally, for quality of life, many validated questionnaires exist, such as the SF-36, JAQQ and QOLS. Because all of these were used in the studies we analysed, comparability was very limited. If quality of life is to be an important outcome measurement, one validated measurement should be settled on.

All selected studies focused on finding parameters, which predicted the outcome significantly ($P < 0.05$). Alternatively, one could construct a clinically usable prediction rule,

which yields an individual score indicating the probability of a certain outcome. In such an approach, statistical significance is not important, but the focus is on the predictive value instead.⁶⁶ So far, no such prediction rules for the outcome in JIA have been developed. This hiatus should be addressed in future studies.

We chose to make a strict selection of articles to be analysed, in order to focus on early predictors, homogenise the results and to be able to draw a consistent conclusion. The drawback of this approach is that we may have missed some predictors, by excluding articles that used other outcome definitions. We feel our approach is justified, because if such studies would have been added to the analysis, comparability would be very limited. In order to be not too strict, we did include some articles with outcome measurements very close to the Wallace criteria, some of which were, indeed, also used by Wallace *et al.* to test their criteria against.²⁵

In conclusion, demographic, clinical and laboratory values are insufficient as early predictors for long-term outcome. Studies that evaluate other early predictors, such as genetics or immunologic parameters, hardly exist. Since standardised criteria for disease remission and functional ability have been developed, future efforts should be directed at performing prospective studies, using these validated outcomes, assessing not only clinical but also immunologic, genetic and imaging variables within 6 months of the disease course. These validated outcomes should therefore routinely be registered in daily clinical practice. Finally, validated and standardised outcome measures should be developed for joint damage and quality of life, if these are to be important outcome measures in longitudinal studies.

FUNDING

This project has received funding from the 7th Framework programme of the EU, SP3-People, support for training and career development for researchers (Marie Curie), Network for Initial Training (ITN), FP7-PEOPLE-2011-ITN, under the Marie Skłodowska-Curie grant agreement No 289903 [grant to EHPvD] and SHARE project, EAHC grant number 2011 1202 [grant to NMW]. The sponsors had no role in the study planning, design, management or data analysis.

SUPPLEMENTARY MATERIAL

Supplementary table 1. Critical appraisal

2

REVIEW OF PREDICTORS IN JIA

Reference	Design	Relevance			Validity					
		Dom	Det	Out	Blind	Rec	SoC	Loss	Mis	
Outcome: disease activity										
34	Prospective	+	+	+	?	+	?	+	+	
41	Cross-sectional	+/-	+/-	+	?	+	?	?	?	
38	Prospective	+	+	+/-	?	+	?	+	+	
37	Retrospective	+/-	+	+	?	+	?	+	+	
39	Retrospective	+/-	+/-	+	?	+	?	+	+	
24	Retrospective	+/-	+/-	+	?	+	?	+	+	
40	Retrospective	+	+/-	+	?	+	?	+/-	+/-	
35	Retrospective	+/-	+/-	+/(-)	?	+	?	-	+	
36	Retrospective	+/-	+	+/-	+/-	+	?	+	+	
67	Retrospective	+/-	+/-	+/-	?	+	?	?	?	
73	Prospective	+/-	-	+/-	?	?	?	+	+	
74	Prospective	+	+/-	+	?	+/-	?	+/-	+/-	
75	Prospective	+/-	+	+/-	-	+/- ^a	?	+	+	
44	Prospective	+	+	-	?	+	?	+	+	
50	Prospective	+/-	+	-	?	+	?	+/-	?	
76	Prospective	+/-	+	-	?	+	?	?	?	
45	Retrospective	+/-	+	-	?	+	?	+	+	
68	Retrospective	+/-	+	-	?	+ ^b	?	+	+/-	
46	Retrospective	+/-	+	-	?	+	?	+/-	+/-	
47	Prospective	+/-	+/-	-	?	+	?	+	+	
52	Prospective	+	+/-	-	?	+	?	?	?	
77	Unclear	+	+/-	-	?	+	?	?	?	
78	Retrospective	+/-	+/-	-	?	+	?	+	+	
71	Retrospective	+/-	+/-	-	?	+	?	?	+	
79	Retrospective	+/-	+/-	-	?	+	?	+/-	+/-	
80	Retrospective	+/-	+/- ^c	-	?	+	?	?	?	
81	Prospective	+/-	+	-	?	-	?	+	+	
Outcome: joint damage										
44	Prospective	+	+/-	+/-	?	+	?	+	+	
48	Prospective	+/-	+	+/-	?	+	?	+	+	
47	Prospective	+/-	+/-	+/-	?	+	?	+	+	
36	Retrospective	+/-	+	+/-	+/-	+	?	+	+	
49	Retrospective	+/-	+	+/-	?	+	?	+	+	
45	Retrospective	+/-	+	+/-	?	+	?	+/-	+	
46	Retrospective	+/-	+	+/-	+/-	+	?	+/-	+/-	
74	Prospective	+	+/-	+	?	+/-	?	+/-	+/-	
75	Prospective	+/-	+	+/-	-	+/- ^a	?	+	+	
82	Retrospective	+/-	+	+/-	+/-	? ^d	?	+/-	+	

Supplementary table 1. (continued)

Reference	Design	Relevance			Validity					
		Dom	Det	Out	Blind	Rec	SoC	Loss	Mis	
73	Prospective	+/-	+/-	+/-	+/-	?	?	+	+	
83	Retrospective	+/-	+	+	?	-	?	+	?	
84	Retrospective	+/-	+	+/-	+/-	-	?	+	?	
81	Prospective	+/-	+	+/-	?	-	?	+/-	+	
Outcome: functional ability										
44	Prospective	+	+	+	?	+	?	+	+	
51	Prospective	+	+	+	?	+	?	+	+	
48	Prospective	+/-	+	+/-	?	+	?	+	+	
50	Prospective	+/-	+	+	?	+	?	+/-	?	
52	Cross-sectional	+	+/-	+	?	+	?	?	?	
47	Prospective	+/-	+/-	+	?	+	?	+	+	
38	Prospective	+	+	+/-	?	+	?	+	+	
68	Retrospective	+/-	+	+	?	+ ^b	?	+	+/-	
45	Retrospective	+/-	+	+	?	+	?	+/-	+	
46	Retrospective	+/-	+	+	?	+	?	+/-	+/-	
69	Retrospective	+/-	+	+	?	+	?	+/-	+/-	
70	Retrospective	+/-	+/-	+	+/-	+	?	+	+	
71	Retrospective	+/-	+/-	+	?	+	?	?	+	
36	Retrospective	+/-	+	+/-	?	+	?	+	+	
37	Retrospective	+/-	+	+/-	?	+	?	+/-	+	
85	Prospective	+/-	+	+	-	+/- ^a	?	+	+	
74	Prospective	+	+/-	+	?	+/-	?	+	+/-	
86	Retrospective	+/-	+	+	?	? ^e	?	+/-	+	
82	Retrospective	+/-	+	+	+/-	? ^d	?	+/-	+/-	
84	Retrospective	+/-	+	+	?	-	?	?	+/-	
87	Prospective	+/-	+	+	?	-	?	+	+	
88	Retrospective	+/-	+	+	?	-	?	+/-	?	
Outcome: Quality of Life										
50	Prospective	+/-	+	+	?	+	?	+/-	?	
38	Prospective	+	+	+/-	?	+	?	+	+	
72	Retrospective	+	+	+	?	+ ^b	?	-	+	
69	Retrospective	+/-	+	+	?	+	?	+/-	+/-	

Abbreviations: Blind, blinding; Det, determinant; Dom, domain; Loss, loss to follow up; Mis, missing predictors; Out, outcome; Rec, recruitment; Ref, reference number; SoC, standardization of care.

Shaded articles were excluded for analysis.

^a Clarified in personal communication with the author; ^b Some covariates in the multivariable analysis were determined after 6 months; ^c Predictors were entered in a multivariable analysis, but only predictors determined after >6 months were entered as covariates; ^d Predictors were determined at the time of the first radiograph. No information about the interval between diagnosis and the first radiograph is given. However, the strongest predictor for all outcomes was radiographic progression in the first year (+/-); ^e Predictors were determined at the time of the first CHAQ. No information about the interval between diagnosis and the first CHAQ is given.

Supplementary table 1. (continued)

Criteria: **Domain:** + Children with confirmed JIA, according to currently valid ILAR criteria +/- Children with JCA/JRA according to previously valid criteria, or children with JIA and additional criteria (e.g. hospitalized, specific categories only) - All other domains; **Determinant:** + Prediction model or single predictors corrected for confounding in multivariable analysis +/- Single predictors in univariate analysis - No predictors; **Outcome:** + Clinical remission: (preliminary) Wallace criteria for clinical inactive disease; joint damage: JADI score; functional ability: CHAQ; quality of life: any validated questionnaire; and follow-up ≥ 1 year +/- Clinical remission: criteria similar to Wallace criteria; Other outcomes: all validated outcome measures, or follow-up <1 year - Clinical remission: no use of Wallace criteria; Other outcomes: no use of validated outcome criteria; **Blinding:** + Both patient and physician blinded (or not applicable in case of objective measurements) +/- Patient or physician not blinded - Not blinded; **Recruitment:** + Predictors determined at time of diagnosis or <6 months (or time of determination does not matter as in genetic evaluations, gender, age at onset, etc.) +/- Predictors determined more than 6 months after diagnosis, but <1 year - Predictors determined after 1 year, or completely at random; **Standardization of care:** + All participants treated according to standards of care - No standardized care; **Loss to follow up (missing outcome):** + <20% and unselective loss to follow up; or >20%, unselective and solved with a statistically valid method (imputation) +/- >20% (not imputed) but unselective loss to follow up - Selective loss to follow up; **Missing predictors:** + <20% and unselective; or >20%, unselective and solved with a statistically valid method (imputation) +/- >20% (not imputed) but unselective - Selective missing predictors.

REFERENCE LIST

1. Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. *Lancet* 2011; 377(9783):2138-2149.
2. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004; 31(2):390-392.
3. Adib N, Silman A, Thomson W. Outcome following onset of juvenile idiopathic inflammatory arthritis: I. Frequency of different outcomes. *Rheumatology* 2005; 44(8):995-1001.
4. Adib N, Silman A, Thomson W. Outcome following onset of juvenile idiopathic inflammatory arthritis: II. Predictors of outcome in juvenile arthritis. *Rheumatology* 2005; 44(8):1002-1007.
5. Ravelli A. Toward an understanding of the long-term outcome of juvenile idiopathic arthritis. *Clin Exp Rheumatol* 2004; 22(3):271-275.
6. Oen K. Long-term outcomes and predictors of outcomes for patients with juvenile idiopathic arthritis. *Best Pract Res Clin Rheumatol* 2002; 16(3):347-360.
7. Foell D, Wulffraat N, Wedderburn LR, Wittkowski H, Frosch M, Gerss J et al. Methotrexate withdrawal at 6 vs 12 months in juvenile idiopathic arthritis in remission: a randomized clinical trial. *JAMA* 2010; 303(13):1266-1273.
8. Lovell DJ, Ruperto N, Goodman S, Reiff A, Jung L, Jarosova K et al. Adalimumab with or without methotrexate in juvenile rheumatoid arthritis. *N Engl J Med* 2008; 359(8):810-820.
9. Ruperto N, Lovell DJ, Quartier P, Paz E, Rubio-Perez N, Silva CA et al. Abatacept in children with juvenile idiopathic arthritis: a randomised, double-blind, placebo-controlled withdrawal trial. *Lancet* 2008; 372(9636):383-391.
10. Lovell DJ, Reiff A, Ilowite NT, Wallace CA, Chon Y, Lin SL et al. Safety and efficacy of up to eight years of continuous etanercept therapy in patients with juvenile rheumatoid arthritis. *Arthritis Rheum* 2008; 58(5):1496-1504.
11. Lovell DJ, Giannini EH, Reiff A, Cawkwell GD, Silverman ED, Nocton JJ et al. Etanercept in children with polyarticular juvenile rheumatoid arthritis. Pediatric Rheumatology Collaborative Study Group. *N Engl J Med* 2000; 342(11):763-769.
12. Beukelman T, Patkar NM, Saag KG, Tolleson-Rinehart S, Cron RQ, DeWitt EM et al. 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: initiation and safety monitoring of therapeutic agents for the treatment of arthritis and systemic features. *Arthritis Care Res (Hoboken)* 2011; 63(4):465-482.
13. De Benedetti F, Brunner HI, Ruperto N, Kenwright A, Wright S, Calvo I et al. Randomized trial of tocilizumab in systemic juvenile idiopathic arthritis. *N Engl J Med* 2012; 367(25):2385-2395.
14. Anderson JJ, Wells G, Verhoeven AC, Felson DT. Factors predicting response to treatment in rheumatoid arthritis: the importance of disease duration. *Arthritis Rheum* 2000; 43(1):22-29.
15. Boers M. Understanding the window of opportunity concept in early rheumatoid arthritis. *Arthritis Rheum* 2003; 48(7):1771-1774.
16. Moreland LW, Bridges SL, Jr. Early rheumatoid arthritis: a medical emergency? *Am J Med* 2001; 111(6):498-500.
17. Mottonen T, Hannonen P, Korpela M, Nissila M, Kautiainen H, Ilonen J et al. Delay to institution of therapy and induction of remission using single-drug or combination-disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. *Arthritis Rheum* 2002; 46(4):894-898.
18. O'Dell JR. Treating rheumatoid arthritis early: a window of opportunity? *Arthritis Rheum* 2002; 46(2):283-285.
19. Quinn MA, Emery P. Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. *Clin Exp Rheumatol* 2003; 21(5 Suppl 31):S154-S157.

20. Albers HM, Wessels JA, van der Straaten RJ, Brinkman DM, Suijlekom-Smit LW, Kamphuis SS et al. Time to treatment as an important factor for the response to methotrexate in juvenile idiopathic arthritis. *Arthritis Rheum* 2009; 61(1):46-51.
21. Wallace CA, Giannini EH, Spalding SJ, Hashkes PJ, O'Neil KM, Zeff AS et al. Trial of early aggressive therapy in polyarticular juvenile idiopathic arthritis. *Arthritis Rheum* 2012; 64(6):2012-2021.
22. Vastert SJ, de JW, Noordman BJ, Holzinger D, Kuis W, Prakken BJ et al. Effectiveness of first line use of recombinant IL-1RA treatment in steroid naive systemic juvenile idiopathic arthritis: Results of a prospective cohort study. *Arthritis Rheum* 2013.
23. Wallace CA, Ruperto N, Giannini E. Preliminary criteria for clinical remission for select categories of juvenile idiopathic arthritis. *J Rheumatol* 2004; 31(11):2290-2294.
24. Wallace CA, Huang B, Bandeira M, Ravelli A, Giannini EH. Patterns of clinical remission in select categories of juvenile idiopathic arthritis. *Arthritis Rheum* 2005; 52(11):3554-3562.
25. Wallace CA, Ravelli A, Huang B, Giannini EH. Preliminary validation of clinical remission criteria using the OMERACT filter for select categories of juvenile idiopathic arthritis. *J Rheumatol* 2006; 33(4):789-795.
26. Wallace CA, Giannini EH, Huang B, Irt L, Ruperto N. American College of Rheumatology provisional criteria for defining clinical inactive disease in select categories of juvenile idiopathic arthritis. *Arthritis Care Res (Hoboken)* 2011; 63(7):929-936.
27. De Benedetti F, Martini A. Is systemic juvenile rheumatoid arthritis an interleukin 6 mediated disease? *J Rheumatol* 1998; 25(2):203-207.
28. Fall N, Barnes M, Thornton S, Luyrink L, Olson J, Ilowite NT et al. Gene expression profiling of peripheral blood from patients with untreated new-onset systemic juvenile idiopathic arthritis reveals molecular heterogeneity that may predict macrophage activation syndrome. *Arthritis Rheum* 2007; 56(11):3793-3804.
29. Gattorno M, Piccini A, Lasiglie D, Tassi S, Brisca G, Carta S et al. The pattern of response to anti-interleukin-1 treatment distinguishes two subsets of patients with systemic-onset juvenile idiopathic arthritis. *Arthritis Rheum* 2008; 58(5):1505-1515.
30. Masters SL, Simon A, Aksentijevich I, Kastner DL. Horror autoinflammaticus: the molecular pathophysiology of autoinflammatory disease (*). *Annu Rev Immunol* 2009; 27:621-668.
31. Ravelli A, Martini A. Juvenile idiopathic arthritis. *Lancet* 2007; 369(9563):767-778.
32. Viola S, Felici E, Magni-Manzoni S, Pistorio A, Buoncompagni A, Ruperto N et al. Development and validation of a clinical index for assessment of long-term damage in juvenile idiopathic arthritis. *Arthritis Rheum* 2005; 52(7):2092-2102.
33. Ruperto N, Ravelli A, Pistorio A, Malattia C, Cavuto S, Gado-West L et al. Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries. Review of the general methodology. *Clin Exp Rheumatol* 2001; 19(4 Suppl 23):S1-S9.
34. Berntson L, Nordal E, Aalto K, Peltoniemi S, Herlin T, Zak M et al. HLA-B27 predicts a more chronic disease course in an 8-year followup cohort of patients with juvenile idiopathic arthritis. *J Rheumatol* 2013; 40(5):725-731.
35. Fantini F, Gerloni V, Gattinara M, Cimaz R, Arnoldi C, Lupi E. Remission in juvenile chronic arthritis: A cohort study of 683 consecutive cases with a mean 10 year followup. *J Rheumatol* 2003; 30(3):579-584.
36. Flato B, Lien G, Smerdel A, Vinje O, Dale K, Johnston V et al. Prognostic factors in juvenile rheumatoid arthritis: A case-control study revealing early predictors and outcome after 14.9 Years. *J Rheumatol* 2003; 30(2):386-393.
37. Flato B, Hoffmann-Vold AM, Reiff A, Forre O, Lien G, Vinje O. Long-term outcome and prognostic factors in enthesitis-related arthritis: A case-control study. *Arthritis Rheum* 2006; 54(11):3573-3582.

38. Oen K, Tucker L, Huber AM, Miettunen P, Scuccimarrì R, Campillo S et al. Predictors of early inactive disease in a juvenile idiopathic arthritis cohort: Results of a Canadian Multicenter, prospective inception cohort study. *Arthritis Care Res* 2009; 61(8):1077-1086.
39. Ringold S, Seidel KD, Koepsell TD, Wallace CA. Inactive disease in polyarticular juvenile idiopathic arthritis: Current patterns and associations. *Rheumatology* 2009; 48(8):972-977.
40. Albers HM, Brinkman DMC, Kamphuis SSM, van Suijlekom-Smit LWA, van Rossum MAJ, Hoppenreijns EPAH et al. Clinical course and prognostic value of disease activity in the first two years in different subtypes of juvenile idiopathic arthritis. *Arthritis Care Res* 2010; 62(2):204-212.
41. Yao TC, Tsai YC, Huang JL. Association of RANTES promoter polymorphism with juvenile rheumatoid arthritis. *Arthritis Rheum* 2009; 60(4):1173-1178.
42. Cassidy JT, Martel W. Juvenile rheumatoid arthritis: clinicoradiologic correlations. *Arthritis Rheum* 1977; 20(2 Suppl):207-211.
43. Dale K, Paus AC, Laires K. A radiographic classification system in juvenile rheumatoid arthritis applied to the knee. *Eur Radiol* 1994; 4:27-32.
44. Pelajo CF, Angeles-Han ST, Prahalad S, Sgarlat CM, Davis TE, Miller LC et al. Evaluation of the association between Hispanic ethnicity and disease activity and severity in a large cohort of patients with juvenile idiopathic arthritis. *Rheumatol Int* 2013;1-6.
45. Al-Matar MJ, Petty RE, Tucker LB, Malleson PN, Schroeder ML, Cabral DA. The early pattern of joint involvement predicts disease progression in children with oligoarticular (pauciarticular) juvenile rheumatoid arthritis. *Arthritis Rheum* 2002; 46(10):2708-2715.
46. Oen K, Malleson PN, Cabral DA, Rosenberg AM, Petty RE, Nickerson P et al. Cytokine genotypes correlate with pain and radiologically defined joint damage in patients with juvenile rheumatoid arthritis. *Rheumatology* 2005; 44(9):1115-1121.
47. Bertilsson L, Andersson-Gare B, Fasth A, Forsblad-d'Elia H. A 5-year prospective population-based study of juvenile chronic arthritis: Onset, disease process, and outcome. *Scand J Rheumatol* 2012; 41(5):379-382.
48. Flato B, Aasland A, Vinje O, Forre O. Outcome and predictive factors in juvenile rheumatoid arthritis and juvenile spondyloarthritis. *J Rheumatol* 1998; 25(2):366-375.
49. Guillaume S, Prieur AM, Coste J, Job-Deslandre C, Sharma L. Long-term outcome and prognosis in oligoarticular-onset juvenile idiopathic arthritis. *Arthritis Rheum* 2000; 43(8):1858-1865.
50. Bertilsson L, Andersson-Gare B, Fasth A, Petersson IF, Forsblad-D'elia H. Disease course, outcome, and predictors of outcome in a population-based juvenile chronic arthritis cohort followed for 17 years. *J Rheumatol* 2013; 40(5):715-724.
51. Hyrich KL, Lal SD, Foster HE, Thornton J, Adib N, Baildam E et al. Disease activity and disability in children with juvenile idiopathic arthritis one year following presentation to paediatric rheumatology. Results from the Childhood Arthritis Prospective Study. *Rheumatology* 2010; 49(1):116-122.
52. Mourao AF, Caetano-Lopes J, Costa P, Canhao H, Santos MJ, Pinto P et al. Tumor necrosis factor-(alpha) -308 genotypes influence inflammatory activity and TNF-(alpha) serum concentrations in children with juvenile idiopathic arthritis. *J Rheumatol* 2009; 36(4):837-842.
53. Duffy CM, Arsenault L, Duffy KN, Paquin JD, Strawczynski H. The Juvenile Arthritis Quality of Life Questionnaire--development of a new responsive index for juvenile rheumatoid arthritis and juvenile spondyloarthritis. *J Rheumatol* 1997; 24(4):738-746.
54. Burckhardt CS, Archenholtz B, Bjelle A. Measuring the quality of life of women with rheumatoid arthritis or systemic lupus erythematosus: a Swedish version of the Quality of Life Scale (QOLS). *Scand J Rheumatol* 1992; 21(4):190-195.
55. Burckhardt CS, Clark SR, Bennett RM. Fibromyalgia and quality of life: a comparative analysis. *J Rheumatol* 1993; 20(3):475-479.

56. de Rotte MC, Luime JJ, Bulatovic M, Hazes JM, Wulffraat NM, de JR. Do snapshot statistics fool us in MTX pharmacogenetic studies in arthritis research? *Rheumatology (Oxford)* 2010; 49(6):1200-1201.
57. Consolaro A, Ruperto N, Bazso A, Pistorio A, Magni-Manzoni S, Filocamo G et al. Development and validation of a composite disease activity score for juvenile idiopathic arthritis. *Arthritis Rheum* 2009; 61(5):658-666.
58. Magni-Manzoni S, Malattia C, Lanni S, Ravelli A. Advances and challenges in imaging in juvenile idiopathic arthritis. *Nat Rev Rheumatol* 2012; 8(6):329-336.
59. Poznanski AK, Hernandez RJ, Guire KE, Bereza UL, Garn SM. Carpal length in children--a useful measurement in the diagnosis of rheumatoid arthritis and some congenital malformation syndromes. *Radiology* 1978; 129(3):661-668.
60. Ravelli A, Ioseliani M, Norambuena X, Sato J, Pistorio A, Rossi F et al. Adapted versions of the Sharp/van der Heijde score are reliable and valid for assessment of radiographic progression in juvenile idiopathic arthritis. *Arthritis Rheum* 2007; 56(9):3087-3095.
61. Rossi F, Di DF, Galipo O, Pistorio A, Valle M, Magni-Manzoni S et al. Use of the Sharp and Larsen scoring methods in the assessment of radiographic progression in juvenile idiopathic arthritis. *Arthritis Rheum* 2006; 55(5):717-723.
62. Lam C, Young N, Marwaha J, McLimont M, Feldman BM. Revised versions of the Childhood Health Assessment Questionnaire (CHAQ) are more sensitive and suffer less from a ceiling effect. *Arthritis Rheum* 2004; 51(6):881-889.
63. Groen W, Unal E, Norgaard M, Maillard S, Scott J, Berggren K et al. Comparing different revisions of the Childhood Health Assessment Questionnaire to reduce the ceiling effect and improve score distribution: Data from a multi-center European cohort study of children with JIA. *Pediatr Rheumatol Online J* 2010; 8:16.
64. Van Dijk M, Groen W, Moors S, Bekkering P, Hegeman A, Janssen A et al. The Dutch translation of the revised Childhood Health Assessment Questionnaire: a preliminary study of score distribution. *Clin Exp Rheumatol* 2010; 28(2):275-280.
65. Norgaard M, Thastum M, Herlin T. The relevance of using the Childhood Health Assessment Questionnaire (CHAQ) in revised versions for the assessment of juvenile idiopathic arthritis. *Scand J Rheumatol* 2013; 42(6):457-464.
66. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer; 2009.
67. Minden K, Kiesslering U, Listing J, Niewerth M, Doring E, Meincke J et al. Prognosis of patients with juvenile chronic arthritis and juvenile spondyloarthropathy. *J Rheumatol* 2000; 27(9):2256-2263.
68. Oen K, Malleson PN, Cabral DA, Rosenberg AM, Petty RE, Reed M et al. Early predictors of longterm outcome in patients with juvenile rheumatoid arthritis: Subset-specific correlations. *J Rheumatol* 2003; 30(3):585-593.
69. Ruperto N, Ravelli A, Levinson JE, Shear ES, Murray K, Tague BL et al. Longterm health outcomes and quality of life in American and Italian inception cohorts of patients with juvenile rheumatoid arthritis. II. Early predictors of outcome. *J Rheumatol* 1997; 24(5):952-958.
70. Vivarelli M, D'Urbano LE, Insalaco A, Lunt M, Jury F, Tozzi AE et al. Macrophage migration inhibitory factor (MIF) and oligoarticular juvenile idiopathic arthritis (o-JIA); association of MIF promoter polymorphisms with response to intra-articular glucocorticoids. *Clin Exp Rheumatol* 2007; 25(5):775-781.
71. Berdeli A, Ozyurek AR, Ulger Z, Gurses D, Levent E, Salar K et al. Association of macrophage migration inhibitory factor gene -173 G/C polymorphism with prognosis in turkish children with juvenile rheumatoid arthritis. *Rheumatol Int* 2006; 26(8):726-731.
72. April KT, Cavallo S, Feldman DE. Children with juvenile idiopathic arthritis: are health outcomes better for those

- diagnosed younger? *Child Care Health Dev* 2013; 39(3):442-448.
73. Lipinska J, Brozik H, Stanczyk J, Smolewska E. Anticitrullinated protein antibodies and radiological progression in juvenile idiopathic arthritis. *J Rheumatol* 2012; 39(5):1078-1087.
 74. Nordal EB, Zak M, Aalto K, Berntson L, Fasth A, Herlin T et al. Validity and predictive ability of the juvenile arthritis disease activity score based on CRP versus ESR in a Nordic population-based setting. *Ann Rheum Dis* 2012; 71(7):1122-1127.
 75. Selvaag AM, Flato B, Dale K, Lien G, Vinje O, Smerdel-Ramoya A et al. Radiographic and clinical outcome in early juvenile rheumatoid arthritis and juvenile spondyloarthropathy: A 3-year prospective study. *J Rheumatol* 2006; 33(7):1382-1391.
 76. Arguedas O, Fasth A, Anderson-Gare B. A prospective population based study on outcome of juvenile chronic arthritis in Costa Rica. *J Rheumatol* 2002; 29(1):174-183.
 77. Ozen S, Alikasifoglu M, Bakkaloglu A, Duzova A, Jarosova K, Nemcova D et al. Tumour necrosis factor (alpha) G -> A -238 and G -> A -308 polymorphisms in juvenile idiopathic arthritis. *Rheumatology* 2002; 41(2):223-227.
 78. Hsu CT, Lin YT, Yang YH, Chiang BL. Factors affecting clinical and therapeutic outcomes of patients with juvenile rheumatoid arthritis. *Scand J Rheumatol* 2004; 33(5):312-317.
 79. Greenwald AG, Zakerzadeh A, Laxer RM, Schneider R, Cameron B, Feldman B et al. Later-onset rheumatoid factor negative polyarticular juvenile idiopathic arthritis (JIA): A unique patient group? *Clin Exp Rheumatol* 2013; 31(4):645-652.
 80. Wu JF, Yang YH, Wang LC, Lee JH, Shen EY, Chiang BL. Comparative usefulness of C-reactive protein and erythrocyte sedimentation rate in juvenile rheumatoid arthritis. *Clin Exp Rheumatol* 2007; 25(5):782-785.
 81. Andersson Gare B, Fasth A. The natural history of juvenile chronic arthritis: A population based cohort study. I. Onset and disease process. *J Rheumatol* 1995; 22(2):295-307.
 82. Magni-Manzoni S, Rossi F, Pistorio A, Temporini F, Viola S, Beluffi G et al. Prognostic Factors for Radiographic Progression, Radiographic Damage, and Disability in Juvenile Idiopathic Arthritis. *Arthritis Rheum* 2003; 48(12):3509-3517.
 83. De Oliveira Sato J, Corrente JE, Saad-Magalhaes C. Progression of articular and extraarticular damage in oligoarticular juvenile idiopathic arthritis. *Clin Exp Rheumatol* 2011; 29(5):871-877.
 84. Oen K, Reed M, Malleson PN, Cabral DA, Petty RE, Rosenberg AM et al. Radiologic outcome and its relationship to functional disability in juvenile rheumatoid arthritis. *J Rheumatol* 2003; 30(4):832-840.
 85. Selvaag AM, Lien G, Sorskaar D, Vinje O, Forre O, Flato B. Early disease course and predictors of disability in juvenile rheumatoid arthritis and juvenile spondyloarthropathy: A 3 year prospective study. *J Rheumatol* 2005; 32(6):1122-1130.
 86. Magni-Manzoni S, Pistorio A, Labo E, Viola S, Garcia-Munitis P, Panigada S et al. A longitudinal analysis of physical functional disability over the course of juvenile idiopathic arthritis. *Ann Rheum Dis* 2008; 67(8):1159-1164.
 87. Andersson Gare B, Fasth A. The natural history of juvenile chronic arthritis: A population based cohort study. II. Outcome. *J Rheumatol* 1995; 22(2):308-319.
 88. Zak M, Pedersen FK. Juvenile chronic arthritis into adulthood: A long-term follow-up study. *Rheumatology* 2000; 39(2):198-204.

Chapter

3

PREDICTION OF INACTIVE DISEASE IN JUVENILE IDIOPATHIC ARTHRITIS: A MULTICENTER OBSERVATIONAL COHORT STUDY

E.H. Pieter van Dijkhuizen, Orfeas Aidonopoulos,
Nienke M. ter Haar, Denise Pires Marafon,
Silvia Magni-Manzoni, Yannis E. Ioannidis,
Lorenza Putignani, Sebastiaan J. Vastert,
Clara Malattia, Fabrizio De Benedetti
and Alberto Martini for the MD-Paedigree consortium

Submitted

ABSTRACT

Objectives

To predict the occurrence of inactive disease in juvenile idiopathic arthritis (JIA) in the first two years of disease.

Methods

An inception cohort of 152 treatment-naïve JIA patients with disease duration less than six months was analyzed. Potential predictors were baseline clinical variables, joint ultrasound, gut microbiota composition and a panel of inflammation-related compounds in blood plasma. Various algorithms were employed to predict inactive disease according to Wallace criteria at six months intervals in the first two years. Performance of the models was evaluated using the split-cohort technique. The cohort was analyzed in its entirety, and separate models were developed for oligoarticular patients, polyarticular rheumatoid factor (RF) negative patients and anti-nuclear antibody (ANA) positive patients.

Results

All models analyzing the cohort as a whole showed poor performance in test data (area under the curve [AUC] <0.65). The subgroup models performed better. Inactive disease was predicted by lower baseline juvenile arthritis disease activity score (JADAS)-71 and lower relative abundance of the operational taxonomic unit *Mogibacteriaceae* for oligoarticular patients (AUC in test data 0.69); shorter duration of morning stiffness, higher hemoglobin and lower CXCL-9 levels at baseline for polyarticular RF negative patients (AUC in test data 0.69); and shorter duration of morning stiffness and higher baseline hemoglobin for ANA positive patients (AUC in test data 0.72).

Conclusions

Inactive disease could not be predicted with satisfactory accuracy in the whole cohort. Interesting predictors were found in more homogeneous subgroups. These need to be validated in future studies.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is an umbrella-term covering chronic arthritis of unknown cause lasting for more than six weeks with an onset before 16 years of age.¹ The prognosis of children with JIA has improved considerably over the last decades. Following the advent of methotrexate and biologic agents, many children achieve disease remission within a reasonable time frame.²⁻⁵ Nevertheless, there is still a substantial proportion of children who do not achieve inactive disease, or fail to do so within a short period of time. Prolonged disease activity causes distress, diminishes the overall well-being of the child and is associated with a higher risk of long term sequelae of the disease, including irreversible joint damage.⁶⁻¹⁰ On top of this, the so-called window of opportunity, known from adult rheumatoid arthritis,¹¹⁻¹⁵ may also be present in pediatric rheumatology, meaning that better outcomes are achieved if effective treatment is started early in the course of disease.

For various reasons, it is not feasible to prescribe biologic agents to all children with JIA. First, this would expose many children to the potential side effects of the drugs and, secondly, this practice would be too costly. Moreover, not all children need biologic agents, since some respond well to intra-articular corticosteroid injections only or, failing that strategy, may benefit from the addition of methotrexate.¹⁶

Therefore, there is a clinical need of predictors capable of stratifying JIA patients according to their prognosis, thus enabling to identify early in the course of disease, ideally at the time of diagnosis, those patients in need of more aggressive therapy. A review of the literature revealed that it is unlikely that clinical predictors suffice to achieve that goal.¹⁷ This conclusion was challenged by a recently published clinical prediction model that performed well in test data.¹⁸ However, its performance was poor in our cohort. Therefore, other categories of predictors need to be taken into consideration for the accurate prediction of the prognosis of JIA.

Our aim was to construct a prediction model of the prognosis of JIA patients in a well-defined, multicentre cohort, using clinical, imaging and microbiota data as well as a panel of inflammation-related compounds in blood plasma.

METHODS

We performed a prospective, observational cohort study at the Istituto Giannina Gaslini (IGG, Genoa, Italy), Ospedale Pediatrico Bambino Gesù (OPBG, Rome, Italy) and the Wilhelmina Children's hospital (WKZ, Utrecht, The Netherlands), between October 2013 and December 2015 (IGG and OPBG) and between March 2014 and December 2015 (WKZ). Consecutive patients with JIA according to international league against rheumatism (ILAR) criteria¹⁹ at disease onset (onset of symptoms at most six months

before enrolment), who had not started any treatment other than non-steroidal anti-inflammatory drugs (NSAIDs) were eligible for enrolment. Systemic JIA patients were not included.

At baseline, clinical parameters regarding demographics and disease manifestations were collected and routine laboratory examinations were performed. The childhood health assessment questionnaire was completed by the child or his parents. Information about ongoing treatment (NSAIDs) was collected. An additional sample of blood was collected, peripheral blood plasma was isolated and frozen at -80 °C. These samples were then shipped to WKZ on dry ice by overnight shipment to perform an in-house developed and validated (ISO9001 certified) multiplex immunoassay (Laboratory of Translational Immunology, University Medical Centre Utrecht) based on Luminex technology (xMAP, Luminex Austin TX USA). The assay was performed as described previously.²⁰ Aspecific heterophilic immunoglobulins were pre-absorbed from all samples with heteroblock (Omega Biologicals, Bozeman MT, USA). Acquisition was performed with the Biorad FlexMAP3D (Biorad laboratories, Hercules USA) in combination with xPONENT software version 4.2 (Luminex). Data were analyzed by 5-parametric curve fitting using Bio-Plex Manager software, version 6.1.1 (Biorad).

Additionally, a stool sample for gut microbiota determination was collected and frozen at -80 °C. These samples were shipped to OPBG on dry ice by overnight shipment for 16S bacterial ribosomal RNA pyrosequencing, as described in literature.²¹ Reads were analyzed by Quantitative Insights into Microbial Ecology (QIIME, v.1.8.0), grouped into operational taxonomic units (OTUs) at a sequence similarity level of 97%.²²⁻²⁵ [Van Dijkhuizen *et al.*, in preparation].

Finally, an extended joint ultrasonography was performed using a pre-specified protocol. Details of the protocol have been published previously [Lanni *et al.* Clin Exp Rheum, in press].

Enrolled patients were evaluated every 6 months and additionally in case of a relapse, for 2 years or until the end of the study period in February 2017, whichever was first. At each visit, disease activity was assessed using the Wallace criteria for inactive disease.²⁶ Ongoing treatment was categorized as 1) no treatment; 2) intra-articular joint injections with or without NSAIDs; 3) methotrexate with or without oral steroids; and 4) biological agents.

Statistical analysis

Three different prediction models were developed: one consisting of clinical (including ultrasound) data only; one containing clinical and microbiota data; and one using clinical and Luminex data. The outcome was inactive disease versus active disease for all models. Missing clinical data were imputed by multiple imputations using chained equations (MICE) producing 10 imputed data sets.²⁷ All other clinical variables, including

the outcome, were used in the imputations, provided they were not linearly related to the variables to be imputed.²⁷ Missing microbiota and Luminex data were only due to missing samples in our cohort and were therefore not imputed.

The cohort was split at random into two thirds for model training and one third for model validation. To eliminate the influence of the data split on the model evaluation, the split was performed ten times and the model was trained and tested in each of these split data sets. The performance was summarized as the mean over the 10 splits.

In the training data, univariate variable selection was performed by fitting a model and retaining the variables with a p-value <0.1 , or <0.2 in case no variable had a p-value <0.1 . If variables correlated in training data (Spearman's $|\rho| >0.6$), the variable with the lowest p-value was retained. The remaining variables were pooled in a multivariable model in the training data and this model was subsequently applied to the test data.

Several different algorithms were tried and model performance was assessed in an appropriate manner for each algorithm. In the primary analysis, to exploit the information contained in all visits for all patients, a repeated measurements analysis using generalized estimating equations (GEE) was performed.^{28;29} Performance of this model was tested using an area under the curve (AUC)-like statistic, adapted to models with repeated measurements.³⁰

Secondary analyses

Various secondary analyses were performed, including Cox proportional hazards models of time to first inactive disease and Cox regressions of time to inactive disease, using recurrent events analysis.³¹ The performance of these models was assessed by predicting the risk score of patients in the test data and correlating this risk score with the observed outcome in the test data set, using Somer's D correlation, adapted to censored data,^{32;33} from which the c-statistic (similar to the AUC) was calculated.

Moreover, next to the longitudinal approach, the outcomes at 6 months and 12 months after baseline were analyzed separately. This approach opened up the possibility of applying more algorithms: we fitted logistic regression models, random forests and support vector machines. Performance of these models was assessed by calculating the AUC.

Finally, to reduce the heterogeneity of JIA, we fitted models to the groups of oligoarticular patients, polyarticular rheumatoid factor (RF) negative patients and antinuclear antibody (ANA) positive patients separately. Patients were pragmatically considered ANA positive, if they had at least one positive ANA determination at baseline (titer $\geq 1:160$).

Statistical analyses were carried out using R 3.3.2 (R foundation for statistical computing, Vienna, Austria).

RESULTS

Of 169 enrolled patients, 10 were enrolled after the start of methotrexate and 7 were lost to follow up (i.e., no follow up data collected at all), leaving 152 patients to be analyzed. Luminex data was available for 121 patients and gut microbiota data for 91 patients. Missing Luminex and microbiota data were due to difficulties in the collection or shipment of samples and were unrelated to disease characteristics. Blood samples for Luminex analysis were collected in EDTA tubes at OPBG (compared to sodium heparin tubes at IGG and WKZ). Consequently, the Luminex results of OPBG patients were incomparable to those of the remainder of patients (data not shown). Therefore, the Luminex results of OPBG patients were excluded from the analysis. Thus, the clinical models were developed in 152 patients, contributing 508 visits, the microbiota models in 91 patients contributing 310 visits and the Luminex models in 80 patients contributing 261 visits.

Baseline characteristics are shown in Table 1. The majority of patients were female, presented with four active joints or fewer and continued to have persistent oligoarthritis. Almost 90% of patients with oligoarticular disease onset and all patients with polyarticular onset presented with high disease activity, according to the cut-off points of the juvenile arthritis disease activity score (JADAS)-71.³⁴ Most patients achieved inactive disease during follow up, but a substantial minority (about 20-40%) showed active disease at the follow up visits (Table 2).

Primary analysis

The best GEE model for all patients combined is presented in supplementary tables 1, showing poor performance in test data (AUC 0.65).

Secondary analyses

The results of the Cox models for time to first remission, Cox models with recurrent events, the logistic regression at 6 and 12 months are shown in supplementary tables 2-5. All models, as well as the random forest and support vector machine algorithms (not shown), performed poorly in test data.

Improved results were obtained when fitting a GEE model to the oligoarticular patients, polyarticular RF negative patients and ANA positive patients separately. For oligoarticular patients, the best model was the model with clinical and microbiota predictors. The odds of achieving inactive disease were decreased by a higher JADAS-71 score at baseline and higher relative abundance of the operational taxonomic unit (OTU) *Mogibacteriaceae*. Other variables were associated in univariate analysis, but lost significance in the multivariable model (Table 3). The mean AUC over the imputed data sets was 0.79 in training data and 0.69 in test data (compared to 0.65 for the model with clinical variables only).

Table 1. Baseline characteristics

Baseline variables	N = 152
Female, <i>n</i> (%)	112 (73.7)
Age at onset, median [1 st , 3 rd quartile]	4.0 [2.1, 7.8]
Disease duration, median [1 st , 3 rd quartile]	0.2 [0.2, 0.3]
Age at diagnosis/enrolment, median [1 st , 3 rd quartile]	4.3 [2.4, 8.2]
Active joints, median [1 st , 3 rd quartile]	2 [1, 5]
More than 4 active joints, <i>n</i> (%)	43 (28.3)
JADAS-71, median [1 st , 3 rd quartile]	13.2 [8.2, 18.6]
ANA positive, <i>n</i> (%) ^a	84 (55.3)
RF positive, <i>n</i> (%) ^a	2 (1.3)
Uveitis, <i>n</i> (%)	8 (5.3)
Luminex data, <i>n</i> (%)	121 (79.6)
Gut microbiota data, <i>n</i> (%)	91 (59.9)
ILAR category at 6 months follow up	
Oligoarthritis persistent, <i>n</i> (%)	90 (59.2)
Oligoarthritis extended, <i>n</i> (%)	5 (3.3)
Polyarthritis RF positive, <i>n</i> (%)	2 (1.3)
Polyarthritis RF negative, <i>n</i> (%)	43 (28.3)
Psoriatic arthritis, <i>n</i> (%)	5 (3.3)
ERA, <i>n</i> (%)	4 (2.6)
Undifferentiated arthritis, <i>n</i> (%)	3 (2.0)

^a Missing clinical variables were imputed: ANA, *n* = 6; JADAS-71, *n* = 10; RF, *n* = 33.

Abbreviations: ANA, antinuclear antibodies; ILAR, international league against rheumatism; JADAS, juvenile arthritis disease activity score; RF, rheumatoid factor.

Table 2. Observed disease activity status according to Wallace criteria in the entire cohort and the three main subgroups.

Time point	Entire cohort		Oligoarticular		Polyarticular RF negative		ANA positive	
	N	ID, <i>n</i> (%)	N	ID, <i>n</i> (%)	N	ID, <i>n</i> (%)	N	ID, <i>n</i> (%)
6 months	151	89 (58.9)	51	34 (66.7)	29	14 (48.3)	87	49 (56.3)
12 months	139	103 (74.1)	47	36 (76.6)	27	18 (66.7)	81	58 (71.6)
18 months	110	87 (79.1)	41	34 (82.9)	23	17 (73.9)	64	51 (79.7)
24 months	82	61 (74.4)	31	24 (77.4)	16	10 (62.5)	53	45 (84.9)
Additional flare visits	26	0 (0)	9	0 (0)	5	0 (0)	13	0 (0)

Abbreviations: ANA, antinuclear antibody; ID, inactive disease; RF, rheumatoid factor.

The best model for the polyarticular RF negative patients was the model with clinical and Luminex predictors. Among the variables associated in univariate analysis, the hemoglobin level and the erythrocyte sedimentation rate (ESR) were correlated (Spearman's rho = -0.71), therefore the ESR was excluded. Likewise, the chemokine

Table 3. Best performing GEE model oligoarticular patients ($N = 52$)

Parameter	Visits ($m = 179$)	OR	95% CI	Wald p-value
Intercept		2.79	0.37-21.10	0.99
Duration of morning stiffness				
None or <15 minutes	105 (58.7)	Reference category		
15 minutes – 2 hours	49 (27.4)	1.28	0.38-4.27	0.69
>2 hours	25 (14.0)	0.26	0.05-1.45	0.12
Knee involvement (count)		0.67	0.29-1.53	0.34
JADAS-71		0.89	0.80-0.997	0.04
Christensenella		0.85	0.69-1.04	0.11
Mogibacteriaceae		0.82	0.68-0.98	0.03
Therapy during follow up				
None	53 (29.6)	Reference category		
IACI +/- NSAIDs	43 (24.0)	2.35	0.78-7.05	0.13
MTX +/- steroids or biologic agents	83 (46.4)	2.60	0.81-8.32	0.11
Interval between baseline and visit, y		2.31	0.89-5.96	0.08
AUC-like statistic				
		Minimum	Mean	Maximum
Training data		0.79	0.79	0.79
Test data		0.69	0.69	0.70

Included variables showed a univariate p-value <0.1 and no bivariate correlations (Spearman's $|\rho| \leq 0.6$).

Abbreviations: AUC, area under the curve; CI, confidence interval; GEE, generalized estimating equations; IACI, intra-articular corticosteroid injections; JADAS, juvenile arthritis disease activity score; MTX, methotrexate; NSAIDs, non-steroidal anti-inflammatory drugs; OR, odds ratio.

CXCL-9 correlated with soluble vascular endothelium derived growth factor receptor 1 (sVEGF-R1; Spearman's $\rho = 0.88$), therefore sVEGF-R1 was excluded. The multivariable model is shown in Table 4. The odds of achieving inactive disease were decreased by a longer duration of morning stiffness, lower hemoglobin levels and higher CXCL-9 levels at baseline (supplementary Figure 1). The mean AUC over the imputed data sets was 0.79 in training data and 0.69 in test data (marginally higher than the model with clinical predictors only, which had an AUC of 0.67). CXCL-9 was subsequently dichotomized. The best cut-off was at 30 pg/ml (supplementary Figure 1), yielding an AUC of 0.78 in training data and 0.69 in test data. The sensitivity of this cut-off to predict patients achieving inactive disease at maximally 50% of follow up visits was 0.31. The specificity was 0.94.

Finally, for the ANA positive patients, the best model was the clinical model (Table 5). The odds of achieving inactive disease were decreased by a duration of morning stiffness >2 hours and lower hemoglobin levels at baseline. The mean AUC over the imputed data sets was 0.79 in training data and 0.72 in test data.

The predicted probabilities of inactive disease according to the three models as a function of the significant predictors are illustrated in Figure 1.

Table 4. Best performing GEE model polyarticular RF negative patients ($N = 29$)

Parameter	Visits ($m = 100$)	OR	95% CI	Wald p-value
Intercept		0.02	0.0002-1.54	0.08
Duration of morning stiffness				
None or <15 minutes	21 (21)	Reference category		
15 minutes – 2 hours	54 (54)	0.26	0.08-0.84	0.02
>2 hours	25 (25)	0.27	0.09-0.85	0.03
Hemoglobin		1.52	1.07-2.16	0.02
CXCL-9		0.98	0.97-0.996	0.009
Therapy during follow up				
Biologic agents	14 (14)	2.80	0.96-8.16	0.06
Interval between baseline and visit, y		2.37	0.84-6.69	0.10
AUC-like statistic				
		Minimum	Mean	Maximum
Training data		0.79	0.79	0.79
Test data		0.69	0.69	0.69

Included variables showed a univariate p-value <0.1 and no bivariate correlations (Spearman's $|\rho| \leq 0.6$).

Abbreviations: AUC, area under the curve; CI, confidence interval; CXCL-9, chemokine C-X-C motif ligand 9; GEE, generalized estimating equations; OR, odds ratio; RF, rheumatoid factor.

DISCUSSION

The aim of this study was to develop a prediction model for the prognosis of JIA patients in the first two years of disease. The strength of the study was that we collected clinical, imaging, microbiota and Luminex data in a well-defined cohort of treatment-naïve JIA patients at onset. The prognostic value of these variables has never been analyzed before. We employed a large number of statistical algorithms, taking a longitudinal approach by incorporating outcome information of all visits, as well as performing survival analysis and cross-sectional analyses at 6 and 12 months after diagnosis and initiation of treatment. Moreover, machine learning algorithms such as random forest and support vector machines were used. Nevertheless, we were unable to construct a prediction model with satisfying accuracy in the prediction of inactive disease according to Wallace criteria, in all patients together.

Despite the failure to develop a model in all patients together, we were able to generate prediction models for oligoarticular, polyarticular RF negative and ANA positive patients separately, even though the performance in test data was still only moderate (AUC 0.69-0.72). For the oligoarticular patients, the odds of attaining inactive disease decreased following a higher JADAS-71 score and higher relative abundance of the OTU *Mogibacteriaceae* at baseline (Table 3). For polyarticular RF negative patients, the odds of inactive disease decreased following a longer duration of morning stiffness, lower hemoglobin levels and higher CXCL-9 levels at baseline (Table 4). Similarly, the odds of

Table 5. Best performing GEE model ANA positive patients ($N = 88$)

Parameter	Visits ($m = 298$)	OR	95% CI	Wald p-value
Intercept		0.001	$6.2 \times 10^{-6} - 0.17$	0.008
Duration of morning stiffness				
None	88 (29.5)	Reference category		
<15 minutes	39 (13.1)	2.16	0.55-8.40	0.27
15 minutes – 1 hour	79 (26.5)	0.78	0.25-2.37	0.66
1 hour – 2 hours	40 (13.4)	1.52	0.39-5.88	0.55
>2 hours	52 (17.4)	0.26	0.10-0.69	0.007
Hemoglobin		1.69	1.16-2.46	0.007
Knee involvement (count)		0.86	0.49-1.49	0.59
Wrist involvement (count)		0.76	0.43-1.33	0.33
Cervical spine involvement	11 (3.7)	0.38	0.07-2.21	0.28
JADAS-71		0.99	0.93-1.06	0.76
Therapy during follow up				
None	50 (16.8)	Reference category		
IACI +/- NSAIDs	56 (18.8)	3.03	0.99-9.26	0.05
MTX +/- steroids	151 (50.7)	2.24	0.79-6.32	0.13
Biologic agents	41 (13.8)	6.90	1.68-28.41	0.007
Interval between baseline and visit, y		3.83	1.75-8.35	0.0007
AUC-like statistic				
		Minimum	Mean	Maximum
Training data		0.78	0.79	0.81
Test data		0.71	0.72	0.74

Included variables showed a univariate p-value <0.1 and no bivariate correlations (Spearman's $|\rho| \leq 0.6$).

Abbreviations: ANA, antinuclear antibodies; AUC, area under the curve; IACI, intra-articular corticosteroid injections; JADAS, juvenile arthritis disease activity score; MTX, methotrexate; NSAIDs, non-steroidal anti-inflammatory drugs.

inactive disease decreased following a duration of morning stiffness >2 hours and lower hemoglobin levels at baseline, for ANA positive patients (Table 5).

Associations in a prediction model do not prove causality. Nonetheless, the predictors in these models merit further attention. Regarding the model for oligoarticular patients, intuitively it is convincing that patients with higher disease activity at baseline experience decreased odds of achieving inactive disease. Nothing is known about *Mogibacteriaceae* in the context of autoimmune diseases. In our previous analysis, there was no difference in the relative abundance of *Mogibacteriaceae* between JIA patients and healthy children [Van Dijkhuizen *et al.*, in preparation]. *Mogibacteriaceae* were less abundant in obese Japanese people, with respect to lean subjects,³⁵ and a decreased abundance of *Mogibacteriaceae* in the bronchial microbiome of asthmatic subjects has been observed.³⁶

Regarding the predictors for the polyarticular RF negative and the ANA positive patients, a longer duration of morning stiffness was associated with decreased

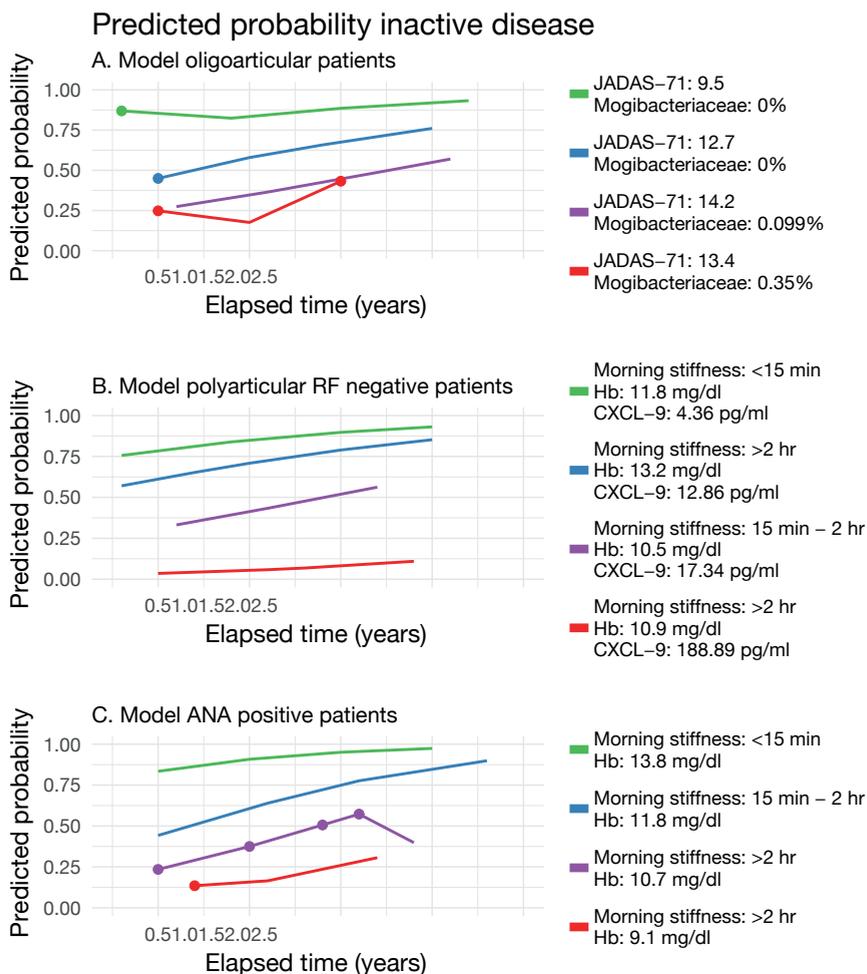


Figure 1. Predicted probabilities of inactive disease during two years of follow up according to the models for oligoarticular (A), polyarticular RF negative (B) and ANA positive (C) JIA patients. Shown are four selected patients with different values for the significant variables in each of the models, illustrating how the predicted probability of inactive disease varies across the values according to the model. Confidence intervals are not shown, but are very wide, indicating poor precision of the prediction. Patients are not the same across the three panels. The dots in panel A and C indicate visits following an intra-articular corticosteroid injection, causing the predicted probability of inactive disease to be higher for those visits.

odds of achieving inactive disease. Notably, this predictor was very consistent in all the algorithms that were employed (supplementary Table 1-5). In a previous analysis, it was also a determinant of the satisfaction of JIA patients with their current illness condition [Del Giudice *et al.*, in preparation]. These findings highlight the importance of the patient's opinion in the evaluation of disease activity and support the notion to take their evaluation into account in therapeutic decision-making. Indeed, the value of

the clinical JADAS as a criterion for treatment escalation improved when the patient's opinion was fully taken into account [Swart *et al.*, submitted]. Hemoglobin level was correlated with the ESR and therefore was thought to be reflective of anemia of chronic disease, making the relationship between a low hemoglobin level and decreased odds of achieving inactive disease convincing.

CXCL-9 is a pro-inflammatory chemokine induced by interferon (IFN)- γ and is a ligand of CXCR-3. It induces T helper 1 and 17 cells and recruits these to sites of inflammation. Indeed, CXCR-3 positive cells are highly enriched in the synovium of inflamed joints in rheumatoid arthritis³⁷⁻⁴⁰ and JIA.⁴¹ Moreover, due to its role in osteoclast activation, it may be involved in the development of bone erosions.⁴² In JIA, synovial expression of CXCL-10, another ligand of CXCR-3 with function very similar to CXCL-9, was increased.^{43,44} Taken together, the potential role of this chemokine in the pathogenesis of JIA merits further investigation. The cut-off of 30 pg/ml may be used to identify patients with worse prognosis, but needs validation in future studies.

A recent review of the literature showed that accurate prediction of JIA prognosis using clinical variables alone is unlikely to be achieved.¹⁷ Nevertheless, a clinical prediction model was published recently, showing good predictive performance in test data (AUC 0.85).¹⁸ The best AUC of this model was 0.68 in our data, obtained when classifying patients as those who never achieved inactive disease *versus* those who achieved it at least once. However, few patients ($n = 11$) were in the unfavorable category and only one of these was predicted correctly. Thus, these results reinforced the doubts about the prognostic value of clinical data. We are now the first to demonstrate that gut microbiota composition and Luminex data are insufficient to predict JIA prognosis as well. Nevertheless, improved prediction accuracy may be expected when separating JIA patients into more homogeneous subgroups. This finding supports previously made suggestions for a revision of the JIA classification criteria.⁴⁵⁻⁴⁸

A limitation of our study is the number of patients and visits with respect to the number of predictors tested. According to some, 40 cases are needed per screened predictor.⁴⁹ However, in our situation, this would amount to a cohort of over 10,000 children. It is not feasible to collect full clinical, immunological and gut microbiota data of such a large cohort of JIA patients. Nevertheless, the prediction models developed in the secondary analyses have to be interpreted with caution (especially the model for polyarticular RF negative patients). Many predictors in the multivariable analysis were not significant. This, coupled with the fact that the number of patients in the subgroups was substantially lower with respect to the full cohort, might indicate that the increased predictive performance was due to overfitting. Further validation and optimization of the models is needed.

In conclusion, the prognosis of JIA could not be predicted accurately in this well-defined cohort of treatment-naïve patients at disease onset, using clinical, imaging, microbiota and Luminex data and a wide array of statistical algorithms. Nevertheless, prediction accuracy improved when analyzing oligoarticular, polyarticular RF negative and ANA positive patients separately, suggesting that the heterogeneity of JIA impeded the accurate

prediction. The subgroup models showed that the duration of morning stiffness was associated with decreased odds of inactive disease, highlighting the importance of the patients' evaluation of disease activity. Moreover, *Mogibacteriaceae* were associated with lower odds of inactive disease in oligoarticular patients and CXCL-9 in polyarticular patients. Efforts should be directed at validating the prognostic value of this OTU and chemokine and at elucidating their potential role in the pathogenesis of JIA.

ACKNOWLEDGEMENTS

Multiplex immunoassays were in-house developed, validated and performed by the Multiplex core facility of the Laboratory for Translational Immunology (ISO9001 certified) of the University Medical Centre Utrecht.

FUNDING

This study has been supported by the European Commission, 7th FP, large integrated project: MD-Paedigree, Model Driven Paediatric European Digital Repository, Information Communication Technologies Programme (Contract Number 600932).

SUPPLEMENTARY MATERIAL

List of available covariates

Demographics

Sex, height, weight, age at onset, age at diagnosis, disease duration at diagnosis.

Disease characteristics

Subtype (at baseline, hence subject to change), onset subtype (oligoarticular or polyarticular, i.e., ≤ 4 joints or >4 joints), subtype at 6 months (ILAR subtype after 6 months), ANA positivity, RF positivity, HLA-b27 positivity, uveitis, enthesitis, presence and duration of morning stiffness.

ACR core set criteria

Physician's global assessment of disease activity, number of active joints, number of limited joints, VAS pain, VAS well-being, CHAQ score.

Joint count and scores

Involvement of the cervical spine, temporomandibular joint, shoulder, elbow, wrist, MCP, PIP, DIP, finger, upper limb, hip, knee, tibiotalar joint, subtalar joint, intertarsal joint, MTP, toe (the number of involved joints of each type was counted), symmetric involvement, JADAS-71, cJADAS.

Routine laboratory

Leukocyte count, neutrophil count, lymphocyte count, hemoglobin, thrombocyte count, ESR, CRP

Ultrasound data

Count of number of joints involved on ultrasound, count of number of tendons involved on ultrasound.

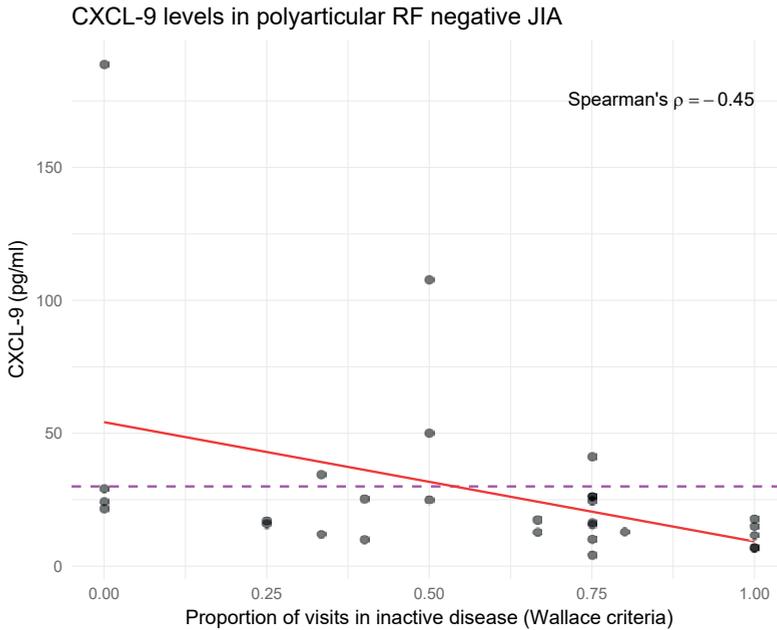
Cytokine data

IL-1RA, IL-1b, IL-6, IL-8, IL-10, IL-12, IL-13, IL-15, IL-17, IL-17F, IL-18, IL-22, IL-23, IL-25, IL-27, TNF α , IFN α , IFN γ , LIGHT, TWEAK, MIF, TSLP, MCP-1, MIP-1 α , MIP-1 β , MCP-2, TARC, MDC, MPIF, TECK, C-TACK, ENA-78, MIG, IP-10, BLC, OPN, SOST, GM-CSF, AR,

NGF, VEGF, sCD40L, HSP70, sPD.1, FAS, FAS-L, LAIR-1, IL.18Bpa, IL-1R1, IL-1R2, TNF-R1, TNF-R2, sCD19, sIL-2Ra, sCD27, sIL-7Ra, sVEGF.R1, Trappin-2, Endoglin, Dkk-1, S100A8, Vimentin, YKL-40, LAP, GDF-15, PARC, Leptin, Resistin, sICAM, sVCAM, MMP-1, MMP-3, MMP-8, MMP-9, TIMP-1, sIL-6R, C5a, sCD14.

Microbiota data

Shannon entropy, Chao 1 richness index, *Actinomyces*, *Akkermansia muciniphila*, *Anaerostipes*, *Anaerotruncus*, *Bacteroidaceae*, *Bacteroides*, *Bacteroides caccae*, *Bacteroides fragilis*, *Bacteroides ovatus*, *Bacteroides uniformis*, *Barnesiellaceae*, *Blautia*, *Blautiaproducta*, *Bulleidia*, *Christensenella*, *Christensenellaceae*, *Clostridiaceae*, *Clostridium*, *Coprobacillus*, *Coprococcus*, *Coriobacteriaceae*, *Dialister*, *Dorea*, *Eggerthella lenta*, *Enterobacteriaceae*, *Erysipelotrichaceae*, *Eubacterium dolichum*, *Faecalibacterium prausnitzii*, *Gemellaceae*, *Granulicatella*, *Haemophilus parainfluenzae*, *Holdemania*, *Klebsiella*, *Lachnospira*, *Lachnospiraceae*, *Mogibacteriaceae*, *Odoribacter*, *Oscillospira*, *Parabacteroides*, *Parabacteroides distasonis*, *Peptostreptococcaceae*, *Phascolarctobacterium*, *Prevotella*, *Prevotella copri*, *Pseudomonas*, *Rikenellaceae*, *Roseburia*, *Ruminococcaceae*, *Ruminococcus*, *Ruminococcus bromii*, *Ruminococcus gnavus*, *Sediminibacterium*, *Streptococcus*, *Sutterella*, *Turicibacter*, *Veillonella dispar*, Others (non-classifiable rest group).



Supplementary Figure 1. Scatterplot of the concentration of CXCL-9 in blood plasma at baseline against the proportion of follow up visits in inactive disease according to Wallace criteria, for polyarticular RF negative JIA patients. The plot shows a negative correlation (red line, Spearman's $\rho = -0.45$) between CXCL-9 and the proportion of visits in inactive disease, which remained even after removal of the two outliers with CXCL-9 >100 pg/ml (Spearman's $\rho = -0.39$). The results remained the same when calculating the correlation weighted on the number of follow up visits per patient.

The horizontal dashed line indicates the optimal cut-off value at 30 pg/ml. Of 13 patients with a proportion of visits in inactive disease $\leq 50\%$, 4 are above the cut-off (sensitivity 0.31). Of 16 patients with a proportion of visits in inactive disease $> 50\%$, 15 are below the cut-off (specificity 0.94).

Supplementary Table 1. Results clinical GEE model all patients ($N = 152$, $m = 508$ visits)

Parameter	OR	95% CI	Wald p-value
Intercept	0.28	0.007-10.68	0.50
Duration of morning stiffness			
None	Reference category		
<15 minutes	1.13	0.44-2.88	0.80
15 minutes – 1 hour	0.52	0.25-1.09	0.08
1 hour – 2 hours	0.71	0.26-1.95	0.50
>2 hours	0.24	0.10-0.59	0.002
Hemoglobin	1.15	0.88-1.50	0.32
Cervical spine involvement	0.50	0.15-1.62	0.25
TMJ involvement	0.53	0.32-0.88	0.01
Pain VAS	1.04	0.93-1.17	0.47
CHAQ score	1.33	0.78-2.25	0.30
JADAS-71	0.95	0.91-1.00	0.04
Therapy during follow up			
None	Reference category		
IACI +/- NSAIDs	1.41	0.72-2.75	0.31
MTX +/- steroids	1.35	0.69-2.62	0.38
Biologic agents	3.35	1.12-9.97	0.03
Interval between baseline and visit, y	2.12	1.31-3.45	0.002
AUC-like statistic			
	Minimum	Mean	Maximum
Training data	0.70	0.70	0.70
Test data	0.65	0.65	0.65

Abbreviations: AUC, area under the curve; CHAQ, childhood health assessment questionnaire; CI, confidence interval; GEE, generalized estimating equations; IACI, intra-articular corticosteroid injections; JADAS, juvenile arthritis disease activity score; MTX, methotrexate; NSAIDs, non-steroidal anti-inflammatory drugs; OR, odds ratio; TMJ, temporomandibular joints; VAS, visual analog scale.

Supplementary Table 2. Results clinical Cox proportional hazards model, all patients ($N = 152$)

Variable	HR	95% CI	P-value		
Duration of morning stiffness					
None	Reference category				
<15 minutes	1.36	0.76-2.41	0.30		
15 minutes – 1 hour	0.81	0.52-1.26	0.34		
1 – 2 hours	0.68	0.37-1.25	0.22		
>2 hours	0.53	0.29-0.96	0.04		
Therapy during follow up					
None/IACI/NSAIDs	Reference category				
MTX +/- steroids	0.49	0.33-0.74	0.0008		
Biological agents	0.41	0.24-0.70	0.001		
Performance in test data					
	C-statistic	Somers' Dxy	HR	95% CI	Log-rank p value
Minimum	0.62	0.24	2.16	1.23-3.81	0.16
Mean	0.62	0.25	2.16	1.24-3.81	0.16
Maximum	0.62	0.25	2.16	1.24-3.82	0.16

Abbreviations: CI, confidence interval; HR, hazard ratio; IACI, intra-articular corticosteroid injection; MTX, methotrexate, NSAIDs, non-steroidal anti-inflammatory drugs.

Supplementary Table 3. Results clinical Cox proportional hazards model with recurrent events, all patients ($N = 152$)

Variable	HR	95% CI	P-value
Presence of morning stiffness	0.64	0.49-0.83	0.0009
TMJ involvement	0.76	0.65-0.89	0.0009
Therapy during follow up			
None	Reference category		
IACI +/- NSAIDs	2.99	1.55-5.75	0.001
MTX +/- steroids	1.92	1.06-3.48	0.03
Biological agents	2.13	1.14-3.97	0.02
Performance in test data			
HR	95%-CI	P-value	R ²
2.30	1.33-4.00	0.03	0.05

Abbreviations: CI, confidence interval; HR, hazard ratio; IACI, intra-articular corticosteroid injection; MTX, methotrexate; NSAIDs, non-steroidal anti-inflammatory drugs; TMJ, temporomandibular joints.

Supplementary Table 4. Results clinical logistic regression 6 months outcome, all patients ($N = 152$)

Variable	OR	95% CI	P-value
Intercept	4.04	1.94-8.40	0.0003
Duration of morning stiffness			
None	Reference category		
<15 minutes	2.06	0.52-8.13	0.30
15 minutes – 1 hour	0.60	0.22-1.60	0.31
1 – 2 hours	1.14	0.34-3.86	0.84
>2 hours	0.25	0.06-1.00	0.05
Parent/patient VAS well-being	0.92	0.77-1.09	0.32
Therapy during follow up			
None/IACI +/- NSAIDs	Reference category		
MTX +/- steroids	0.43	0.19-0.94	0.04
Biological agents	0.54	0.12-2.37	0.42

AUC

	Minimum	Mean	Maximum
Training data	0.73	0.74	0.75
Test data	0.65	0.67	0.68

Abbreviations: AUC, area under the curve; CI, confidence interval; IACI, intra-articular corticosteroids injection; MTX, methotrexate; NSAIDs, non-steroidal anti-inflammatory drugs; OR, odds ratio; VAS, visual analog scale.

Supplementary Table 5. Results Luminex logistic regression 12 months outcome, all patients ($N = 121$)

Variable	OR	95% CI	P-value
Intercept	0.21	0.02-2.70	0.23
Duration of morning stiffness			
None	Reference category		
<15 minutes	0.36	0.04-3.28	0.37
15 minutes – 1 hour	0.31	0.07-1.41	0.13
1 – 2 hours	0.21	0.04-1.20	0.08
>2 hours	0.10	0.01-0.67	0.02
sICAM ($\times 1000$ pg/mL)	1.005	1.001-1.008	0.008
Therapy during follow up			
None/IACI +/- NSAIDs	Reference category		
MTX +/- steroids	0.75	0.17-3.31	0.70
Biological agents	0.92	0.12-7.01	0.93

AUC

	Minimum	Mean	Maximum
Training data	0.78	0.80	0.83
Test data	0.63	0.65	0.68

Abbreviations: AUC, area under the curve; CI, confidence interval; IACI, intra-articular corticosteroids injection; MTX, methotrexate; NSAIDs, non-steroidal anti-inflammatory drugs; OR, odds ratio; VAS, visual analog scale.

REFERENCE LIST

1. Ravelli A, Martini A. Juvenile idiopathic arthritis. *Lancet* 2007; 369(9563):767-778.
2. Guzman J, Oen K, Tucker LB, Huber AM, Shiff N, Boire G et al. The outcomes of juvenile idiopathic arthritis in children managed with contemporary treatments: results from the ReACCh-Out cohort. *Ann Rheum Dis* 2015; 74(10):1854-1860.
3. Foell D, Wulffraat N, Wedderburn LR, Wittkowski H, Frosch M, Gerss J et al. Methotrexate withdrawal at 6 vs 12 months in juvenile idiopathic arthritis in remission: a randomized clinical trial. *JAMA* 2010; 303(13):1266-1273.
4. Ruperto N, Murray KJ, Gerloni V, Wulffraat N, deOliveira SK, Falcini F et al. A randomized trial of parenteral methotrexate comparing an intermediate dose with a higher dose in children with juvenile idiopathic arthritis who failed to respond to standard doses of methotrexate. *Arthritis Rheum* 2004; 50(7):2191-2201.
5. Solari N, Palmisani E, Consolaro A, Pistorio A, Viola S, Buoncompagni A et al. Factors associated with achievement of inactive disease in children with juvenile idiopathic arthritis treated with etanercept. *J Rheumatol* 2013; 40(2):192-200.
6. Palmisani E, Solari N, Magni-Manzoni S, Pistorio A, Labo E, Panigada S et al. Correlation between juvenile idiopathic arthritis activity and damage measures in early, advanced, and longstanding disease. *Arthritis Rheum* 2006; 55(6):843-849.
7. Nielsen S, Ruperto N, Gerloni V, Simonini G, Cortis E, Lepore L et al. Preliminary evidence that etanercept may reduce radiographic progression in juvenile idiopathic arthritis. *Clin Exp Rheumatol* 2008; 26(4):688-692.
8. Taxter AJ, Wileyto EP, Behrens EM, Weiss PF. Patient-reported Outcomes across Categories of Juvenile Idiopathic Arthritis. *J Rheumatol* 2015; 42(10):1914-1921.
9. Magni-Manzoni S, Rossi F, Pistorio A, Temporini F, Viola S, Beluffi G et al. Prognostic factors for radiographic progression, radiographic damage, and disability in juvenile idiopathic arthritis. *Arthritis Rheum* 2003; 48(12):3509-3517.
10. Oen K, Reed M, Malleson PN, Cabral DA, Petty RE, Rosenberg AM et al. Radiologic outcome and its relationship to functional disability in juvenile rheumatoid arthritis. *J Rheumatol* 2003; 30(4):832-840.
11. Boers M. Understanding the window of opportunity concept in early rheumatoid arthritis. *Arthritis Rheum* 2003; 48(7):1771-1774.
12. Moreland LW, Bridges SL, Jr. Early rheumatoid arthritis: a medical emergency? *Am J Med* 2001; 111(6):498-500.
13. Mottonen T, Hannonen P, Korpela M, Nissila M, Kautiainen H, Ilonen J et al. Delay to institution of therapy and induction of remission using single-drug or combination-disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. *Arthritis Rheum* 2002; 46(4):894-898.
14. O'Dell JR. Treating rheumatoid arthritis early: a window of opportunity? *Arthritis Rheum* 2002; 46(2):283-285.
15. Quinn MA, Emery P. Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. *Clin Exp Rheumatol* 2003; 21(5 Suppl 31):S154-S157.
16. Ravelli A, Davi S, Bracciolini G, Pistorio A, Consolaro A, van Dijkhuizen EH et al. Intra-articular corticosteroids versus intra-articular corticosteroids plus methotrexate in oligoarticular juvenile idiopathic arthritis: a multicentre, prospective, randomised, open-label trial. *Lancet* 2017; 389(10072):909-916.
17. van Dijkhuizen EHP, Wulffraat NM. Early predictors of prognosis in juvenile idiopathic arthritis: a systematic literature review. *Ann Rheum Dis* 2015; 74(11):1996-2005.
18. Guzman J, Henrey A, Loughin T, Berard RA, Shiff NJ, Jurecsek R et al. Predicting Which Children with Juvenile Idiopathic Arthritis Will Have a Severe Disease Course: Results from the ReACCh-Out Cohort. *J Rheumatol* 2017; 44(2):230-240.

19. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004; 31(2):390-392.
20. de Jager W, Prakken BJ, Bijlsma JW, Kuis W, Rijkers GT. Improved multiplex immunoassay performance in human plasma and synovial fluid following removal of interfering heterophilic antibodies. *J Immunol Methods* 2005; 300(1-2):124-135.
21. De Filippis F, La SA, Villani F, Ercolini D. Exploring the sources of bacterial spoilers in beefsteaks by culture-independent high-throughput sequencing. *PLoS One* 2013; 8(7):e70222.
22. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010; 7(5):335-336.
23. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; 26(19):2460-2461.
24. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 2012; 6(3):610-618.
25. Navas-Molina JA, Peralta-Sanchez JM, Gonzalez A, McMurdie PJ, Vazquez-Baeza Y, Xu Z et al. Advancing our understanding of the human microbiome using QIIME. *Methods Enzymol* 2013; 531:371-444.
26. Wallace CA, Giannini EH, Huang B, Irtter L, Ruperto N. American College of Rheumatology provisional criteria for defining clinical inactive disease in select categories of juvenile idiopathic arthritis. *Arthritis Care Res (Hoboken)* 2011; 63(7):929-936.
27. Van Buuren S, Groothuis-Oudshoorn K. MICE: Multivariate Imputation by Chained Equations in R. *J Stat Soft* 2011; 45(3):1-67.
28. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; 73(1):13-22.
29. Lim LS, Pullenayegum E, Moineddin R, Gladman DD, Silverman ED, Feldman BM. Methods for analyzing observational longitudinal prognosis studies for rheumatic diseases: a review & worked example using a clinic-based cohort of juvenile dermatomyositis patients. *Pediatr Rheumatol Online J* 2017; 15(1):18.
30. Davarzani N, Peeters R, Smirnov E, Karel J, Brunner-La Rocca HP. Ranking Accuracy for Logistic-GEE Models. In: Boström H, Knobbe A, Soares C, Papapetrou P, editors. *Advances in Intelligent Data Analysis XV. IDA 2016. Lecture Notes in Computer Science*. Cham: Springer; 2016. 14-25.
31. Amorim LD, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. *Int J Epidemiol* 2015; 44(1):324-333.
32. Chen HC, Kodell RL, Cheng KF, Chen JJ. Assessment of performance of survival prediction models for cancer prognosis. *BMC Med Res Methodol* 2012; 12:102.
33. Harrell FE, Jr., Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA* 1982; 247(18):2543-2546.
34. Consolaro A, Ruperto N, Bracciolini G, Frisina A, Gallo MC, Pistorio A et al. Defining criteria for high disease activity in juvenile idiopathic arthritis based on the juvenile arthritis disease activity score. *Ann Rheum Dis* 2014; 73(7):1380-1383.
35. Oki K, Toyama M, Banno T, Chonan O, Benno Y, Watanabe K. Comprehensive analysis of the fecal microbiota of healthy Japanese adults reveals a new bacterial lineage associated with a phenotype characterized by a high frequency of bowel movements and a lean body type. *BMC Microbiol* 2016; 16(1):284.
36. Durack J, Lynch SV, Nariya S, Bhakta NR, Beigelman A, Castro M et al. Features of the bronchial bacterial microbiome associated with atopy, asthma, and responsiveness to inhaled corticosteroid treatment. *J Allergy Clin Immunol* 2016.
37. Karin N, Wildbaum G, Thelen M. Biased signaling pathways via CXCR3 control the development and function of CD4+ T cell subsets. *J Leukoc Biol* 2016; 99(6):857-862.

38. Karin N, Wildbaum G. The role of chemokines in adjusting the balance between CD4+ effector T cell subsets and FOXP3-negative regulatory T cells. *Int Immunopharmacol* 2015; 28(2):829-835.
39. Van Raemdonck K, Van den Steen PE, Liekens S, Van Damme J, Struyf S. CXCR3 ligands in disease and therapy. *Cytokine Growth Factor Rev* 2015; 26(3):311-327.
40. Lacotte S, Brun S, Muller S, Dumortier H. CXCR3, inflammation, and autoimmune diseases. *Ann N Y Acad Sci* 2009; 1173:310-317.
41. Wedderburn LR, Robinson N, Patel A, Varsani H, Woo P. Selective recruitment of polarized T cells expressing CCR5 and CXCR3 to the inflamed joints of children with juvenile idiopathic arthritis. *Arthritis Rheum* 2000; 43(4):765-774.
42. Maruotti N, d'Onofrio F, Corrado A, Cantatore FP. Why TNF-alpha inhibition is not sufficient to avoid juxta-articular erosions in chronic arthritis? *Intern Emerg Med* 2012; 7(1):15-20.
43. Pharoah DS, Varsani H, Tatham RW, Newton KR, de JW, Prakken BJ et al. Expression of the inflammatory chemokines CCL5, CCL3 and CXCL10 in juvenile idiopathic arthritis, and demonstration of CCL5 production by an atypical subset of CD8+ T cells. *Arthritis Res Ther* 2006; 8(2):R50.
44. Martini G, Zulian F, Calabrese F, Bortoli M, Facco M, Cabrelle A et al. CXCR3/CXCL10 expression in the synovium of children with juvenile idiopathic arthritis. *Arthritis Res Ther* 2005; 7(2):R241-R249.
45. Ravelli A, Varnier GC, Oliveira S, Castell E, Arguedas O, Magnani A et al. Antinuclear antibody-positive patients should be grouped as a separate category in the classification of juvenile idiopathic arthritis. *Arthritis Rheum* 2011; 63(1):267-275.
46. van den Broek T, Hoppenreijs E, Meerding J, Scholman R, Otten HG, Swart JF et al. Cytokine profiling at disease onset: support for classification of young antinuclear antibody-positive patients as a separate category of juvenile idiopathic arthritis. *Ann Rheum Dis* 2015; 74(2):470-472.
47. Martini A. Are the number of joints involved or the presence of psoriasis still useful tools to identify homogeneous disease entities in juvenile idiopathic arthritis? *J Rheumatol* 2003; 30(9):1900-1903.
48. Martini A. It is time to rethink juvenile idiopathic arthritis classification and nomenclature. *Ann Rheum Dis* 2012; 71(9):1437-1439.
49. Tabachnick BG, Fidell LS. Using multivariate statistics. 6 ed. Boston: Pearson Education; 2013.

Chapter

4

MICROBIOME ANALYTICS OF THE GUT MICROBIOTA IN JUVENILE IDIOPATHIC ARTHRITIS PATIENTS: AN OBSERVATIONAL, LONGITUDINAL COHORT STUDY

E.H. Pieter van Dijkhuizen, Federica Del Chierico,
Clara Malattia, Alessandra Russo, Denise Pires Marafon,
Nienke M. ter Haar, Silvia Magni-Manzoni,
Sebastiaan J. Vastert, Bruno Dallapiccola, Berent Prakken,
Alberto Martini, Fabrizio De Benedetti
and Lorenza Putignani for the MD-Paedigree consortium

Submitted

ABSTRACT

Objectives

To assess the composition of gut microbiota in Italian and Dutch juvenile idiopathic arthritis (JIA) patients at baseline, in inactive disease and persistent activity, compared to healthy controls.

Methods

In a prospective, multicenter, observational cohort study, fecal samples were collected of 78 Italian and 21 Dutch treatment-naïve JIA patients at baseline with less than 6 months disease duration and compared to 107 geography-matched healthy control samples. Furthermore, 44 follow up samples in inactive disease and 25 in persistent activity were analyzed. Gut microbiota composition was determined by 16S rRNA-based-metagenomics. The α - and β -diversity were computed, and log-ratios of relative abundance were compared between patients and healthy controls using random forest models and logistic regression.

Results

Patient samples at baseline showed reduced richness compared to healthy controls ($p = 0.004$). Random forest distinguished Italian and Dutch baseline samples from controls (area under the curve in test data >0.99 and 0.71 , respectively). Mainly the Operational Taxonomic Units (OTUs) *Erysipelotrichaceae*, *Allobaculum* and *Faecalibacterium prausnitzii* showed different relative abundance in Italian baseline samples compared to controls (false discovery rate [FDR] correction by the Benjamini-Hochberg procedure). Some OTUs differed between Dutch samples and healthy controls, but no evidence remained after FDR correction. No differences were found in paired analysis between Italian baseline and inactive disease samples.

Conclusions

We found evidence for dysbiosis in JIA patients. Only patient/control status, age and geographical origin appeared to be drivers of the microbiota profiles, regardless of disease activity stage, inflammation and autoimmunity markers.

INTRODUCTION

The human body is host to a myriad of non-pathogenic micro-organisms, the so-called microbiota. Many factors contribute to its composition, among which are mode of delivery, diet and drug use.¹ The interaction between host and microbiota is complex and not yet fully elucidated. What is known is that the microbiota plays a role in host-defense against pathogens and provides the host with nutrients and vitamins.^{1,2}

Recent studies have demonstrated alterations in the composition of gut microbiota in patients with autoimmune disorders, such as rheumatoid arthritis and type 1 diabetes mellitus.^{1,3-5} These alterations of the microbial profile induce a dysbiosis state of the microbiota also affecting functional activities of the gut.^{3,6} It is still unclear if there is any causal relationship underlying the association between dysbiosis and autoimmune disorders. However, various observations in germ-free mice are compatible with such a pathogenic role of gut dysbiosis.^{1,7,8}

Preliminary observations have suggested the presence of dysbiosis also in pediatric patients with juvenile idiopathic arthritis (JIA).⁹⁻¹² JIA is defined as any arthritis of unknown etiology, lasting for more than six weeks, with an onset before 16 years of age.¹³ The findings in JIA are interesting because of the hypothesized pathogenic role of the gut microbiota and potential therapeutic ramifications. However, the previous studies of microbiota in JIA were inconclusive, since they were small cross-sectional studies, and some included patients with long-standing disease treated with disease-modifying anti-rheumatic drugs (DMARDs). Differences between patients with active and inactive disease have only been explored in small groups. Finally, the statistical analysis used in these studies can lead to a marked increase of the false discovery rate (FDR), up to 68%.¹⁴

The hypothesis of dysbiosis in JIA patients needs therefore to be explored further. This study addresses the aforementioned problems as it is a large, multicenter, longitudinal study of the gut microbiota composition in a well-defined cohort of JIA patients at baseline and in inactive disease, compared to healthy children.

METHODS

We conducted a multicenter, observational, prospective cohort study in Italy (Istituto Giannina Gaslini [IGG] and Ospedale Pediatrico Bambino Gesù [OPBG]) and The Netherlands (Wilhelmina Children's hospital [WKZ]). All consecutive patients were enrolled who presented between October 2013 and December 2015 with new-onset juvenile idiopathic arthritis (JIA) according to International League Against Rheumatism (ILAR) criteria¹⁵ and less than six months disease duration since the first disease symptoms. Patients with systemic JIA and patients previously treated with disease modifying anti-rheumatic drugs (DMARDs), intra-articular steroid injections or systemic steroids were excluded.

Patients were included at the time of diagnosis and followed up every six months until 24 months. Collected clinical data included patient demographics, JIA category, full joint count, juvenile arthritis disease activity score (JADAS) -71,¹⁶ physician's global assessment of disease activity, parent or patient global assessment of well-being, the presence of active uveitis, anti-nuclear antibody (ANA) positivity, rheumatoid factor (RF) positivity, erythrocyte sedimentation rate (ESR), c-reactive protein (CRP) and the use of non-steroidal anti-inflammatory drugs (NSAIDs). At the follow up visits, data was collected regarding treatment and disease activity status, defined according to the Wallace criteria for inactive disease.¹⁷

The study was approved by the local ethics committees and was conducted according to the principles of good clinical practice and the declaration of Helsinki. Written informed consent was obtained of all participants.

Fecal samples

A fecal sample was collected at baseline and during follow up in case of inactive disease, persistent activity or disease flare. Details of sample collection and analysis are specified in the supplementary material.

At the time of withdrawal of the samples, information was collected regarding diet, previous antibiotic, prebiotic and probiotic use, as well as gastro-intestinal infections and diagnosis of inflammatory bowel disorder.

Additionally, healthy control samples, matched on geography with baseline samples of patients, were collected. Data of Italian healthy controls were available in the biobank of OPBG, whereas Dutch healthy control samples were collected at primary and secondary schools in the vicinity of the hospital.

Statistical analysis

To assess overall differences in microbial community structure between patient groups and healthy controls, we calculated measures of α -diversity by Chaol and Shannon indices using the "alpha_rarefaction.py" script of the QIIME software.¹⁸ The former summarizes microbial diversity (richness) per sample accounting for rare species, which might have been missed due to undersampling. The Shannon index is a measure of entropy taking account of species richness and evenness in the sample. These indices were compared between patients and controls in linear models using age, sex and geography as covariates. Moreover, measures of β -diversity were calculated by unweighted UniFrac,¹⁹ using the scripts, "beta_diversity_through_plots.py" and "compare_categories.py" of the QIIME software.¹⁸ They were compared across groups using principal coordinates analysis and Adonis. These measures of β -diversity, as opposed to α -diversity, take account of the distance of OTUs along the phylogenetic tree.

OTUs present in less than 20% of the samples were removed from the analysis. Since microbiota data are expressed as relative abundances of the OTUs, the sum of which equals to 100% for each sample, they cannot be analyzed using standard statistical techniques.^{14;20;21} Rather, the analysis should be based on log-ratios of the individual parts with the geometric mean of the sample, according to the framework of compositional data analysis.^{20;21} This was performed using the package “compositions” in the statistical software R. Since it is impossible to take the logarithm of 0, zeros were imputed by means of robust imputation using the package “robCompositions”, as specified in literature.^{22;23} Next, the log-ratio-transformed OTU relative abundances were analyzed using principal components analysis (PCA), showing that Italian and Dutch samples clustered in two different clusters (supplementary figure 1). Therefore they were analyzed separately. Likewise, baseline, inactive disease and persistent activity samples were analyzed separately.

A random forest classification model was employed to classify Italian baseline and control samples, using the microbiota composition, age and sex. To this end, the data were split at random in 70% for model training and 30% for model testing. The best model was selected in the training set, using 10-fold cross-validation, repeated 10 times. This model was then applied to the test set to assess its performance. The same procedure was used for the Dutch samples.

Differences in the log-ratio-transformed relative abundances of all OTUs between baseline samples and healthy controls, inactive disease samples and healthy controls and persistent activity samples and healthy controls were assessed with logistic regression, using age and sex as covariates and patient/control status as dependent variable. A separate model was fitted for each of the OTUs and each of the disease states. As before, Italian patients and healthy controls were analyzed separately from Dutch patients and healthy controls. The false discovery rate (FDR) was controlled by the Benjamini-Hochberg procedure.

Finally, associations between the gut microbiota compositions and various clinical variables were explored using PCA, principal components regression and k-means clustering.

Statistical analysis was carried out using R statistics version 3.3.2 (The R foundation for statistical computing, Vienna, Austria).

RESULTS

In total, 169 patients were included. Baseline fecal samples were collected of 99 patients (78 Italian and 21 Dutch). Samples of the remainder of patients were not collected due to difficulties collecting or shipping the samples. There were no differences between patients of whom a sample was taken and patients of whom no sample was taken (data

not shown). Baseline characteristics of the analyzed patients are shown in Table 1 and reflect characteristics common to a population with recent onset JIA. The majority of participants were female and most had oligoarticular JIA. Disease activity at baseline was moderate to high according to the JADAS-71.^{16;24} Corresponding to clinical observations, the age at disease onset was higher in Dutch patients compared to Italians.

The baseline samples were matched with 107 healthy control samples, with similar age and sex (Table 1). Furthermore, 44 patients contributed at least one sample in inactive disease (39 Italian and 5 Dutch; supplementary Table 1). These samples were collected a median time of 1.0 year after diagnosis (range 0.3-2.0 years). Persistent activity or flare samples were available of 25 patients (24 Italian and 1 Dutch; supplementary Table 2), collected a median time of 0.5 years after diagnosis (range 0.1-1.1 years).

After data filtering, a total of 998,583 sequence reads of 16S rRNA gene amplicons were obtained with an average of 3,210 reads/sample and an average length of 487 base pairs, calculated after primer removal.

In all samples together, 321 different operational taxonomic units (OTUs) were discovered. After removal of those OTUs present in less than 20% of the samples, 57 remained for analysis (supplementary Table 3).

Baseline samples

Linear models with age, gender and geography as covariates showed strong evidence for a reduced richness indicated by a lower Chaol index for baseline samples with respect to healthy controls (β -73; 95% confidence interval [CI] -124, -23; $p=0.005$; Figure 1A). The model also showed evidence that richness was reduced for Dutch female healthy controls samples, compared to the average (Figure 1A). There was no evidence of differences in the Shannon index. β -diversity differed between Italian baseline samples and healthy controls ($p=0.001$) as well as between Dutch baseline samples and healthy controls ($p=0.013$; Figure 1B).

The random forest classification algorithm separated Italian baseline patients from healthy controls almost perfectly, with an area under the curve (AUC) of the receiver operating characteristics (ROC) curve >0.99 in both training and test data. This was mainly due to the OTUs *Allobaculum*, *Erysipelotrichaceae*, *Propionibacterium acnes* and *Barnesiellaceae*.

In the comparison of Italian samples at baseline and healthy controls, *Erysipelotrichaceae*, *F. prausnitzii*, *Parabacteroides*, *Enterococcus* and *Ruminococcaceae* were increased, whereas *Allobaculum*, *Gemellaceae*, *Propionibacterium acnes* and *Turcibacter* were less abundant in patients with respect to healthy controls (false discovery rate [FDR] corrected p -value <0.05 ; see Table 2 for OTUs with crude p -value <0.05 , supplementary Table 4 for full list). Figure 2 shows the relative abundances of the four OTUs with the strongest evidence of differences between patients and healthy controls.

Table 1. Clinical features of patients and healthy controls at baseline.

	Italian		Dutch	
	Patients	Controls	Patients	Controls
N	78	79	21	28
Female, <i>n</i> (%)	61 (78.2)	38 (48.1)	16 (76.2)	20 (71.4)
Age, median [IQR]	3.9 [2.3, 6.7]	8 [5, 11]	8.5 [4.8, 11.7]	8.0 [4.8, 11.2]
JIA category at onset				
Oligoarticular, <i>n</i> (%)	47 (60.3)		14 (66.7)	
Polyarticular RF positive, <i>n</i> (%)	2 (2.6)		0 (0)	
Polyarticular RF negative, <i>n</i> (%)	24 (30.8)		4 (19.0)	
Enthesitis-related, <i>n</i> (%)	0 (0)		2 (9.5)	
Psoriatic, <i>n</i> (%)	3 (3.8)		1 (4.8)	
Undifferentiated, <i>n</i> (%)	2 (2.6)		0 (0)	
Number of active joints, median [IQR]	3 [1, 5]		2 [1, 4]	
PGA, median [IQR]	6 [4, 8]		2.0 [1.5, 3.0]	
Parent GA, median [IQR]	3.1 [1.0, 5.2]		5.7 [2.0, 8.1]	
JADAS-71, median [IQR]	13.2 [8.6, 19.0]		9.5 [6.6, 13.5]	
ESR, median [IQR]	23 [14, 43]		11 [6, 31]	
ANA, <i>n</i> (%)	51 (65.4)		8 (38.1)	
RF, <i>n</i> (%)	2 (2.6)		0 (0)	
NSAIDs, <i>n</i> (%)	62 (79.5)		17 (81.0)	
Antibiotics, <i>n</i> (%)	13 (16.7)	0 (0)	0 (0)	0 (0)
Prebiotics, <i>n</i> (%)	14 (17.9)	0 (0)	8 (38.1)	9 (32.1)
Probiotics, <i>n</i> (%)	14 (17.9)	0 (0)	0 (0)	3 (10.7)
GI infections, <i>n</i> (%)	11 (14.1)	0 (0)	3 (14.3)	1 (3.6)
GI disorders, <i>n</i> (%) ^a	6 (7.7)	0 (0)	2 (9.5)	0 (0)

Abbreviations: ANA, anti-nuclear antibodies; ESR, erythrocyte sedimentation rate; GA, global assessment of disease activity; GI, gastro-intestinal; IQR, interquartile range; JADAS, juvenile arthritis disease activity score; JIA, juvenile idiopathic arthritis; NSAID, non-steroidal anti-inflammatory drug; PGA, physician global assessment of disease activity; RF, rheumatoid factor.

^a GI disorders were: constipation (*n* = 3), celiac disease (*n* = 2), milk and egg allergy (*n* = 1), abdominal bloating (*n* = 1), irritable bowel syndrome and lactose intolerance (*n* = 1).

Dutch baseline samples were separated well from healthy control samples by the random forest algorithm in the training set (AUC of the ROC curve >0.99). The most important OTUs for this separation were *Eggerthella lenta*, *Rikenellaceae*, *Coprobacillus* and *Mogibacteriaceae*. In the test data set (*n* = 14), the AUC of the ROC curve was 0.71 (95% CI 0.41, >0.99). Various OTUs showed a difference in relative abundance in logistic regression models, but no evidence of differences was maintained after FDR correction (Figure 3, supplementary Table 5).

Figure 1. Plots of α - (A) and β - (B-C) diversity. A: Boxplots of the Chaol index for Italian and Dutch samples of male and female subjects. Shown are the distributions for healthy controls, baseline samples, inactive disease and persistent activity. Linear models of the Chaol index with age, sex and geography as covariates showed evidence of reduced richness in baseline samples, inactive disease and persistent activity as compared to healthy controls. The plot show that sex, even though not a determinant of the microbiota composition in itself, confounded the relation between disease activity state and microbiota richness. B-C: Principal coordinates analysis of measures of β -diversity, calculated by unweighted UniFrac for Italian (B) and Dutch (C) subjects. Shown are patients at baseline and controls (left panels), and baseline, inactive disease, persistent activity and healthy controls (right panels). Adonis showed evidence of pairwise differences between all subgroups, except for persistent activity vs. inactive disease samples. **Abbreviations:** ID, inactive disease; PA, persistent activity; PC, principal coordinate; PCoA, principal coordinates analysis. ▶

Table 2. Differences in relative abundance between Italian baseline samples and healthy controls by separate logistic regressions.

OTUs ^a	Relative abundance in patients	Crude p-value	Adjusted p-value ^b
Erysipelotrichaceae	Increased	<0.001	<0.001
Allobaculum	Decreased	<0.001	<0.001
F. prausnitzii	Increased	<0.001	0.009
Gemellaceae	Decreased	0.001	0.016
Propionibacterium acnes	Decreased	0.002	0.016
Parabacteroides	Increased	0.002	0.016
Enterococcus	Decreased	0.003	0.025
Turicibacter	Decreased	0.004	0.025
Ruminococcaceae	Increased	0.005	0.033
Phascolarctobacterium	Increased	0.012	0.07
Blautia	Decreased	0.017	0.09
Barnesiellaceae	Decreased	0.026	0.12
Dorea	Increased	0.037	0.16

^a Selected list of OTUs with crude p-value < 0.05. The full list can be accessed in supplementary Table 4.

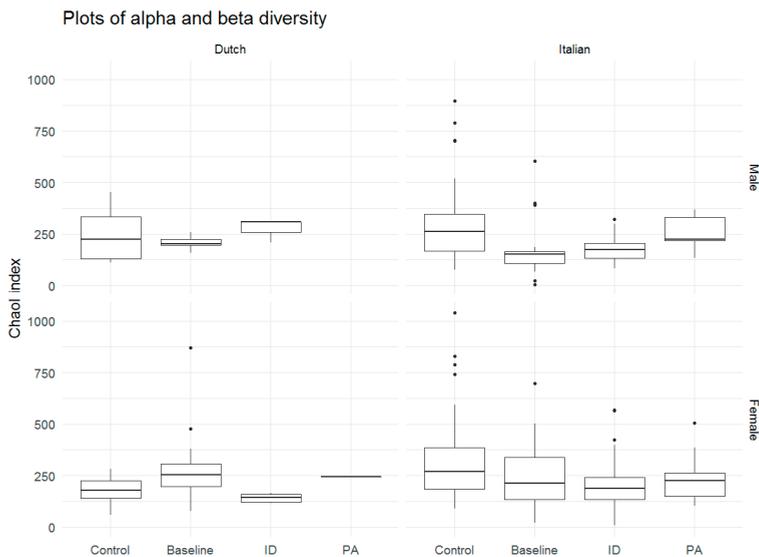
^b False discovery rate control using the Benjamini-Hochberg procedure.

Inactive disease samples

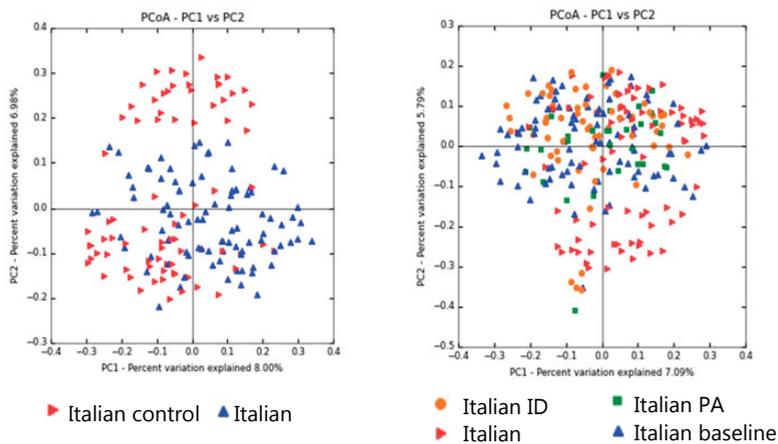
There was strong evidence of a reduced richness as evidenced by a lower Chaol measure in inactive disease samples versus healthy controls using a linear model with age, gender and geography as covariates (β -141; 95% CI -202, -81; $p < 0.001$; Figure 1). There was no evidence of differences in the Shannon measure. The analysis of β -diversity showed evidence of differences between inactive disease samples and healthy controls ($p = 0.001$) as well as between inactive disease samples and baseline samples ($p = 0.035$; Figure 1)

Various OTUs in Italian inactive disease samples were differently abundant compared to control samples (supplementary Table 6). Figure 4 shows the regression coefficients and false coverage rate-corrected CIs in comparison with those obtained in

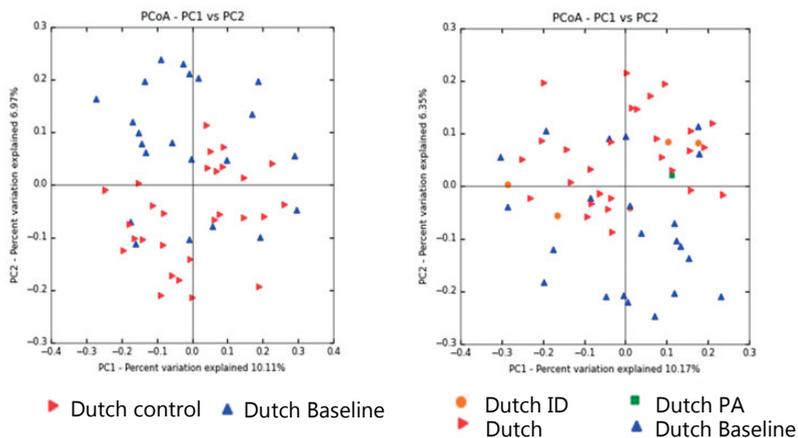
A



B



C



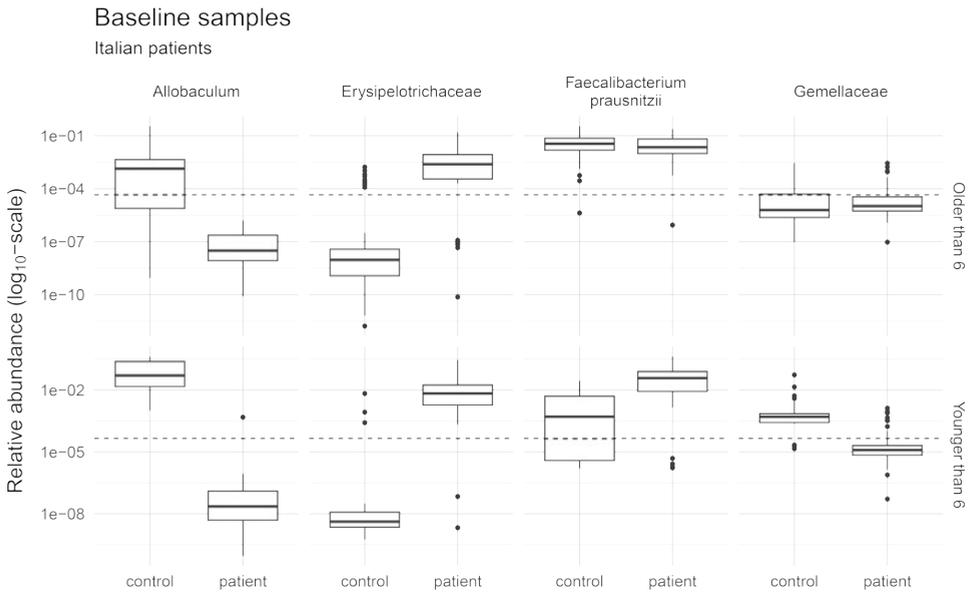


Figure 2. Boxplots of the relative abundance of the four OTUs with strongest evidence of differences between Italian baseline samples and healthy controls. The age of the subjects confounded the relationship between patient-control status and the relative abundance. Therefore, the plots were divided into patients older than the median age of 6 years (upper row) and younger than the median age (bottom row). Shown are the relative abundances; the analysis was performed on the log-ratio of the relative abundances. The horizontal dashed line represents the limit of detection at a relative abundance of 4.5×10^{-5} . Relative abundances below this line were not observed, but were the results of imputation. Note the \log_{10} -transformation of the y-axis.

the analysis of baseline samples. *Erysipelotrichaceae* and *F. prausnitzii* were increased in both baseline and inactive disease samples compared to healthy controls, whereas *Allobaculum* and *Propionibacterium acnes* were decreased. *Enterococcus*, *Gemellaceae*, *Parabacteroides*, *Ruminococcaceae* and *Turicibacter* showed differences at baseline, but not in inactive disease, whereas *Bacteroides caccae*, *Barnesiellaceae* and *Dorea* showed differences in inactive disease samples, but not at baseline. A paired analysis of selected OTUs was performed of 24 Italian patients who contributed a sample at baseline and inactive disease, showing no differences between baseline and inactive disease (supplementary figure 2).

No evidence for differences between Dutch inactive disease samples and healthy controls was found (supplementary Table 7).

Persistent activity

Analysis of α -diversity showed decreased richness of persistent activity samples compared to healthy controls according to the Chaol index ($p=0.004$; Figure 1). There

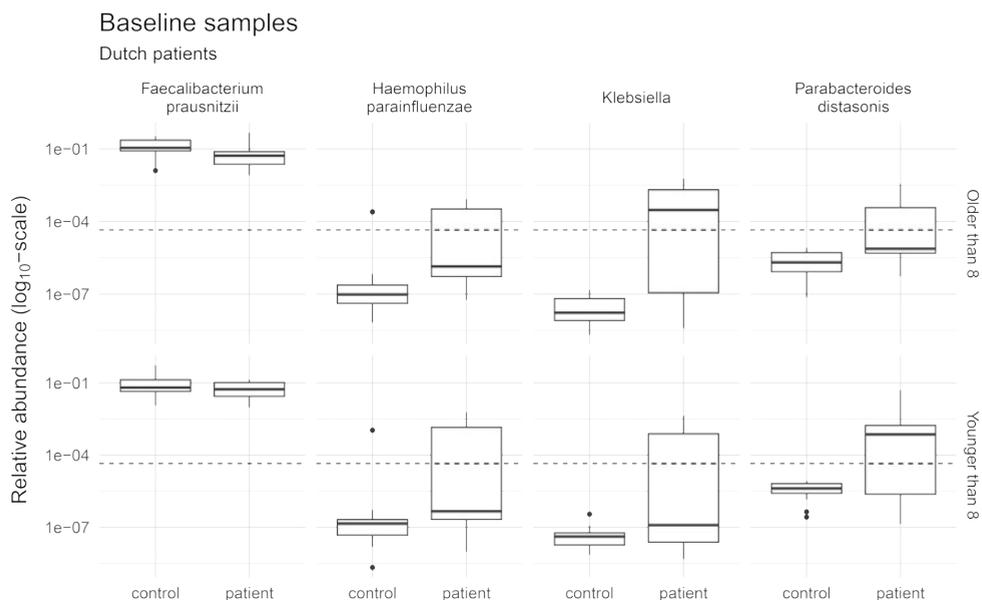


Figure 3. Boxplots of the relative abundance of the four OTUs with strongest evidence of differences between Dutch baseline samples and healthy controls. The age of the subjects confounded the relationship between patient-control status and the relative abundance. Therefore, the plots were divided into patients older than the median age of 8 years (upper row) and younger than the median age (bottom row). No evidence of differences was maintained after false discovery rate control by the Benjamini-Hochberg procedure. Shown are the relative abundances; the analysis was performed on the log-ratio of the relative abundances. The horizontal dashed line represents the limit of detection at a relative abundance of 4.5×10^{-5} . Relative abundances below this line were not observed, but were the results of imputation. Note the \log_{10} -transformation of the y-axis.

was no evidence of differences according to the Shannon index. Analysis of β -diversity showed evidence of differences between persistent activity samples and healthy controls ($p=0.001$). There was weak evidence of differences between persistent activity samples and baseline samples ($p=0.02$) and very weak evidence of differences between persistent activity and inactive disease samples ($p=0.07$; Figure 1).

Results of the logistic regressions comparing Italian persistent activity and healthy control samples are reported in supplementary Table 8. Evidence of differences was found for *Allobaculum* and *Erysipelotrichaceae* only. Figure 4 compares the results with those obtained in the baseline and inactive disease analyses.

Clinical features

Principal components regression showed strong evidence that the microbiota composition differed with age (association with the first and second principal component, $p=0.02$ and $p<0.001$ respectively). A weak association was found with the presence of uveitis at

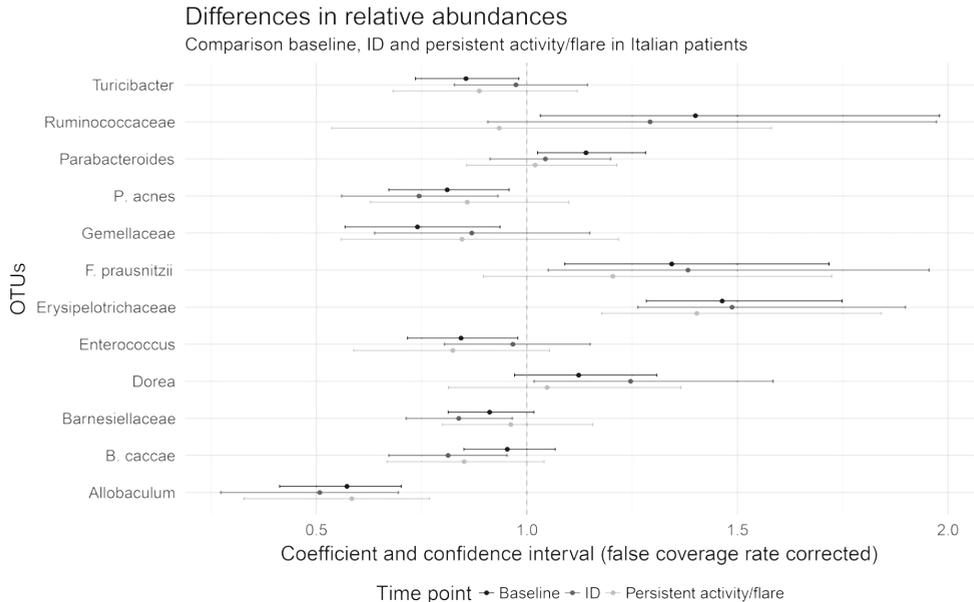


Figure 4. Comparison of differences in relative abundance between baseline samples and healthy controls (black lines), inactive disease samples and healthy controls (dark grey lines) and persistent activity samples and healthy controls (light grey lines). Shown are the regression coefficients and false coverage rate-corrected confidence intervals for the depicted OTUs, obtained in a separate multivariable logistic regression analysis for each OTU, with age and sex as covariates and patient/control status as dependent variable. Only OTUs with an adjusted p -value < 0.05 for at least one of the baseline, the inactive disease or the persistent activity analysis are shown. Regression coefficients and confidence intervals to the left of the vertical dashed line (at $x = 1$) denote decreased relative abundance in patients and those to the right denote increased relative abundance in patients. **Abbreviations:** ID, inactive disease; OTUs, operational taxonomic units.

baseline. Differently abundant OTUs were *Blautia* ($p=0.003$), *Coriobacteriaceae* ($p=0.009$), *Coprococcus* ($p=0.03$) and *Lachnospiraceae* ($p=0.04$; unadjusted p -values; no differences remained after FDR correction). However, only seven patients had uveitis at baseline. No or very weak associations were found with gender, JIA subtype, physician's global assessment of disease activity, number of active joints, ANA, NSAIDs, probiotics, prior gastrointestinal infections or gastrointestinal disorders (see Table 1 for gastrointestinal disorders in the cohort).

Sensitivity analyses

The administration of antibiotics or probiotics in the month prior to sample collection was weakly associated with the gut microbiota composition. Therefore, all abovementioned analyses were repeated after exclusion of subjects having taken either, without any substantial differences (data not shown). A further analysis was performed calculating

log-ratios using simple multiplicative replacement of zeros in the data,²² instead of robust imputation, without major differences (data not shown).

DISCUSSION

This is the first report of microbiota composition in a prospective inception cohort of JIA patients, taking into account JIA disease activity state.¹⁰⁻¹² The results of our study lent support to the hypothesis that the composition of microbiota is different in JIA patients compared to healthy controls. First, evidence was found that microbiota richness, in terms of rare OTUs, as measured by the Chaol index was reduced in JIA patients compared to healthy controls, at baseline, in inactive disease and during persistent disease activity in both Italian and Dutch patients. Measures of β -diversity were different in all subgroups of patient samples compared to healthy controls. Differences in gut microbiota composition were observed between JIA patients at baseline and healthy children, both when classifying samples using the random forest algorithm (AUC >0.99 for Italian samples and 0.71 for Dutch samples) and when analyzing the OTUs individually. Most notably *Erysipelotrichaceae* and *F. prausnitzii* increased, while *Allobaculum* decreased (Table 2; Figure 4) in Italian patient samples. Dutch samples were dissimilar to Italian samples, confirming the prior belief that the gut microbiota composition differs between these populations, potentially due to diet and other environmental factors. However, it could not be ruled out that the observed differences were due to different sample collection procedures used in the Italian and Dutch cohorts.

Few OTUs were differently abundant between controls and patients across all disease activity states. Nonetheless, when comparing baseline samples with paired inactive disease samples, no differences in relative abundance were found (supplementary Figure 2). This may indicate that the gut microbiota profile is specific to the inflammatory process underlying JIA, rather than disease activity status. Yet, it should be taken into account that JIA is very heterogeneous, clinically, and that different diseases may be underlying the clinical symptoms. Therefore, this conclusion is still preliminary. Finally, the results also indicated a strong relationship of the microbiota composition with age, as found previously.^{25,26}

The results lead towards the question whether the microbiota composition is involved in the pathogenesis of JIA. Observations compatible with a pathogenic role of microbiota are a reduced number of T helper 17 (Th17) and Forkhead box P3 (FoxP3) positive regulatory T cells in germ-free mice.^{1,2,7} Administration of stools from healthy mice to these germ-free mice led to a reconstitution of Th17 numbers,²⁷ suggesting that microbiota influence the development of the immune system. These cell populations are of critical importance in the immune homeostasis and play a key role in the pathogenesis of autoimmune diseases, among which JIA.^{2,28} Furthermore, the risk to develop autoimmune disorders appeared increased following the use of antibiotics in early childhood. This

risk was dependent on the number, dose and timing of the antibiotic prescriptions.^{29;30} Finally, fecal microbiota transplantation has led to promising, though varying results for inflammatory bowel disorders.³¹⁻³⁴ Nevertheless, more research is needed to corroborate these findings.

The OTUs *Erysipelotrichaceae*, *F. prausnitzii* and *Allobaculum* have been associated with human disease previously. *Erysipelotrichaceae* appears to be highly immunogenic^{35;36} and can potentially overgrow in the gut after treatment with broad spectrum antibiotics.³⁷ *F. prausnitzii* was identified to have anti-inflammatory properties, and a decreased abundance of this microorganism in the gut has been associated with Crohn's disease³⁸ and JIA.^{11;12} Yet, in our study it was increased in Italian patients at baseline. However, previous studies examined enthesitis-related arthritis only, whereas our cohort contained a majority of oligoarticular and polyarticular patients. The discrepancy might also be occasioned by differences in the statistical analysis. Furthermore, the balance among different subspecies of *F. prausnitzii* can result in changes in short-chain fatty acids (SCFA) production, therefore a fine-level characterization of *F. prausnitzii* strains is important to better understand their real function in the microbiome.³⁹ Such a fine-level characterization was not available in our study. *Allobaculum* is associated with obesity. Several studies reported that its abundance was lower in the gut of obese mice^{40;41} and increased in weight-reduced mice.⁴² Furthermore, *Allobaculum* was positively correlated with plasma high-density lipoprotein concentrations in hamsters⁴³ and negatively correlated with leptin concentration.⁴²

The strengths of our study are that we studied a large cohort of treatment-naïve JIA patients compared to healthy controls. Patients were followed longitudinally, allowing comparing the microbiota between various disease activity states. Finally, we used compositional data analysis, to analyze data expressed as proportions of a whole, such as microbiota relative abundances.^{20;21} Failure to do so, can lead to an increased FDR, up to 68%.¹⁴ Despite this being the largest study of gut microbiota composition in JIA patients, a limitation is the relatively small number of Dutch inactive disease and persistent disease activity samples, as well as the number of paired samples, re-emphasizing the need for replication in future cohorts.

In conclusion, our results supported the hypothesis of gut dysbiosis in JIA patients, in terms of richness and compositional deviation from healthy subjects. Since the microbiome profile appeared to be independent from disease stage, inflammation and autoimmunity indices, it seems plausible to hypothesize that the JIA microbiome affects the disease, rather than vice versa. Clearly, future investigations should be aimed at replicating these results in other populations, elucidating the potential causal role of gut microbiota in the pathogenesis of JIA and investigating the interaction amongst genetics, microbiota and environmental factors that concur in the development of JIA.

DATA AVAILABILITY

All raw 454 sequencing reads and the associated metadata are available at NCBI: Bioprojects: PRJNA379123 and PRJNA280490 (<http://www.ncbi.nlm.nih.gov/bioproject/?term=>).

FUNDING

This study has been supported by the European Commission, 7th FP, large integrated project: MD-Paedigree, Model Driven Paediatric European Digital Repository, Information Communication Technologies Programme (Contract Number 600932).

SUPPLEMENTARY MATERIAL

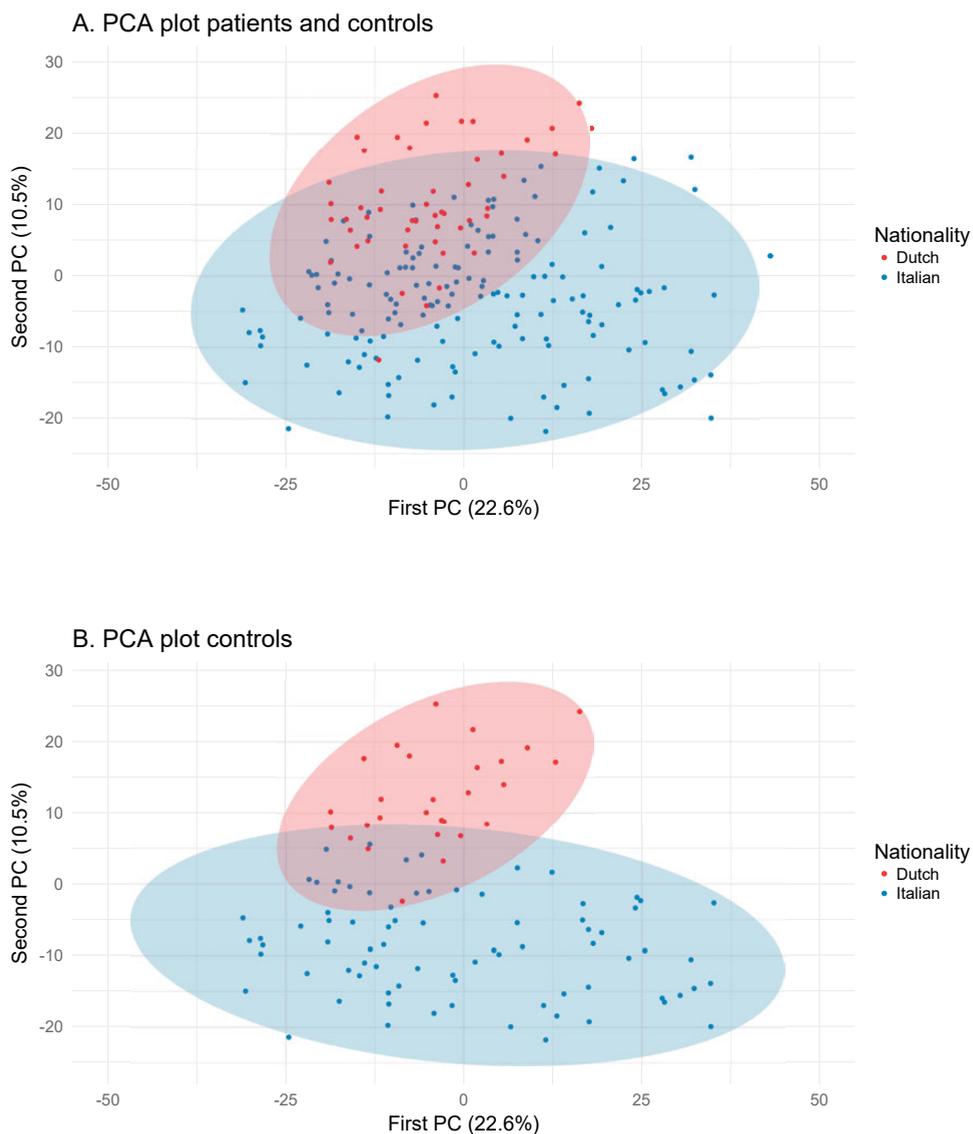
Detailed methods

Fecal sample collection

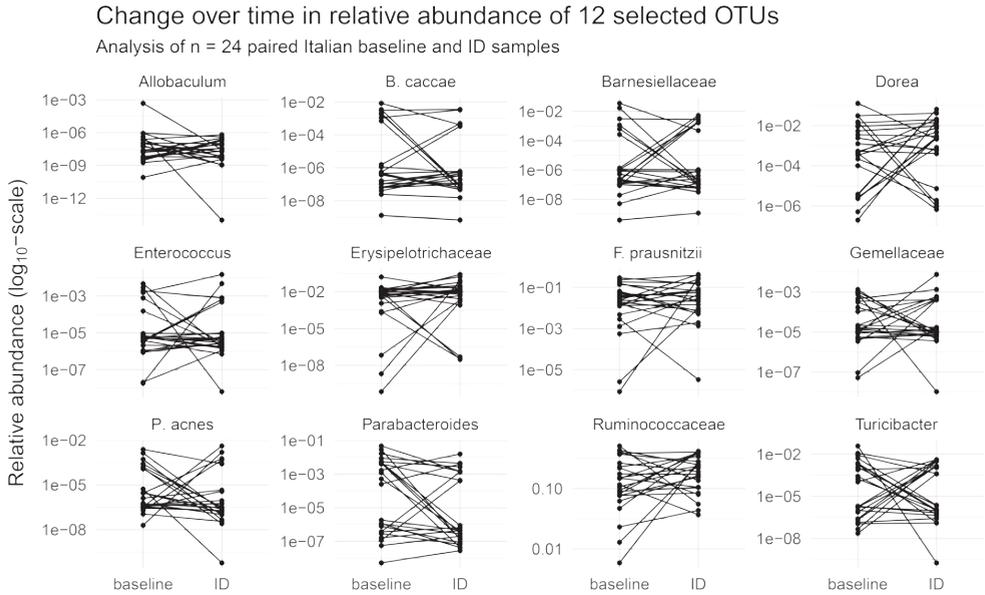
A fecal sample was collected at baseline and during follow up in case of inactive disease, persistent activity or disease flare. Samples were collected on-site in sterile containers and stored at -80 °C within 24 hours. Alternatively, samples were collected at home, shipped to the hospital and stored at -80 °C within 24 hours. Dutch samples were collected at home, stored in the domestic freezer and brought to the hospital in frozen condition at the time of the next scheduled visit. Samples were stored at -80 °C in the hospital. Samples from IGG and WKZ were shipped on dry ice to OPBG for analysis.

Genomic DNA Extraction, Pyrosequencing and Quantitative Analysis of the Microbiome Composition

Genomic DNA was isolated from the entire set of samples, using the QIAamp DNA Stool Mini Kit (Qiagen, Germany). The V1-V3 regions of the 16S ribosomal RNA (rRNA) locus were amplified for the next pyrosequencing step on a 454-Junior Genome Sequencer (Roche 454 Life Sciences, Branford, USA)¹⁸. Reads were analyzed by Quantitative Insights into Microbial Ecology (QIIME, v.1.8.0),¹⁹ grouped into operational taxonomic units (OTUs) at a sequence similarity level of 97% by PyNAST¹⁹ for sequence alignment and by UCLUST⁴⁴ for the matching of OTUs against Greengenes database (v. 13.8).⁴⁵ The α - and β -diversity and the ANOSIM tests were carried out by QIIME software, using *alpha_rarefaction.py*, *beta_diversity_through_plots.py* and *compare_categories.py* scripts.⁴⁶



Supplementary figure 1. Scatterplot of the scores of the first and second principal component of all baseline patient and healthy control samples together (A). Panel B shows samples of healthy controls. Dutch samples cluster together, therefore these subgroups were analyzed separately.



Supplementary figure 2. Relative abundance of 12 selected OTUs in 24 Italian patients who contributed a sample at baseline and in inactive disease. OTUs were selected if they showed a significantly different relative abundance between patients and healthy controls at baseline or in inactive disease. Paired differences of the log-ratio of the relative abundance were analyzed using the Wilcoxon signed-rank test. All $p > 0.05$. Note the different scale and the \log_{10} -transformation of the y-axis.

Supplementary Table 1. Clinical features of patients with inactive disease sample.

	Italian	Dutch
N	39	5
Female, <i>n</i> (%)	33 (84.6)	3 (60)
Age, median [IQR]	5.5 [3.8-9.6]	7.5 [5.9-13.0]
JIA category at onset		
Oligoarticular persistent, <i>n</i> (%)	30 (76.9)	3 (60)
Oligoarticular extended, <i>n</i> (%)	1 (2.6)	0 (0)
Polyarticular RF positive, <i>n</i> (%)	0 (0)	0 (0)
Polyarticular RF negative, <i>n</i> (%)	8 (20.5)	2 (40)
Enthesitis-related, <i>n</i> (%)	0 (0)	0 (0)
Psoriatic, <i>n</i> (%)	0 (0)	0 (0)
Undifferentiated, <i>n</i> (%)	0 (0)	0 (0)
Number of active joints, median [IQR]	0 [0, 0]	0 [0, 0]
PGA, median [IQR]	0 [0, 0]	0 [0, 0]
Parent GA, median [IQR]	0 [0, 0]	0 [0, 0.1]
JADAS-71, median [IQR]	0 [0, 0.5]	0 [0, 0.1]
ESR, median [IQR]	9 [6, 15]	6 [5, 6]
ANA, <i>n</i> (%)	28 (71.8)	1 (20)
RF, <i>n</i> (%)	0 (0)	0 (0)
NSAIDs, <i>n</i> (%)	2 (5.1)	3 (60)
Intra-articular corticosteroids, <i>n</i> (%)	8 (20.5)	3 (60)
Oral steroids, <i>n</i> (%)	2 (5.1)	1 (20)
MTX, <i>n</i> (%)	27 (69.2)	3 (60)
Anti-TNF, <i>n</i> (%)	7 (17.9)	1 (20)
Antibiotics, <i>n</i> (%)	0 (0)	0 (0)
Prebiotics, <i>n</i> (%)	3 (7.7)	2 (40)
Probiotics, <i>n</i> (%)	2 (5.1)	0 (0)
GI infections, <i>n</i> (%)	6 (15.4)	0 (0)
GI disorders, <i>n</i> (%)	0 (0)	1 (20)

Abbreviations: ANA, anti-nuclear antibodies; anti-TNF, anti-tumor necrosis factor; ESR, erythrocyte sedimentation rate; GA, global assessment of disease activity; GI, gastro-intestinal; IQR, interquartile range; JADAS, juvenile arthritis disease activity score; JIA, juvenile idiopathic arthritis; NSAID, non-steroidal anti-inflammatory drug; MTX, methotrexate; PGA, physician global assessment of disease activity; RF, rheumatoid factor.

Supplementary Table 2. Clinical features of patients with persistent activity or flare sample.

	Italian
N	24
Female, <i>n</i> (%)	17 (70.8)
Age, median [IQR]	4.7 [3.2, 7.9]
JIA category at onset	
Oligoarticular persistent, <i>n</i> (%)	18 (75.0)
Oligoarticular extended, <i>n</i> (%)	1 (4.2)
Polyarticular RF positive, <i>n</i> (%)	1 (4.2)
Polyarticular RF negative, <i>n</i> (%)	4 (16.7)
Enthesitis-related, <i>n</i> (%)	0 (0)
Psoriatic, <i>n</i> (%)	0 (0)
Undifferentiated, <i>n</i> (%)	0 (0)
Number of active joints, median [IQR]	2 [1, 3]
PGA, median [IQR]	3.0 [1.5, 4.0]
Parent GA, median [IQR]	2.0 [0.5, 3.9]
JADAS-71, median [IQR]	7.8 [5.7, 11.1]
ESR, median [IQR]	14 [8, 26]
ANA, <i>n</i> (%)	19 (79.2)
RF, <i>n</i> (%)	1 (4.2)
NSAIDs, <i>n</i> (%)	8 (33.3)
Intra-articular corticosteroids, <i>n</i> (%)	4 (16.7)
Oral steroids, <i>n</i> (%)	0 (0)
MTX, <i>n</i> (%)	12 (50.0)
Anti-TNF, <i>n</i> (%)	3 (12.5)
Antibiotics, <i>n</i> (%)	0 (0)
Prebiotics, <i>n</i> (%)	1 (4.2)
Probiotics, <i>n</i> (%)	0 (0)
GI infections, <i>n</i> (%)	1 (4.2)
GI disorders, <i>n</i> (%)	0 (0)

Abbreviations: ANA, anti-nuclear antibodies; anti-TNF, anti-tumor necrosis factor; ESR, erythrocyte sedimentation rate; GA, global assessment of disease activity; GI, gastro-intestinal; IQR, interquartile range; JADAS, juvenile arthritis disease activity score; JIA, juvenile idiopathic arthritis; NSAID, non-steroidal anti-inflammatory drug; MTX, methotrexate; PGA, physician global assessment of disease activity; RF, rheumatoid factor.

Supplementary Table 3. List of OTUs.

OTUs	Frequency (%) ^a	Mean ^b	Included ^c
Abiotrophia	16.3	0.0007	No
Acidaminococcus	7.5	0.0022	No
Acinetobacter	4.2	0.0009	No
Acinetobacter rhizosphaerae	0.7	0.0001	No
Acinetobacter venetianus	0.7	0.0006	No
Actinobacillus parahaemolyticus	2.3	0.0007	No
Actinomyces	36.3	0.0010	Yes
Actinotalea	0.3	0.0002	No
Adlercreutzia	10.8	0.0008	No
Aerococcaceae	2.6	0.0003	No
Aeromicrobium	1.0	0.0026	No
Aeromonadaceae	1.6	0.0004	No
Aggregatibacter	0.7	0.0019	No
Aggregatibacter segnis	1.0	0.0005	No
Agrobacterium	2.6	0.0024	No
Agrococcus jenensis	0.3	0.0002	No
Akkermansia muciniphila	49.3	0.0046	Yes
Alcaligenaceae	1.3	0.0014	No
Alistipes indistinctus	0.3	0.0002	No
Alistipes massiliensis	1.0	0.0005	No
Allobaculum	21.2	0.0077	Yes
Alloiococcus	0.7	0.0002	No
Alloiococcus otitis	1.0	0.0004	No
Alphaproteobacteria	0.3	0.0009	No
Alteromonadales	0.3	0.0002	No
Anaerococcus	8.5	0.0004	No
Anaerofustis	10.8	0.0006	No
Anaeroplasmataceae	0.7	0.0009	No
Anaerostipes	51.6	0.0022	Yes
Anaerotruncus	19.3	0.0007	No
Anoxybacillus kestanbolensis	0.3	0.0003	No
Arcanobacterium	0.3	0.0006	No
Arthrobacter psychrolactophilus	0.7	0.0003	No
Atopobium	15.4	0.0008	No
Aurantimonadaceae	0.3	0.0026	No
Bacillaceae	0.3	0.0004	No
Bacillus	1.0	0.0008	No
Bacillus clausii	0.3	0.0004	No
Bacteroidaceae	22.2	0.0013	Yes
Bacteroides	93.1	0.0309	Yes
Bacteroides acidifaciens	0.3	0.0003	No
Bacteroides caccae	37.9	0.0020	Yes
Bacteroides coprophilus	0.3	0.0051	No

Supplementary Table 3. (continued)

OTUs	Frequency (%) ^a	Mean ^b	Included ^c
Bacteroides eggerthii	8.8	0.0027	No
Bacteroides fragilis	53.9	0.0063	Yes
Bacteroides ovatus	55.2	0.0050	Yes
Bacteroides plebeius	4.2	0.0012	No
Bacteroides uniformis	69.9	0.0072	Yes
Barnesiellaceae	44.4	0.0037	Yes
Bifidobacteriaceae	0.3	0.0002	No
Bifidobacterium	7.5	0.0010	No
Bifidobacterium adolescentis	2.6	0.0010	No
Bifidobacterium bifidum	0.7	0.0003	No
Bifidobacterium longum	10.5	0.0013	No
Blautia	85	0.0044	Yes
Blautiaproducta	17	0.0008	No
Bradyrhizobiaceae	3.3	0.0003	No
Brochothrix	0.3	0.0006	No
Bulleidia	20.3	0.0007	Yes
Bulleidia moorei	3.3	0.0006	No
Butyricimonas	15.4	0.0012	No
Butyrivibrio	1.3	0.0002	No
Campylobacter	2.6	0.0004	No
Candidatus Rhodoluna	1.3	0.0007	No
Capnocytophaga	0.7	0.0003	No
Carnobacterium	0.7	0.0076	No
Catenibacterium	1.6	0.0021	No
Catonella	0.3	0.0003	No
Caulobacter	1.6	0.0010	No
Caulobacteraceae	3.9	0.0020	No
Cellulomonas	0.3	0.0002	No
Cellvibrio	0.3	0.0004	No
Cerasicoccaceae	0.3	0.0002	No
Cetobacterium somerae	0.3	0.0005	No
Christensenella	23.9	0.0010	Yes
Christensenellaceae	35.3	0.0019	Yes
Chryseobacterium	2.9	0.0041	No
Citrobacter	4.2	0.0009	No
Cloacibacterium	1.0	0.0002	No
Clostridia	1.6	0.0005	No
Clostridiaceae	82.4	0.0039	Yes
Clostridium	65.4	0.0023	Yes
Clostridium hiranonis	0.3	0.0003	No
Clostridium methylpentosum	0.3	0.0004	No
Clostridium perfringens	0.7	0.0006	No
Clostridium piliforme	0.3	0.0002	No

Supplementary Table 3. (continued)

OTUs	Frequency (%) ^a	Mean ^b	Included ^c
<i>Clostridium saccharogumia</i>	1.3	0.0003	No
<i>Collinsella</i>	0.7	0.0005	No
<i>Collinsella aerofaciens</i>	18.3	0.0028	No
<i>Collinsella stercoris</i>	2.0	0.0013	No
Comamonadaceae	7.2	0.0017	No
<i>Comamonas</i>	3.3	0.0033	No
<i>Coprobacillus</i>	27.1	0.0012	Yes
<i>Coprobacillus cateniformis</i>	1.0	0.0014	No
<i>Coprococcus</i>	88.2	0.0068	Yes
Coriobacteriaceae	64.7	0.0019	Yes
Corynebacteriaceae	2.6	0.0004	No
<i>Corynebacterium</i>	15.4	0.0008	No
<i>Corynebacterium durum</i>	3.3	0.0004	No
<i>Corynebacterium variabile</i>	0.7	0.0008	No
<i>Defluviibacter</i>	0.3	0.0010	No
Dehalobacteriaceae	0.3	0.0003	No
<i>Dehalobacterium</i>	15.4	0.0007	No
<i>Delftia</i>	0.7	0.0020	No
<i>Dermabacter</i>	0.3	0.0012	No
<i>Desulfitobacter</i>	0.3	0.0040	No
<i>Devosia</i>	1.3	0.0005	No
<i>Dialister</i>	72.9	0.0201	Yes
<i>Dietzia</i>	0.3	0.0019	No
Dietziaceae	0.3	0.0019	No
<i>Dorea</i>	66.3	0.0020	Yes
<i>Dorea formicigenerans</i>	0.7	0.0003	No
<i>Dyadobacter</i>	1.3	0.0003	No
<i>Dysgonomonas</i>	1.0	0.0015	No
<i>Dysgonomonas gadei</i>	0.3	0.0019	No
<i>Eggerthella lenta</i>	34	0.0015	Yes
<i>Enhydrobacter</i>	1.3	0.0002	No
Enterobacteriaceae	82	0.0140	Yes
Enterococcaceae	16	0.0020	No
<i>Enterococcus</i>	20.9	0.0023	Yes
<i>Epulopiscium</i>	1.0	0.0011	No
Erysipelotrichaceae	70.6	0.0048	Yes
<i>Ethanoligenens</i>	0.3	0.0002	No
<i>Eubacterium</i>	1.3	0.0004	No
<i>Eubacterium bifforme</i>	8.8	0.0081	No
<i>Eubacterium cylindroides</i>	0.7	0.0051	No
<i>Eubacterium dolichum</i>	39.2	0.0014	Yes
<i>Exiguobacterium</i>	0.3	0.0011	No
<i>Faecalibacterium prausnitzii</i>	93.5	0.0317	Yes

Supplementary Table 3. (continued)

OTUs	Frequency (%) ^a	Mean ^b	Included ^c
Filifactor	0.3	0.0002	No
Finegoldia	4.2	0.0003	No
Flavobacteriaceae	1.6	0.0006	No
Flavobacterium	4.9	0.0045	No
Flavobacterium succinicans	2.0	0.0136	No
Flectobacillus	0.7	0.0008	No
Fluviicola	0.3	0.0010	No
Fusobacteriales	0.3	0.0003	No
Fusobacterium	5.6	0.0005	No
Gallicola	0.7	0.0003	No
Gemella	2.6	0.0004	No
Gemellaceae	25.8	0.0007	Yes
Geobacillus	0.3	0.0006	No
Gordonia	1.0	0.0003	No
Granulicatella	36.9	0.0010	Yes
Haemophilus parainfluenzae	30.4	0.0018	Yes
Holdemania	24.8	0.0007	Yes
Hydrogenophaga	0.7	0.0001	No
Hydrogenophilus	2.3	0.0005	No
Hymenobacter	1.3	0.0002	No
Hyphomicrobium	3.9	0.0004	No
Janibacter	1.6	0.0005	No
Janthinobacterium	2.3	0.0022	No
Janthinobacterium lividum	1.3	0.0012	No
Jonquetella anthropi	0.3	0.0002	No
Kineococcus	0.3	0.0007	No
Klebsiella	23.2	0.0018	Yes
Lachnospira	57.2	0.0019	Yes
Lachnospiraceae	97.7	0.0457	Yes
Lactobacillaceae	3.9	0.0011	No
Lactobacillus	16.3	0.0012	No
Lactobacillus brevis	0.3	0.0006	No
Lactobacillus reuteri	0.3	0.0007	No
Lactobacillus zeae	3.9	0.0010	No
Lactococcus	10.1	0.0010	No
Lactococcus garvieae	1.0	0.0052	No
Legionellaceae	0.3	0.0004	No
Leptotrichia	1.0	0.0007	No
Leucobacter	0.3	0.0002	No
Leuconostoc	1.6	0.0004	No
Leuconostocaceae	1.3	0.0006	No
Listeriaceae	0.3	0.0016	No
Macrococcus caseolyticus	0.3	0.0003	No

Supplementary Table 3. (continued)

OTUs	Frequency (%) ^a	Mean ^b	Included ^c
Megamonas	4.9	0.0011	No
Megasphaera	4.6	0.0018	No
Mesorhizobium	2.0	0.0004	No
Methylibium	0.7	0.0006	No
Methylobacteriaceae	2.0	0.0038	No
Methylobacterium	0.3	0.0002	No
Methylophilaceae	2.3	0.0048	No
Microbacteriaceae	3.3	0.0007	No
Microbacterium	0.7	0.0003	No
Micrococcaceae	2.0	0.0005	No
Mobiluncus	1.0	0.0005	No
Mogibacteriaceae	63.4	0.0024	Yes
Mogibacterium	7.5	0.0007	No
Moraxella	1.6	0.0007	No
Moraxellaceae	7.8	0.0012	No
Morganella	1.0	0.0008	No
Morganella morganii	1.3	0.0008	No
Mycobacteriaceae	0.3	0.0003	No
Mycobacterium	0.7	0.0006	No
Mycobacterium vaccae	0.3	0.0003	No
Nannocystis	0.3	0.0004	No
Neisseria	0.3	0.0005	No
Neisseria subflava	1.6	0.0004	No
Neisseriaceae	0.7	0.0011	No
Nitrospira	0.3	0.0007	No
Nocardioideaceae	1.0	0.0004	No
Novosphingobium	0.3	0.0002	No
Odoribacter	32	0.0010	Yes
Odoribacteraceae	0.3	0.0012	No
Oligella	0.3	0.0003	No
Oribacterium	2.0	0.0002	No
Oscillospira	91.5	0.0066	Yes
Oxalobacter formigenes	7.2	0.0012	No
Paenibacillaceae	0.3	0.0005	No
Parabacteroides	52	0.0043	Yes
Parabacteroides distasonis	28.4	0.0016	Yes
Parabacteroides gordonii	0.7	0.0008	No
Paracoccus	2.3	0.0005	No
Paraprevotella	12.1	0.0024	No
Paraprevotellaceae	0.7	0.0009	No
Parvimonas	13.7	0.0006	No
Pasteurellaceae	3.6	0.0010	No
Pediococcus	1.6	0.0008	No

Supplementary Table 3. (continued)

OTUs	Frequency (%) ^a	Mean ^b	Included ^c
Pedobacter	2.0	0.0328	No
Pelomonas	0.7	0.0002	No
Peptococcaceae	1.0	0.0005	No
Peptococcus	1.6	0.0006	No
Peptoniphilus	8.2	0.0007	No
Peptostreptococcaceae	64.7	0.0038	Yes
Peptostreptococcus	2.9	0.0004	No
Phascolarctobacterium	30.4	0.0038	Yes
Phenylobacterium	1.0	0.0007	No
Phycoccus	0.3	0.0002	No
Phyllobacteriaceae	9.8	0.0005	No
Pirellulaceae	0.3	0.0001	No
Planococcaceae	4.9	0.0003	No
Polaromonas	0.3	0.0004	No
Porphyromonadaceae	3.9	0.0006	No
Porphyromonas	12.4	0.0008	No
Prevotella	30.1	0.0012	Yes
Prevotella	3.6	0.0031	No
Prevotella copri	23.5	0.0046	Yes
Prevotella intermedia	0.3	0.0001	No
Prevotella melaninogenica	1.3	0.0004	No
Prevotella stercorea	0.7	0.0023	No
Prevotellaceae	1.0	0.0003	No
Propionibacteriaceae	1.0	0.0008	No
Propionibacterium acnes	20.9	0.0007	Yes
Propionibacterium granulosum	0.3	0.0004	No
Propionivibrio	1.0	0.0003	No
Proteus	0.7	0.0012	No
Pseudomonadaceae	7.5	0.0007	No
Pseudomonadales	1.0	0.0002	No
Pseudomonas	22.2	0.0015	Yes
Pseudoramibacter	10.1	0.0020	No
Pseudoxanthomonas mexicana	0.3	0.0004	No
Pyramidobacter piscolens	2.6	0.0010	No
Ralstonia	1.3	0.0005	No
Renibacterium	0.3	0.0004	No
Rheinheimera	2.3	0.0007	No
Rhizobiaceae	0.3	0.0019	No
Rhodobacter	0.3	0.0002	No
Rhodobacteraceae	3.9	0.0007	No
Rhodococcus	1.6	0.0016	No
Rhodococcus fascians	2.0	0.0032	No
Rhodocyclaceae	2.3	0.0014	No

Supplementary Table 3. (continued)

OTUs	Frequency (%) ^a	Mean ^b	Included ^c
Rhodoferax	0.3	0.0003	No
Rhodoplanes	0.3	0.0003	No
Rhodospirillaceae	4.2	0.0007	No
Rikenella	0.7	0.0024	No
Rikenellaceae	76.5	0.0088	Yes
Roseburia	64.7	0.0023	Yes
Roseomonas mucosa	0.3	0.0004	No
Rothia aeria	0.7	0.0003	No
Rothia mucilaginosa	0.7	0.0005	No
Rubelli microbium	0.3	0.0008	No
Rubrobacter	2.3	0.0005	No
Ruminococcaceae	99	0.1234	Yes
Ruminococcus	94.8	0.0122	Yes
Ruminococcus bromii	62.7	0.0087	Yes
Ruminococcus flavefaciens	1.0	0.0008	No
Ruminococcus gnavus	49.3	0.0020	Yes
Ruminococcus torques	3.9	0.0005	No
Salinispora	1.3	0.0002	No
Schwartzia	0.3	0.0005	No
Sediminibacterium	13.7	0.0005	No
Selenomonas	0.7	0.0003	No
Selenomonas noxia	0.3	0.0004	No
Serratia	0.7	0.0016	No
Shewanella	0.3	0.0001	No
Simplicispira	0.3	0.0003	No
Sinobacteraceae	0.7	0.0004	No
Slackia	5.6	0.0010	No
Solibacillus	0.7	0.0003	No
Sphingobacteriaceae	0.3	0.0023	No
Sphingobium	1.0	0.0020	No
Sphingomonadaceae	5.9	0.0013	No
Sphingomonas	11.4	0.0006	No
Sphingomonas wittichii	1.3	0.0009	No
Staphylococcaceae	6.5	0.0004	No
Staphylococcus	2.9	0.0005	No
Staphylococcus aureus	5.9	0.0005	No
Staphylococcus epidermidis	0.3	0.0008	No
Streptococcaceae	1.6	0.0008	No
Streptococcus	84.6	0.0068	Yes
Streptococcus anginosus	0.3	0.0003	No
Streptococcus luteciae	0.3	0.0003	No
Succiniclasticum	2.3	0.0017	No
Succinivibrio	0.3	0.0567	No

Supplementary Table 3. (continued)

OTUs	Frequency (%) ^a	Mean ^b	Included ^c
Sutterella	46.4	0.0020	Yes
Synergistaceae	1.0	0.0012	No
Tannerella	0.3	0.0002	No
Tatlockia	0.3	0.0018	No
Tissierellaceae	9.8	0.0010	No
Trabulsiella	0.3	0.0004	No
Turicibacter	46.1	0.0012	Yes
Varibaculum	2.9	0.0007	No
Variovorax paradoxus	7.5	0.0008	No
Veillonella	7.5	0.0046	No
Veillonella dispar	54.9	0.0022	Yes
Veillonella parvula	1.3	0.0003	No
Veillonellaceae	9.2	0.0007	No
Victivallaceae	3.3	0.0008	No
Weeksellaceae	1.0	0.0002	No
Xanthobacteraceae	0.3	0.0003	No
Xanthomonadaceae	2.3	0.0008	No
Yonghaparkia	0.3	0.0008	No
Others	99	0.0301	Yes

^a The percentage of samples in which the OTU was present.

^b The geometric mean of the samples in which the OTU was present.

^c OTUs were included in the analysis if they were present in at least 20% of the samples.

Abbreviations: OTU, operational taxonomic unit.

Supplementary Table 4. Differences in relative abundance between Italian baseline samples and healthy controls.

OTUs	Relative abundance in patients	Crude p-value	Adjusted p-value ^a
Erysipelotrichaceae	Increased	<0.001	<0.001
Allobaculum	Decreased	<0.001	<0.001
Faecalibacterium prausnitzii	Increased	<0.001	0.009
Gemellaceae	Decreased	0.001	0.016
Propionibacterium acnes	Decreased	0.002	0.016
Parabacteroides	Increased	0.002	0.016
Enterococcus	Decreased	0.003	0.025
Turicibacter	Decreased	0.004	0.025
Ruminococcaceae	Increased	0.005	0.033
Phascolarctobacterium	Increased	0.012	0.07
Blautia	Decreased	0.017	0.09
Barnesiellaceae	Decreased	0.026	0.12
Dorea	Increased	0.037	0.16
Christensenella	Increased	0.06	0.26
Granulicatella	Decreased	0.07	0.26
Clostridium	Decreased	0.08	0.28
Clostridiaceae	Increased	0.11	0.34
Bacteroides uniformis	Increased	0.11	0.34
Bacteroides ovatus	Increased	0.11	0.34
Bacteroides	Increased	0.19	0.53
Prevotella copri	Increased	0.19	0.53
Oscillospira	Increased	0.20	0.53
Coriobacteriaceae	Decreased	0.25	0.61
Bacteroides caccae	Decreased	0.26	0.61
Eubacterium dolichum	Decreased	0.27	0.61
Eggerthella lenta	Decreased	0.28	0.61
Klebsiella	Increased	0.29	0.62
Lachnospira	Decreased	0.33	0.67
Bacteroides fragilis	Increased	0.39	0.73
Holdemania	Increased	0.39	0.73
Christensenellaceae	Increased	0.40	0.73
Sutterella	Increased	0.41	0.73
Veillonella dispar	Decreased	0.44	0.76
Pseudomonas	Decreased	0.46	0.77
Ruminococcus gnavus	Increased	0.48	0.78
Ruminococcus bromii	Increased	0.52	0.78
Odoribacter	Increased	0.54	0.78
Bulleidia	Increased	0.54	0.78
Coprococcus	Decreased	0.55	0.78
Coprobacillus	Decreased	0.56	0.78
Actinomyces	Decreased	0.57	0.78
Ruminococcus	Decreased	0.58	0.78

Supplementary Table 4. (continued)

OTUs	Relative abundance in patients	Crude p-value	Adjusted p-value ^a
Peptostreptococcaceae	Decreased	0.62	0.82
Bacteroidaceae	Increased	0.66	0.84
Haemophilus parainfluenzae	Decreased	0.67	0.84
Parabacteroides distasonis	Increased	0.69	0.86
Others	Decreased	0.75	0.90
Streptococcus	Increased	0.76	0.90
Rikenellaceae	Increased	0.80	0.93
Lachnospiraceae	Increased	0.85	0.95
Akkermansia muciniphila	Decreased	0.87	0.95
Dialister	Increased	0.87	0.95
Prevotella	Increased	0.91	0.96
Mogibacteriaceae	Decreased	0.91	0.96
Roseburia	Increased	0.93	0.96
Enterobacteriaceae	Decreased	0.97	0.98
Anaerostipes	Increased	0.98	0.98

^a False discovery rate control using the Benjamini-Hochberg procedure.

Abbreviations: OTU, operational taxonomic unit.

Supplementary Table 5. Differences in relative abundance between Dutch baseline samples and healthy controls.

OTUs	Relative abundance in patients	Crude p-value	Adjusted p-value ^a
Parabacteroides distasonis	Increased	0.003	0.12
Klebsiella	Increased	0.004	0.12
Faecalibacterium prausnitzii	Decreased	0.010	0.16
Haemophilus parainfluenzae	Increased	0.011	0.16
Actinomyces	Decreased	0.015	0.17
Veillonella dispar	Increased	0.018	0.17
Mogibacteriaceae	Decreased	0.029	0.17
Others	Decreased	0.029	0.17
Eggerthella lenta	Decreased	0.030	0.17
Ruminococcus	Decreased	0.031	0.17
Rikenellaceae	Increased	0.032	0.17
Bacteroidaceae	Increased	0.040	0.18
Ruminococcaceae	Decreased	0.046	0.18
Enterobacteriaceae	Increased	0.047	0.18
Bacteroides caccae	Increased	0.048	0.18
Allobaculum	Decreased	0.07	0.25
Christensenella	Decreased	0.07	0.25
Bacteroides	Increased	0.09	0.26
Blautia	Decreased	0.09	0.26
Dialister	Decreased	0.09	0.26
Coprobacillus	Decreased	0.11	0.29
Bacteroides uniformis	Increased	0.13	0.32
Clostridium	Decreased	0.15	0.35
Pseudomonas	Increased	0.16	0.35
Eubacterium dolichum	Decreased	0.16	0.35
Bulleidia	Decreased	0.16	0.35
Prevotella	Decreased	0.17	0.36
Lachnospira	Decreased	0.19	0.36
Granulicatella	Decreased	0.19	0.36
Enterococcus	Decreased	0.20	0.36
Coriobacteriaceae	Decreased	0.20	0.36
Sutterella	Increased	0.21	0.36
Barnesiellaceae	Increased	0.21	0.36
Lachnospiraceae	Decreased	0.23	0.38
Prevotella copri	Decreased	0.25	0.42
Bacteroides fragilis	Increased	0.33	0.52
Dorea	Increased	0.34	0.52
Propionibacterium acnes	Decreased	0.35	0.52
Parabacteroides	Increased	0.39	0.57
Oscillospira	Decreased	0.40	0.57
Roseburia	Increased	0.41	0.57
Turicibacter	Decreased	0.46	0.62

Supplementary Table 5. (continued)

OTUs	Relative abundance in patients	Crude p-value	Adjusted p-value ^a
Christensenellaceae	Decreased	0.51	0.68
Anaerostipes	Increased	0.53	0.68
Odoribacter	Decreased	0.53	0.68
Erysipelotrichaceae	Decreased	0.59	0.71
Akkermansia muciniphila	Decreased	0.60	0.71
Streptococcus	Decreased	0.61	0.71
Peptostreptococcaceae	Decreased	0.62	0.71
Ruminococcus bromii	Decreased	0.64	0.71
Clostridiaceae	Decreased	0.66	0.71
Phascolarctobacterium	Increased	0.66	0.71
Gemellaceae	Increased	0.66	0.71
Coprococcus	Increased	0.68	0.72
Ruminococcus gnavus	Decreased	0.69	0.72
Holdemania	Decreased	0.80	0.82
Bacteroides ovatus	Increased	0.98	0.98

^a False discovery rate control using the Benjamini-Hochberg procedure.

Abbreviations: OTU, operational taxonomic unit.

Supplementary Table 6. Differences in relative abundance between Italian inactive disease samples and healthy controls.

OTUs	Relative abundance in patients	Crude p-value	Adjusted p-value ^a
Erysipelotrichaceae	Increased	<0.001	<0.001
Allobaculum	Decreased	<0.001	<0.001
Bacteroides caccae	Decreased	<0.001	0.013
Propionibacterium acnes	Decreased	0.001	0.013
Barnesiellaceae	Decreased	0.001	0.013
Faecalibacterium prausnitzii	Increased	0.003	0.033
Dorea	Increased	0.005	0.044
Clostridiaceae	Increased	0.007	0.050
Christensenella	Increased	0.013	0.09
Bacteroidaceae	Decreased	0.020	0.11
Pseudomonas	Decreased	0.027	0.14
Phascolarctobacterium	Increased	0.047	0.21
Coprococcus	Increased	0.049	0.21
Bacteroides fragilis	Increased	0.052	0.21
Coriobacteriaceae	Decreased	0.06	0.22
Streptococcus	Increased	0.06	0.22
Ruminococcaceae	Increased	0.07	0.22
Lachnospiraceae	Increased	0.09	0.28
Haemophilus parainfluenzae	Decreased	0.10	0.29
Roseburia	Increased	0.12	0.35
Lachnospira	Decreased	0.14	0.38
Peptostreptococcaceae	Increased	0.15	0.39
Gemellaceae	Decreased	0.18	0.43
Bacteroides	Increased	0.18	0.43
Bulleidia	Increased	0.20	0.46
Clostridium	Decreased	0.21	0.46
Anaerostipes	Increased	0.22	0.46
Bacteroides uniformis	Increased	0.22	0.46
Mogibacteriaceae	Increased	0.24	0.47
Oscillospira	Increased	0.33	0.62
Parabacteroides	Increased	0.38	0.70
Dialister	Decreased	0.40	0.72
Eggerthella lenta	Decreased	0.42	0.72
Klebsiella	Decreased	0.45	0.76
Blautia	Decreased	0.47	0.76
Odoribacter	Decreased	0.49	0.78
Rikenellaceae	Decreased	0.52	0.79
Others	Decreased	0.53	0.79
Ruminococcus gnavus	Increased	0.57	0.81
Akkermansia muciniphila	Decreased	0.57	0.81
Enterococcus	Decreased	0.59	0.82
Bacteroides ovatus	Increased	0.60	0.82

Supplementary Table 6. (continued)

OTUs	Relative abundance in patients	Crude p-value	Adjusted p-value ^a
<i>Veillonella dispar</i>	Decreased	0.62	0.82
<i>Turicibacter</i>	Decreased	0.65	0.83
<i>Prevotella copri</i>	Increased	0.66	0.83
<i>Prevotella</i>	Increased	0.68	0.83
<i>Eubacterium dolichum</i>	Increased	0.70	0.83
<i>Granulicatella</i>	Increased	0.70	0.83
Enterobacteriaceae	Decreased	0.75	0.87
<i>Ruminococcus</i>	Decreased	0.79	0.88
<i>Coprobacillus</i>	Increased	0.79	0.88
<i>Parabacteroides distasonis</i>	Decreased	0.80	0.88
<i>Ruminococcus bromii</i>	Increased	0.84	0.89
Christensenellaceae	Decreased	0.84	0.89
<i>Sutterella</i>	Decreased	0.88	0.91
<i>Actinomyces</i>	Increased	0.95	0.96
<i>Holdemania</i>	Increased	0.96	0.96

^a False discovery rate control using the Benjamini-Hochberg procedure.

Abbreviations: OTU, operational taxonomic unit.

Supplementary Table 7. Differences in relative abundance between Dutch inactive disease samples and healthy controls.

OTUs	Relative abundance in patients	Crude p-value	Adjusted p-value ^a
Veillonella dispar	Increased	0.07	0.85
Parabacteroides distasonis	Increased	0.09	0.85
Bacteroidaceae	Increased	0.09	0.85
Coprobacillus	Decreased	0.10	0.85
Ruminococcus bromii	Increased	0.10	0.85
Bacteroides uniformis	Decreased	0.14	0.85
Haemophilus parainfluenzae	Increased	0.15	0.85
Dorea	Decreased	0.15	0.85
Actinomyces	Decreased	0.15	0.85
Klebsiella	Increased	0.16	0.85
Eggerthella lenta	Decreased	0.18	0.85
Dialister	Increased	0.18	0.85
Oscillospira	Decreased	0.21	0.87
Bacteroides ovatus	Decreased	0.21	0.87
Bacteroides caccae	Decreased	0.25	0.91
Sutterella	Increased	0.26	0.91
Coprococcus	Increased	0.27	0.91
Anaerostipes	Increased	0.29	0.91
Coriobacteriaceae	Decreased	0.30	0.91
Allobaculum	Decreased	0.34	0.93
Propionibacterium acnes	Decreased	0.39	0.93
Christensenella	Decreased	0.40	0.93
Others	Decreased	0.42	0.93
Prevotella copri	Decreased	0.45	0.93
Mogibacteriaceae	Decreased	0.45	0.93
Parabacteroides	Increased	0.45	0.93
Bulleidia	Decreased	0.47	0.93
Ruminococcus	Increased	0.51	0.93
Akkermansia muciniphila	Decreased	0.53	0.93
Rikenellaceae	Increased	0.53	0.93
Christensenellaceae	Increased	0.53	0.93
Odoribacter	Increased	0.55	0.93
Faecalibacterium prausnitzii	Decreased	0.56	0.93
Enterococcus	Decreased	0.57	0.93
Peptostreptococcaceae	Decreased	0.58	0.93
Gemellaceae	Decreased	0.59	0.93
Eubacterium dolichum	Increased	0.65	0.96
Phascolarctobacterium	Decreased	0.65	0.96
Turicibacter	Increased	0.68	0.96
Clostridiaceae	Decreased	0.69	0.96
Streptococcus	Decreased	0.69	0.96
Pseudomonas	Decreased	0.73	0.96

Supplementary Table 7. (continued)

OTUs	Relative abundance in patients	Crude p-value	Adjusted p-value ^a
Clostridium	Decreased	0.76	0.96
Granulicatella	Decreased	0.78	0.96
Lachnospiraceae	Increased	0.78	0.96
Holdemania	Increased	0.80	0.96
Erysipelotrichaceae	Decreased	0.81	0.96
Prevotella	Increased	0.81	0.96
Bacteroides	Increased	0.86	0.97
Bacteroides fragilis	Increased	0.87	0.97
Ruminococcaceae	Increased	0.90	0.97
Blautia	Increased	0.91	0.97
Lachnospira	Increased	0.91	0.97
Ruminococcus gnavus	Decreased	0.91	0.97
Barnesiellaceae	Increased	0.98	0.99
Roseburia	Decreased	0.98	0.99
Enterobacteriaceae	Decreased	0.99	0.99

^a False discovery rate control using the Benjamini-Hochberg procedure.

Abbreviations: OTU, operational taxonomic unit.

Supplementary Table 8. Differences in relative abundance between Italian persistent activity samples and healthy controls.

OTUs	Relative abundance in patients	Crude p-value	Adjusted p-value ^a
Erysipelotrichaceae	Increased	<0.001	<0.001
Allobaculum	Decreased	<0.001	<0.001
Blautia	Decreased	0.003	0.06
Klebsiella	Increased	0.006	0.09
Phascolarctobacterium	Increased	0.008	0.10
Coriobacteriaceae	Decreased	0.016	0.16
Bacteroides caccae	Decreased	0.019	0.16
Enterococcus	Decreased	0.031	0.22
Parabacteroides distasonis	Increased	0.047	0.30
Faecalibacterium prausnitzii	Increased	0.07	0.34
Prevotella copri	Increased	0.07	0.34
Ruminococcus	Decreased	0.08	0.34
Propionibacterium acnes	Decreased	0.08	0.34
Pseudomonas	Increased	0.09	0.38
Rikenellaceae	Decreased	0.11	0.41
Turicibacter	Decreased	0.12	0.43
Others	Decreased	0.13	0.44
Anaerostipes	Increased	0.16	0.50
Gemellaceae	Decreased	0.17	0.50
Bacteroidaceae	Increased	0.18	0.52
Veillonella dispar	Increased	0.19	0.52
Eggerthella lenta	Decreased	0.21	0.54
Bacteroides	Increased	0.23	0.56
Christensenellaceae	Decreased	0.24	0.57
Bacteroides fragilis	Increased	0.28	0.62
Haemophilus parainfluenzae	Increased	0.28	0.62
Holdemania	Decreased	0.31	0.65
Lachnospira	Decreased	0.33	0.67
Enterobacteriaceae	Increased	0.34	0.67
Clostridium	Decreased	0.37	0.69
Oscillospira	Increased	0.39	0.69
Coprobacillus	Decreased	0.40	0.69
Coprococcus	Decreased	0.40	0.69
Actinomyces	Decreased	0.47	0.76
Lachnospiraceae	Decreased	0.48	0.76
Bacteroides uniformis	Increased	0.48	0.76
Barnesiellaceae	Decreased	0.50	0.76
Ruminococcus gnavus	Increased	0.53	0.78
Dialister	Decreased	0.54	0.78
Granulicatella	Decreased	0.55	0.78
Dorea	Increased	0.56	0.78
Bulleidia	Decreased	0.57	0.78

Supplementary Table 8. (continued)

OTUs	Relative abundance in patients	Crude p-value	Adjusted p-value ^a
<i>Bacteroides ovatus</i>	Decreased	0.59	0.78
<i>Roseburia</i>	Decreased	0.64	0.82
Clostridiaceae	Increased	0.67	0.83
Ruminococcaceae	Decreased	0.69	0.83
<i>Streptococcus</i>	Increased	0.70	0.83
<i>Odoribacter</i>	Increased	0.70	0.83
<i>Parabacteroides</i>	Increased	0.72	0.84
Peptostreptococcaceae	Increased	0.79	0.89
Mogibacteriaceae	Decreased	0.79	0.89
<i>Ruminococcus bromii</i>	Decreased	0.86	0.94
<i>Akkermansia muciniphila</i>	Decreased	0.91	0.97
<i>Eubacterium dolichum</i>	Increased	0.92	0.97
<i>Christensenella</i>	Decreased	0.95	0.98
<i>Sutterella</i>	Decreased	0.97	0.98
<i>Prevotella</i>	Increased	0.98	0.98

^a False discovery rate control using the Benjamini-Hochberg procedure.

Abbreviations: OTU, operational taxonomic unit.

REFERENCE LIST

1. Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. *Nat Rev Genet* 2012; 13(4):260-270.
2. Verwoerd A, Ter Haar NM, de RS, Vastert SJ, Bogaert D. The human microbiome and juvenile idiopathic arthritis. *Pediatr Rheumatol Online J* 2016; 14(1):55.
3. Zhang X, Zhang D, Jia H, Feng Q, Wang D, Liang D et al. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat Med* 2015; 21(8):895-905.
4. Giongo A, Gano KA, Crabb DB, Mukherjee N, Novelo LL, Casella G et al. Toward defining the autoimmune microbiome for type 1 diabetes. *ISME J* 2011; 5(1):82-91.
5. Missaghi B, Barkema HW, Madsen KL, Ghosh S. Perturbation of the human microbiome as a contributor to inflammatory bowel disease. *Pathogens* 2014; 3(3):510-527.
6. Scher JU, Ubeda C, Artacho A, Attur M, Isaac S, Reddy SM et al. Decreased bacterial diversity characterizes the altered gut microbiota in patients with psoriatic arthritis, resembling dysbiosis in inflammatory bowel disease. *Arthritis Rheumatol* 2015; 67(1):128-139.
7. Kamada N, Seo SU, Chen GY, Nunez G. Role of the gut microbiota in immunity and inflammatory disease. *Nat Rev Immunol* 2013; 13(5):321-335.
8. Wu HJ, Ivanov II, Darce J, Hattori K, Shima T, Umesaki Y et al. Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells. *Immunity* 2010; 32(6):815-827.
9. Hissink Muller PC, Westedt PM, Budding AE, Allaart CF, Brinkman DM, Kuijpers TW et al. PReS-FINAL-2160: Intestinal microbiome in polyarticular juvenile idiopathic arthritis: a pilot study. *Pediatr Rheumatol Online J* 2013; 11(Suppl 2):P172.
10. Tejesvi MV, Arvonen M, Kangas SM, Keskitalo PL, Pirttila AM, Karttunen TJ et al. Faecal microbiome in new-onset juvenile idiopathic arthritis. *Eur J Clin Microbiol Infect Dis* 2016; 35(3):363-370.
11. Di Paola M, Cavalieri D, Albanese D, Sordo M, Pindo M, Donati C et al. Alteration of Fecal Microbiota Profiles in Juvenile Idiopathic Arthritis. Associations with HLA-B27 Allele and Disease Status. *Front Microbiol* 2016; 7:1703.
12. Stoll ML, Kumar R, Morrow CD, Lefkowitz EJ, Cui X, Genin A et al. Altered microbiota associated with abnormal humoral immune responses to commensal organisms in enthesitis-related arthritis. *Arthritis Res Ther* 2014; 16(6):486.
13. Ravelli A, Martini A. Juvenile idiopathic arthritis. *Lancet* 2007; 369(9563):767-778.
14. Mandal S, Van TW, White RA, Eggesbo M, Knight R, Peddada SD. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb Ecol Health Dis* 2015; 26:27663.
15. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004; 31(2):390-392.
16. Consolaro A, Ruperto N, Bazso A, Pistorio A, Magni-Manzoni S, Filocamo G et al. Development and validation of a composite disease activity score for juvenile idiopathic arthritis. *Arthritis Rheum* 2009; 61(5):658-666.
17. Wallace CA, Giannini EH, Huang B, Irtter L, Ruperto N. American College of Rheumatology provisional criteria for defining clinical inactive disease in select categories of juvenile idiopathic arthritis. *Arthritis Care Res (Hoboken)* 2011; 63(7):929-936.
18. De Filippis F, La SA, Villani F, Ercolini D. Exploring the sources of bacterial spoilers in beefsteaks by culture-independent high-throughput sequencing. *PLoS One* 2013; 8(7):e70222.
19. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010; 7(5):335-336.

20. Aitchison J. The statistical analysis of compositional data. London: Chapman & Hall; 1986.
21. Gloor GB, Wu JR, Pawlowsky-Glahn V, Egozcue JJ. It's all relative: analyzing microbiome data as compositions. *Ann Epidemiol* 2016; 26(5):322-329.
22. Martín-Fernández JA, Barceló-Vidal C, Pawlowsky-Glahn V. Dealing with zeros and missing values in compositional data sets using nonparametric imputation. *Math Geol* 2003; 35(3):253-278.
23. Palarea-Albaladejo J, Martín-Fernández JA. A modified EM algorithm for replacing rounded zeros in compositional data sets. *Comput Geosci* 2008; 34:902-917.
24. Consolaro A, Ruperto N, Bracciolini G, Frisina A, Gallo MC, Pistorio A et al. Defining criteria for high disease activity in juvenile idiopathic arthritis based on the juvenile arthritis disease activity score. *Ann Rheum Dis* 2014; 73(7):1380-1383.
25. Odamaki T, Kato K, Sugahara H, Hashikura N, Takahashi S, Xiao JZ et al. Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. *BMC Microbiol* 2016; 16:90.
26. Rodriguez JM, Murphy K, Stanton C, Ross RP, Kober OI, Juge N et al. The composition of the gut microbiota throughout life, with an emphasis on early life. *Microb Ecol Health Dis* 2015; 26:26050.
27. Ivanov II, Frutos RL, Manel N, Yoshinaga K, Rifkin DB, Sartor RB et al. Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine. *Cell Host Microbe* 2008; 4(4):337-349.
28. Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. *Lancet* 2011; 377(9783):2138-2149.
29. Arvonen M, Virta LJ, Pokka T, Kroger L, Vahasalo P. Repeated exposure to antibiotics in infancy: a predisposing factor for juvenile idiopathic arthritis or a sign of this group's greater susceptibility to infections? *J Rheumatol* 2015; 42(3):521-526.
30. Horton DB, Scott FI, Haynes K, Putt ME, Rose CD, Lewis JD et al. Antibiotic Exposure and Juvenile Idiopathic Arthritis: A Case-Control Study. *Pediatrics* 2015; 136(2):e333-e343.
31. Mattner J, Schmidt F, Siegmund B. Faecal microbiota transplantation-A clinical view. *Int J Med Microbiol* 2016; 306(5):310-315.
32. Moayyedi P, Surette MG, Kim PT, Libertucci J, Wolfe M, Onischi C et al. Fecal Microbiota Transplantation Induces Remission in Patients With Active Ulcerative Colitis in a Randomized Controlled Trial. *Gastroenterology* 2015; 149(1):102-109.
33. Rossen NG, Fuentes S, van der Spek MJ, Tijssen JG, Hartman JH, Dufflou A et al. Findings From a Randomized Controlled Trial of Fecal Transplantation for Patients With Ulcerative Colitis. *Gastroenterology* 2015; 149(1):110-118.
34. Vermeire S, Joossens M, Verbeke K, Wang J, Machiels K, Sabino J et al. Donor Species Richness Determines Faecal Microbiota Transplantation Success in Inflammatory Bowel Disease. *J Crohns Colitis* 2016; 10(4):387-394.
35. Dinh DM, Volpe GE, Duffalo C, Bhalchandra S, Tai AK, Kane AV et al. Intestinal microbiota, microbial translocation, and systemic inflammation in chronic HIV infection. *J Infect Dis* 2015; 211(1):19-27.
36. Palm NW, de Zoete MR, Cullen TW, Barry NA, Stefanowski J, Hao L et al. Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* 2014; 158(5):1000-1010.
37. Zhao Y, Wu J, Li JV, Zhou NY, Tang H, Wang Y. Gut microbiota composition modifies fecal metabolic profiles in mice. *J Proteome Res* 2013; 12(6):2987-2999.
38. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermudez-Humaran LG, Gratadoux JJ et al. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci U S A* 2008; 105(43):16731-16736.
39. Song H, Yoo Y, Hwang J, Na YC, Kim HS. Faecalibacterium prausnitzii subspecies-level dysbiosis in the human gut microbiome underlying atopic dermatitis. *J Allergy Clin Immunol* 2016; 137(3):852-860.

40. Cani PD, Bibiloni R, Knauf C, Waget A, Neyrinck AM, Delzenne NM et al. Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet-induced obesity and diabetes in mice. *Diabetes* 2008; 57(6):1470-1481.
41. Fleissner CK, Huebel N, Abd El-Bary MM, Loh G, Klaus S, Blaut M. Absence of intestinal microbiota does not protect mice from diet-induced obesity. *Br J Nutr* 2010; 104(6):919-929.
42. Ravussin Y, Koren O, Spor A, LeDuc C, Gutman R, Stombaugh J et al. Responses of gut microbiota to diet composition and weight loss in lean and obese mice. *Obesity (Silver Spring)* 2012; 20(4):738-747.
43. Martinez I, Wallace G, Zhang C, Legge R, Benson AK, Carr TP et al. Diet-induced metabolic improvements in a hamster model of hypercholesterolemia are strongly linked to alterations of the gut microbiota. *Appl Environ Microbiol* 2009; 75(12):4175-4184.
44. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010; 26(19):2460-2461.
45. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 2012; 6(3):610-618.
46. Navas-Molina JA, Peralta-Sanchez JM, Gonzalez A, McMurdie PJ, Vazquez-Baeza Y, Xu Z et al. Advancing our understanding of the human microbiome using QIIME. *Methods Enzymol* 2013; 531:371-444.

Part

JUVENILE DERMATOMYOSITIS



Chapter

5

MODELLING DISEASE ACTIVITY IN JUVENILE DERMATOMYOSITIS: A BAYESIAN APPROACH

EH Pieter van Dijkhuizen, Claire T Deakin,
Lucy R Wedderburn and Maria De Iorio

ABSTRACT

Juvenile dermatomyositis is the most common form of the juvenile idiopathic inflammatory myopathies characterised by muscle and skin inflammation, leading to symmetric proximal muscle weakness and cutaneous symptoms. It has a fluctuating course and varying prognosis. In a Bayesian framework, we develop a joint model for four longitudinal outcomes, which accounts for within individual variability as well as inter-individual variability. Correlations among the outcome variables are introduced through a subject-specific random effect. Moreover, we exploit an approach similar to a hurdle model to account for excess of a specific outcome in the response. Clinical markers and symptoms are used as covariates in a regression set-up. Data from an ongoing observational cohort study are available, providing information on 340 subjects, who contributed 2725 clinical visits. The model shows good performance and yields efficient estimations of model parameters, as well as accurate predictions of the disease activity parameters, corresponding well to observed clinical patterns over time. The posterior distribution of the by-subject random intercepts shows a substantial correlation between two of the outcome variables. A subset of clinical markers and symptoms are identified as associated with disease activity. These findings have the potential to influence clinical practice as they can be used to stratify patients according to their prognosis and guide treatment decisions, as well as contribute to on-going research about the most relevant outcome markers for patients affected by juvenile dermatomyositis.

INTRODUCTION

The juvenile idiopathic inflammatory myopathies (IIM) are a group of heterogeneous chronic diseases whose prime manifestation is skeletal muscle inflammation. Juvenile dermatomyositis (JDM) is the most common of the juvenile IIM, with an incidence of about 1.9–4.1 per million children per year.¹ Besides the skeletal muscle inflammation, JDM is characterised by skin inflammation and manifests itself clinically as symmetric proximal muscle weakness and cutaneous symptoms such as papules over the knuckles (Gottron's papules), skin rashes, nail fold erythema, ulceration and calcinosis (deposition of subcutaneous or intramuscular calcium). JDM is a systemic disease, potentially involving other organ systems as well. If not adequately treated, the inflammation may cause muscular damage, leading to permanent muscle weakness and life-long disabilities, and skin atrophy.^{2,3} The prognosis of JDM is variable with about 24–40% of patients experiencing a monocyclic course, and the majority of patients (50–60%) showing a chronic disease pattern. The mortality rate of JDM is about 2–3%.³

This clinical picture highlights that JDM is a disease with a significant medical, social and economic burden.^{1,4} Some studies suggest that early aggressive treatment improves the outcome for children with JDM, thereby alleviating this burden.⁵ On the other hand, the treatment itself is associated with serious side effects, such as growth failure, pubertal delay and osteoporosis. As a consequence, it is good practice to administer as few drugs as possible, especially in paediatrics, in order to limit the negative effects on the normal physiology of the developing child. As such, the clinical challenge is to stratify patients according to their prognosis and adapt their treatment plans in concordance with their expected prognosis.

To answer the clinical question of interest, we need to identify clinical markers associated with disease activity and capable of predicting its evolution. Insight has been gained by the study of the various antibody patterns in the juvenile IIM^{2,3,6} mainly to distinguish between JDM and the other subtypes of the juvenile IIM. Two previous studies have evaluated the potential of clinical variables to predict disease activity^{1,7} but these studies have the limitation of measuring disease activity as a binary outcome (active vs. non-active). Moreover, whereas Stringer *et al.*⁷ analyse the time to remission, Ravelli *et al.*¹ evaluate disease activity cross-sectionally at approximately 7.7 years after disease onset.

In contrast, the aim of the current study is to identify variables associated with disease activity of JDM by modelling clinically relevant, continuous response variables over time using clinical markers and symptoms of the disease as covariates. Such a model could be used to inform physicians about features associated with disease activity, as well as guide clinicians in the management of the disease.

DATA

Data collection

Data are available from a national longitudinal cohort study, in which children with definite or probable JDM according to Bohan and Peter criteria^{8,9} are enrolled across the UK. The study has been described previously.⁴

Enrolment started in 2000 and is ongoing. Patients are enrolled at diagnosis, or as soon as they come to the attention of one of the participating centres, and are followed up every 3 months for 2 years and subsequently at least annually. Data collected concern disease manifestations, such as muscular involvement, skin disease and other organ involvement. Routine laboratory blood tests are also performed.

At each visit four outcome parameters¹⁰ are measured as proxy of disease activity:

- Serum creatine kinase level (CK), a muscle enzyme which is released into the blood stream in increased amounts in case of muscular inflammation. The serum level of this enzyme is usually elevated in active JDM, but tends to be relatively low.³ Its theoretical range is in $(0, \infty)$ and values exceeding 150 U/L are considered pathological.¹⁰ Many laboratories report a censored value of <20 U/L if the level is below the laboratory's detection limit.
- Childhood myositis assessment scale (CMAS) which assesses muscular strength and endurance.¹¹ The CMAS consists of a series of little tasks, which the child is asked to perform, such as raising his/her head for at least 2 minutes while laying supine, raising to his/her feet from a sitting position and performing sit ups, with and without counterbalance. A predefined number of points is assigned based on the performance of the child and the total CMAS score is the sum over all tasks, ranging from 0 to 53, with a lower score indicating more active disease. Scores of at least 48 points are considered normal.¹⁰
- Manual muscle testing of 8 muscle groups (MMT8), which consists in the assessment of muscle strength of 8 predefined muscle groups. Each muscle group is then assigned by the examiner a score ranging from 0 (equal to a complete loss of muscle contraction) to 10 (equal to full muscle strength). The total MMT8 is the sum over all muscle groups, thus ranging from 0 to 80, with a lower score indicating higher disease activity. A value of at least 78 points is considered normal.¹⁰
- Physician's global assessment of disease activity (PGA), which is the physician's rating of disease activity, taking into account muscular disease, skin involvement and all other organ involvement. The PGA ranges from 0.0, indicating no disease activity, to 10.0, indicating maximal disease activity. The PGA should be at most 0.2 points to be considered normal.¹⁰

Ethical approval of the study was obtained and all participants and their parents (as appropriate) provided informed consent. The study was conducted according to good clinical practice guidelines and the declaration of Helsinki.

Data cleaning

At the time of analysis, 4122 visits of 469 patients are available. Patients not fulfilling the inclusion criteria of having probable or definite JDM according to Bohan and Peter criteria are excluded, leaving 413 patients contributing 3881 visits. Missing data are imputed in the Bayesian model, except for missing data on history variables, because these are based on patient's recollection. This leads to the exclusion of 73 patients. Therefore, the final model includes 340 patients, contributing 2725 visits. Baseline characteristics of included and excluded patients are compared in Table 1. As expected, due to the exclusion of visits with missing data, patients having contributed fewer visits are more likely to be excluded altogether from the study. This, in turn, is more likely to occur in patients enrolled later in the course of their disease, therefore presenting less disease activity at enrolment (Table 1).

Of the 340 included subjects, five fully observed participants (i.e., no missing values in the dependent and independent variables for all visits) are randomly selected for the out of sample predictions and are therefore excluded from the training set.

The original dataset contains 83 covariates that can be used as predictor variables, including the time elapsed since diagnosis at the visit, disease signs and symptoms and treatment variables. The number of covariates is reduced by excluding those hardly containing any information. To this end, a pre-selection is made by fitting linear mixed effects models, using the package `lme4`¹² available in the R software.¹³ More in details, for each covariate we fit four independent linear mixed models, one for each outcome, including also time from diagnosis and a random intercept for the time from diagnosis. We obtain a p-value for the covariate by employing a Wald *t* test and retain the smallest one

Table 1. Baseline table

Parameter	Included <i>N</i> = 340	Excluded <i>N</i> = 73
Female, <i>n</i> (%)	236 (69.4)	54 (74.0)
Age at diagnosis, y	7.4 [4.5, 10.5]	7.3 [4.1, 11.1]
Disease duration at diagnosis, y	0.3 [0.2, 0.6]	0.3 [0.2, 1.0]
Time after diagnosis at enrolment, y	0.2 [0.1, 1.1]	2.3 [0.4, 5.4]
Duration of follow up, y	4.1 [1.6, 7.1]	1.2 [0.1, 2.6]
Disease activity at enrolment:		
CK, U/L	103 [64, 440]	98 [45, 256]
CMAS, points	41 [21, 50]	46 [37, 52]
MMT8, points	65 [45, 80]	80 [64, 80]
PGA, cm	3 [1.3, 6.0]	2.3 [0.5, 4.0]

Values are the median [1st quartile, 3rd quartile], except where indicated otherwise.

Abbreviations: CK, creatine kinase; cm, centimetre; CMAS, childhood myositis assessment scale; MMT8, manual muscle testing of 8 muscle groups; PGA, physician's global assessment of disease activity; U/L, units per litre; y, year.

among the four models. We then rank the covariates according to their most significant p-value and in such a way we select the 50% most significant predictors.

MODEL

Model specification

As discussed previously, the data consist of $N = 340$ patients, each of them contributing multiple visits. For each patient i at visit j , a vector of four outcome variables is measured, $Y_{ij} = (Y_{ij}^{(\text{CK})}, Y_{ij}^{(\text{CMAS})}, Y_{ij}^{(\text{MMT8})}, Y_{ij}^{(\text{PGA})})$, $i = 1, 2, \dots, N, j = 1, 2, \dots, m_i$, where m_i is the number of observations for the i th patient. Additionally, for each subject i at visit j , a vector of covariates X_{ij} is available and t_{ij} denotes the time elapsed since diagnosis. To account for the longitudinal nature of the study and for dependency between visits for the same patients, we introduce a subject-specific random effect for each of the response variables. We then link the response variables by assuming a joint model for the random effects.

The continuous outcome variables are transformed to ensure normality: we take a log-transformation for the CK levels, whereas we choose a square-root transformation for CMAS, MMT8 and PGA, as it is a well-known variance stabilising transformation that can deal with the presence of zeros in the data. For ease of notation, in what follows $\log(\text{CK})$, $\sqrt{\text{CMAS}}$, $\sqrt{\text{MMT8}}$ and $\sqrt{\text{PGA}}$ are referred to as CK, CMAS, MMT8 and PGA, respectively. The empirical distribution of the transformed outcome measures is shown in Figure 1. These variables are modelled using a Bayesian approach.

Let us first consider the outcome variable CK, which we assume normally distributed:

$$Y_{ij}^{(\text{CK})} | \mu_{ij}^{(\text{CK})}, \sigma_{(\text{CK})}^2 \sim N(\mu_{ij}^{(\text{CK})}, \sigma_{(\text{CK})}^2)$$

where

$$\mu_{ij}^{(\text{CK})} = \alpha^{(\text{CK})} + \eta_i^{(\text{CK})} + (\gamma^{(\text{CK})} + \theta_i^{(\text{CK})}) \times t_{ij} + X_{ij} \beta^{(\text{CK})}$$

Here, $\alpha^{(\text{CK})}$ is a mean effect common to all subjects, $\beta^{(\text{CK})}$ is the vector of coefficients for the covariates X_{ij} , $\gamma^{(\text{CK})}$ is the common effect for the time elapsed since diagnosis t_{ij} , $\eta_i^{(\text{CK})}$ is a subject-specific random intercept, while $\theta_i^{(\text{CK})}$ is a patient-specific random coefficient for the time elapsed since diagnosis.

As mentioned above, some of the CK values ($n = 14$, 0.5% of the observations) are below the laboratory's detection level and reported as <20 or <21 (depending on the laboratory). These instances are modelled as truncated values, by specifying this observation as right-truncated at 20 or 21.

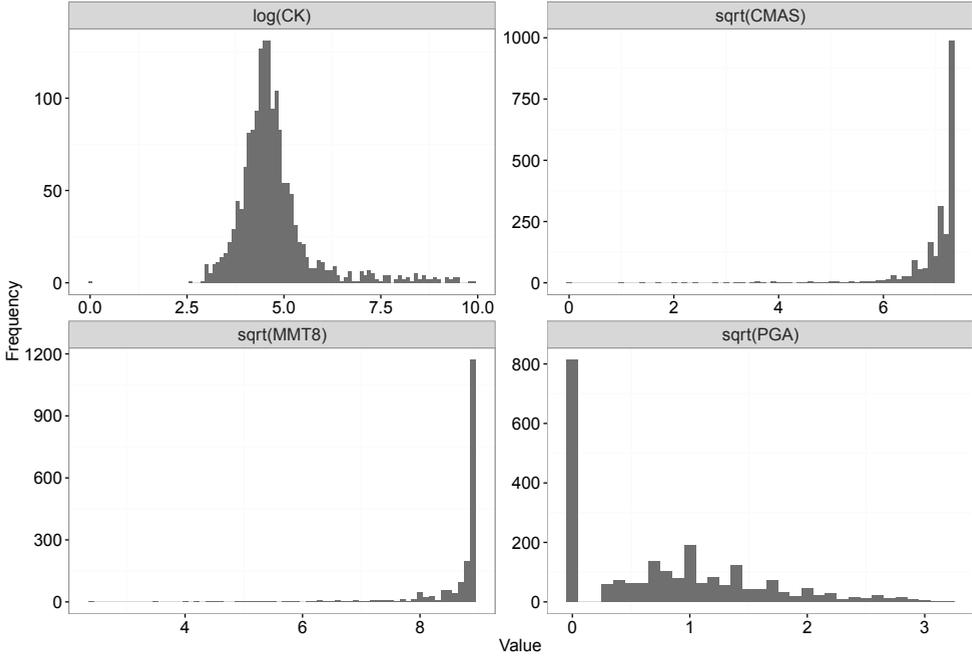


Figure 1. Empirical distribution of the four transformed outcomes.

Most patients reach disease remission after a certain period of time and a large number of non-pathological values for CMAS, MMT8 and PGA are observed. From Figure 1, it is clear that these variables are not normally distributed, but they show an excess of non-pathological values. For these clinical outcomes we assume a mixture of a point mass at the most frequent value ($\sqrt{53}$ for CMAS, $\sqrt{80}$ for MMT8 and 0 for PGA) and a truncated Normal distribution for the other observations. This strategy is similar to the hurdle model¹⁴ as we recognise the possibility that the mechanism determining non-pathological values may differ from those that influence the distribution of pathological observations. We specify a probit regression to link the probability of observing a non-pathological value for subject i at visit j to the time elapsed since diagnosis. Moreover, we assume a linear regression model for the continuous part of the mixture.

More in details, let $k \in \{\text{CMAS}, \text{MMT8}, \text{PGA}\}$ denote one of the clinical outcomes of interest. Then,

$$Y_{ij}^{(k)} | w_{ij}^{(k)}, \mu_{ij}^{(k)}, \sigma_{(k)}^2 \sim w_{ij}^{(k)} \delta_{y_{ij}^{(k)}} + (1 - w_{ij}^{(k)}) \times TN(\mu_{ij}^{(k)}, \sigma_{(k)}^2) \quad (1)$$

where δ is the Kronecker delta and the non-pathological values are:

$$Y_*^{(k)} = \begin{cases} \sqrt{53}, & \text{if } k = \text{CMAS} \\ \sqrt{80}, & \text{if } k = \text{MMT8} \\ 0, & \text{if } k = \text{PGA} \end{cases}$$

Here, $TN(\mu, \sigma^2)$ denotes a truncated Normal distribution with mean μ and variance σ^2 . The truncation limit changes depending on the outcome variable, but the idea is to truncate the normal distribution according to the range of possible values the response can assume (since they represent clinical scores) and the non-pathological value. As such, on the transformed scale, the distribution of CMAS is truncated on the interval $(0, \sqrt{53})$, the distribution of MMT8 on $(0, \sqrt{80})$ and the distribution of PGA on the interval $(0, \sqrt{10})$.

We assume a probit regression for the weights of the mixture:

$$w_{ij}^{(k)} = \Phi(z_{ij}^{(k)})$$

where Φ is the cumulative distribution function of the standard Normal distribution and

$$z_{ij}^{(k)} = \alpha^{(k)} + \eta_i^{(k)} + \gamma^{(k)} \times t_{ij} \quad (2)$$

As in the case of creatine kinase, $\alpha^{(k)}$ and $\gamma^{(k)}$ are the common intercept and regression coefficient, respectively, common to all subjects, while η_i is the patient specific random effect. We complete the model by specifying a linear regression model for the continuous component of the mixture (see equation (1)):

$$\mu_{ij}^{(k)} = \psi^{(k)} + \lambda^{(k)} \times t_{ij} + X_{ij} \beta^{(k)}$$

Note that we do not include an individual random effect for $\mu_{ij}^{(k)}$.

The four disease outcomes are measured in the same patient at the same time and capture different aspects of the same biological entity, i.e. disease activity. As such, they are correlated, especially CMAS and MMT8. Therefore, it is important to account for these correlations in the model. Following Li *et al.*,¹⁵ we specify a multivariate Normal distribution on subject-specific random effects to take the inherent correlation of the longitudinal outcomes into account. This approach leads to a more efficient estimation of the model parameters. Hence, the four by-subject random intercepts for the outcome variables are modelled as follows:

$$\eta_i = \begin{pmatrix} \eta_i^{(\text{CK})} \\ \eta_i^{(\text{CMAS})} \\ \eta_i^{(\text{MMT8})} \\ \eta_i^{(\text{PGA})} \end{pmatrix} \sim \mathbf{N}(\mathbf{0}, \Omega) \quad (3)$$

where $\mathbf{0}$ is a four-dimensional vector of zeros and Ω denotes the precision matrix.

Prior specification

The model is completed by specifying uninformative prior distributions on the remaining parameters. We use independent Normal(0, 1000) priors for the intercepts $\alpha^{(k)}$ and $\psi^{(k)}$, as well as for the regression coefficients $\gamma^{(k)}$ and $\lambda^{(k)}$, with $k \in \{\text{CK}, \text{CMAS}, \text{MMT8}, \text{PGA}\}$. We place a Normal(0, σ_θ^2) prior on the patient-specific random slope $\theta_i^{(\text{CK})}$ for the CK model, as well as a Normal(0, σ_β^2) prior on the regression coefficients $\beta^{(k)}$, with common variance σ_β^2 shared by all $\beta^{(k)}$. We elicit independent Gamma distributions with hyper-parameters (0.001, 0.001) for the observation precision $\tau^{(k)} = \frac{1}{\sigma_{(k)}^2}$, $k \in \{\text{CK}, \text{CMAS}, \text{MMT8}, \text{PGA}\}$, $\tau_\theta = \frac{1}{\sigma_\theta^2}$ and $\tau_\beta = \frac{1}{\sigma_\beta^2}$.

We assume a Multivariate Normal prior as random effects distribution. We choose for computational reasons a conjugate Wishart prior for the precision matrix Ω in equation (3) with degrees of freedom $\nu = 6$ and centering matrix $S = I \times \nu \times 10$, where I is the identity matrix of appropriate dimension. As such the prior mean is given by $\mathcal{E}(\Omega) = 0.1I$.¹⁵

Bayesian variable selection

To identify the most influential variables on disease progression and to obtain a sparser and more interpretable model, we perform Bayesian variable selection of the 47 covariates entered into the model (except for time from diagnosis), following the approach proposed by Kuo and Mallick.¹⁶ See O'Hara and Sillanpää¹⁷ for a review of Bayesian variable selection methods.

In short, for each covariate we introduce an indicator variable $\delta_p, p=1,2,\dots,P$ taking values in $\{0, 1\}$. Here, P denotes the total number of available covariates. When $\delta_p = 1$, we include the p th predictor in the regression model. When $\delta_p = 0$, we omit the p th predictor. We then add the indicator functions to the regression term by defining new coefficients

$$\tilde{\beta}_p = \delta_p \beta_p$$

We complete the model by eliciting independent prior distributions on the δ_p :

$$\delta_p \sim \text{Bernoulli}(\pi_p), \quad p = 1, 2, \dots, P$$

$$\pi_p \sim \text{Beta}(0.1, 0.1)$$

The Beta hyperprior induces sparsity as it places most of its mass on 0 or 1. The same indicator variable is used for all instances of each covariate in the model, so that the covariate would be included or not simultaneously in the models describing the four longitudinal outcomes, i.e. the mean of the CK model, and the linear parts of the CMAS, MMT8 and PGA models. The prior on the coefficients β_p is still a Normal distribution with mean 0 and common precision τ_β , to which we assign a Gamma distribution with parameters (0.001, 0.001), as described previously.

RESULTS

Posterior inference is performed in the R software version 3.2.2¹³ and JAGS, using the package rjags¹⁸ as interface. We run the chain for 30,000 iterations with a burn-in of 5000 and 2000 adaptation iterations. We thin the chain every 10 iterations.

The goodness of fit of the model is assessed visually in six randomly selected patients, each having contributed at least 10 visits. Their observed values of CK, CMAS, MMT8 and PGA are plotted over time, together with the model fit and 95% credible intervals (CI). The plot for PGA is shown in Figure 2, plots for CK, CMAS and MMT8 were similar. The plots show a good fit of the model to the data. In particular, fluctuating patterns over time are predicted well by the model. Some observed values in Figure 2 are outside the posterior 95% CI, however, this happens only for 1.8% of the observations.

Moreover, to better understand the predictive ability of the model, we consider the posterior predictive probability of an observation falling beyond the cut-off point for pathological values for each of the four longitudinal outcomes, as defined in literature (i.e., >150 for CK, <48 for CMAS, <78 for MMT8 and >0.2 for PGA; note that these cut-off points are defined on the original measurement scale¹⁰). These predictions concern pathological values and are, therefore, relevant from a clinical point of view.

In order to assess the predictive performance of the proposed model we employ the Brier statistic.¹⁹ This statistic assesses the quality of predictions when the response variable is binary:

$$\text{Brier} = \frac{1}{N} \sum_{i=1}^N (f_i - o_i)^2$$

where o_i is a binary observation, f_i is its predicted probability and N is the number of observations in the sample.

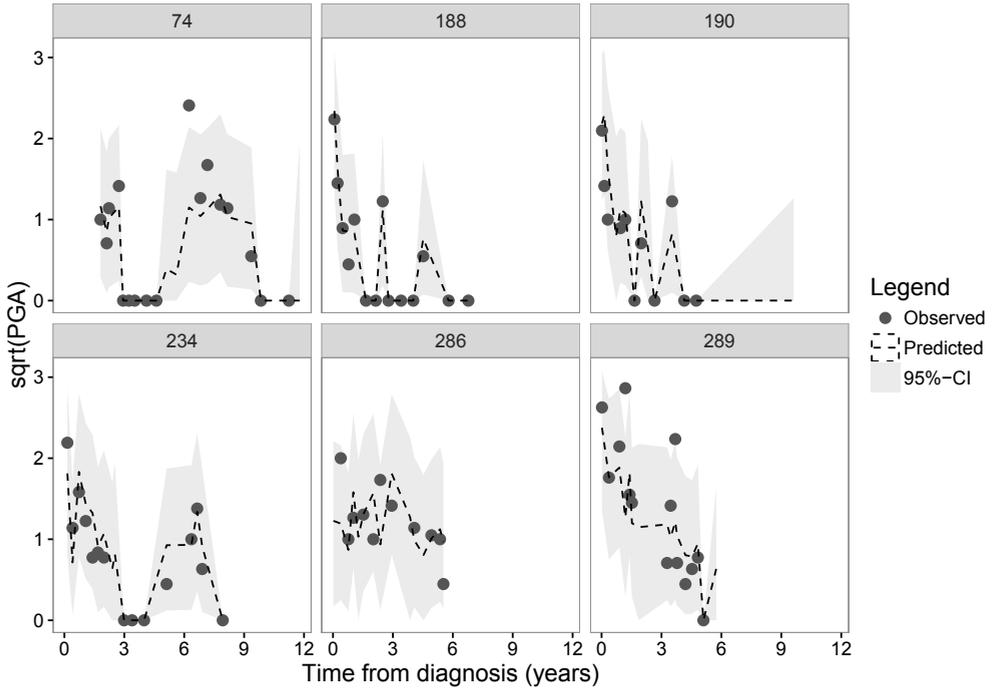


Figure 2. Model fit for PGA. Observed values of sqrt(PGA) over time of six randomly selected subjects are represented by dark grey dots, the fitted trajectory by the dashed line and the 95% credible interval around the fit is given by the shaded area.

In our case, the response variable is continuous, so in order to apply this statistic we discretise the response variable. We are interested in predicting whether each clinical outcome is above or below the above-mentioned thresholds. Then, the predicted f_{ij} is the predictive probability of obtaining values larger or smaller than the specified threshold for each patient i in the sample and visit j . Hence, the Brier statistic becomes:

$$\text{Brier} = \frac{1}{\sum_{i=1}^N m_i} \sum_{i=1}^N \sum_{j=1}^{m_i} (f_{ij} - Y_{ij}^d)^2$$

where Y_{ij}^d is the ij th response discretised with respect to the critical threshold (equal to 1 if Y_{ij} assumes pathological values and 0 otherwise), N is the number of patients and m_i is the number of observations for individual i . The sum is taken over all the patients in the sample and the observation times. The Brier score assumes values between 0 and 1, with lower scores indicating a better performance.²⁰ We obtain Brier scores equal to 0.12, 0.12, 0.08 and 0.05 for CK, CMAS, MMT8 and PGA, respectively. Following Steyerberg *et al.*,²⁰ we calculate rescaled Brier scores, which are normalised with respect to

the maximum obtainable Brier score for a non-informative model based on the observed frequency of the outcomes. The scaled Brier score is very similar to Pearson's R^2 statistic (high values indicate a good predictive performance). In our analysis, we obtain 49%, 42%, 63% and 80% for CK, CMAS, MMT8 and PGA, respectively, indicating moderate to good performance of the model.

We also perform out of sample predictions on the four outcomes for five randomly selected, fully observed patients collectively contributing seven observations (Figure 3). Once again, the results show good accuracy, with the model's predictions being close to the observed values and almost always on the same side of the cut-off point discriminating pathological values from non-pathological values, as established in the literature (150 U/L for CK, 48 points for CMAS, 78 points for MMT8 and 0.2 points for PGA).¹⁰ In only two cases involving CMAS, the observed value is below the cut-off point, indicating a pathological value, whereas the predicted value is above. However, on a closer inspection we notice that in both these cases the observed values are close to the cut-off point (46 and 47, respectively). The relatively wide CIs reflect the uncertainty in the predictions, the estimation of the parameters, as well as in the imputation of missing values.

The posterior estimates of the regression coefficients of the time elapsed since diagnosis and the 46 covariates are reported in Figure 4. The variability associated with the estimates is reasonable as shown by the CI. As expected, CMAS, MMT8 and PGA tend

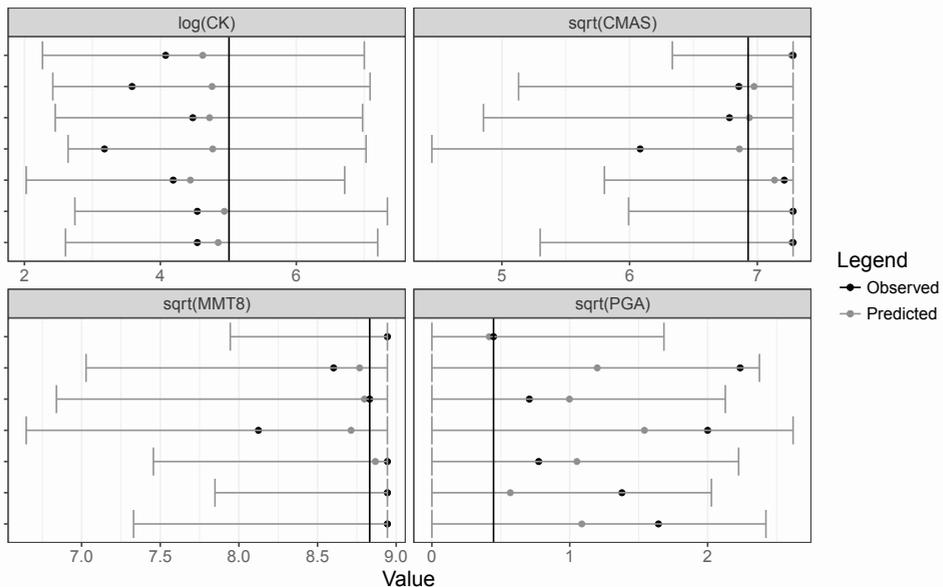


Figure 3. Out-of-sample predictions. Observed values are represented by black dots, while predicted values with corresponding 95% credible intervals are represented by grey dots with error bars for all four outcomes in seven visits of five randomly selected patients. The vertical line indicates the pathological value as defined in literature.

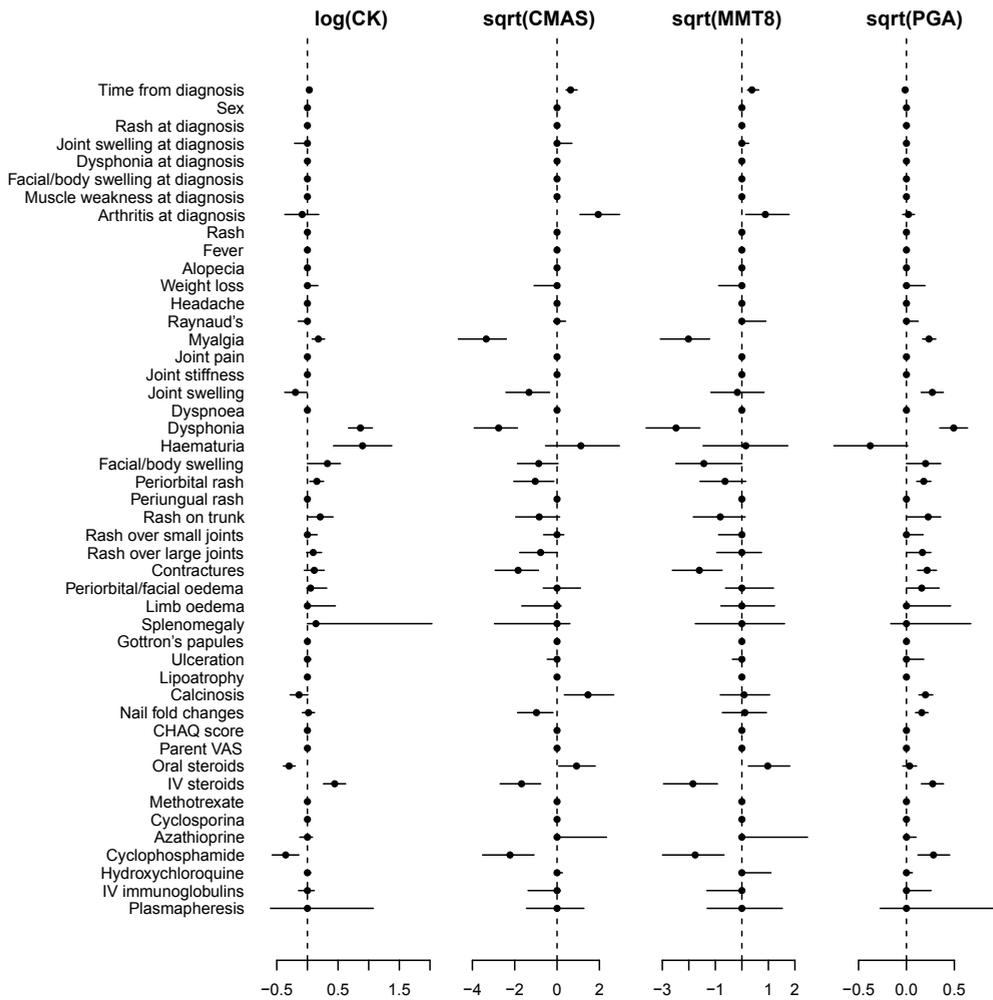


Figure 4. Medians (dots) and 95% credible intervals (horizontal lines) for the fixed effects for the time elapsed since diagnosis and all the covariates. We plot the results for each of the four outcomes.

to normalise over time, whereas the CK level does not significantly change from the time from diagnosis. Various clinical signs and symptoms are associated with the four outcomes, most notably a range of manifestations of skin disease, such as periorbital rash, rash on the trunk, rash over large joints, nail fold changes and facial swelling. These parameters are not only associated with PGA, but also with the strictly muscular outcomes CK and CMAS, indicating that patients who are doing worse with respect to muscular disease activity, tend to have higher cutaneous disease activity as well. As expected, patients having myalgia (muscle pain) or dysphonia (altered voice pitch due to inflammation of the muscles of the vocal cords) have much worse disease activity. Of signs and symptoms

present at diagnosis, only arthritis is associated with improved disease activity with regard to the CMAS and MMT8 at the follow up visits, probably reflecting a subset of patients with an overlap syndrome between JDM and juvenile idiopathic arthritis (JIA).

The methodology developed by Li *et al.*¹⁵ for modelling multiple correlated outcomes is adopted, because of the prior belief that the four outcomes for JDM are correlated. The posterior distribution of the correlation between the by-subject random intercepts for the four outcomes supports this prior belief. The estimates are shown in Table 2, evidencing a substantial correlation between $\eta_i^{(\text{CMAS})}$ and $\eta_i^{(\text{MMT8})}$ (0.54, 95% CI 0.42–0.65), highlighting the appropriateness of the adopted methodology.

In Appendix A, we present four extensions of the proposed modelling strategy and compare their performance to the one of the model presented in the previous section. Appendix B contains the JAGS code used to fit the original model.

DISCUSSION

The aim of this work is to jointly model four correlated longitudinal outcomes that measure disease activity in JDM, including clinical markers and symptoms as covariates. To the best of our knowledge, this is the first study to model disease activity in JDM, taking account of (i) longitudinal disease activity; (ii) continuous outcome measures; (iii) four different response variables, measured simultaneously, accounting for correlations among them; and (iv) the particular characteristics of the distribution of the outcome variables due to the fact that many patients attain disease remission at a certain time point.

To achieve this, we propose a statistical approach based on the work of Li *et al.*¹⁵ and specify a joint random effects distribution for the four outcomes (see equation (3)). Moreover, since most enrolled patients present disease remission at a certain point in time, normal (i.e., non-pathological) values are observed for CMAS, MMT8 and PGA for a large number of follow up visits. Therefore, we observe an excess of non-pathological values

Table 2. Estimated correlations with 95% credible intervals of subject random intercepts for the four outcomes CK, CMAS, MMT8 and PGA.

Comparison	$\hat{\rho}$	95% CI
CK vs. CMAS	0.07	(-0.07, 0.23)
CK vs. MMT8	-0.001	(-0.16, 0.15)
CK vs. PGA	0.0007	(-0.15, 0.16)
CMAS vs. MMT8	0.54	(0.42, 0.65)
CMAS vs. PGA	0.17	(0.01, 0.32)
MMT8 vs. PGA	0.23	(0.08, 0.32)

Abbreviations: CI: credible interval; CK: creatine kinase; CMAS: childhood myositis assessment scale; MMT8: manual muscle testing of 8 muscle groups; PGA: physician's global assessment.

in the empirical distribution of these variables. To account for this peak in the empirical distribution, we adopt an approach similar to a hurdle model for CMAS, MMT8 and PGA.¹⁴

Our results show that the proposed model leads to accurate prediction of the outcome variables, corresponding well to observed patterns in the data over time. Moreover, even though the CIs of the predictions are often large, the variability associated with the estimates for the regression coefficients is reasonable, as shown by relatively narrow CI. Indeed, Li *et al.*¹⁵ demonstrated that their methodology leads to a more efficient estimation of the model parameters compared to a strategy that considers each outcome individually. Moreover, the posterior distribution of Ω shows evidence of correlation between the by-subject random intercepts, especially between those for CMAS and MMT8 ($\hat{\rho} = 0.54$, 95% CI 0.42–0.65, Table 2), indicating that those subjects with a higher CMAS tend to have a higher MMT8 as well. This corresponds to existing clinical knowledge.

Our analysis shows that disease activity tends to decrease over time, especially with respect to CMAS, MMT8 and PGA, whereas CK levels in general do not change significantly over time (the change in CK level is estimated to be $e^{0.031} = 1.03$ U/L per year follow up; Figure 4). This is consistent with the clinically observed fact that CK levels tend to be relatively low in patients with JDM.³ In the described cohort, the median CK level at baseline, shortly after diagnosis, is 103 U/L (Table 1), which is considered a normal value.¹⁰ Nonetheless, patients with myalgia, dysphonia or haematuria, all signs of more severe (systemic) disease, present more elevated CK levels (Figure 4).

Furthermore, we find that different cutaneous manifestations of the disease, most notably periorbital rash, rash on the trunk, rash over large joints, nail fold changes and facial swelling, are associated with disease activity, especially with PGA, which is the only clinical response taking cutaneous disease activity into account directly. However, some of these skin manifestations, such as periorbital rash and nail fold changes, are associated with the strictly muscular disease activity responses as well. These findings are in line with previous studies, even though based on a different methodology.^{1,7}

While Ravelli *et al.*¹ found associations between disease activity and sex, age at onset and dysphonia/dysphagia at onset, this study does not replicate these results (although we find an association between dysphonia at the visit and all four outcome variables). These differences may be explained by differences in ethnic composition of the samples (a mix of European and Latin-American patients for Ravelli *et al.*¹ while this study recruits only UK patients) and methodology. Specifically, Ravelli *et al.*¹ analysed each outcome individually using logistic regression at a median of 7.7 years after disease onset, therefore discretising the response and not accounting for the development over time, while we propose joint modelling of the longitudinal outcome on their original scale. Furthermore, this analysis includes time-varying clinical markers and symptoms determined at each study visit, whereas Ravelli *et al.*¹ mainly considered baseline variables. It can be hypothesised that the association between current signs and symptoms and disease activity parameters

overwhelms the associations between baseline parameters and disease activity at the study visits. Indeed, most baseline variables are not shown to be influential when we perform Bayesian variable selection (Figure 4). This is probably due to the fact that we integrate in the model information from different sources and across time, allowing us to better explain variability in the responses.

Since we did not impute history variables, we have excluded visits with missing values for these variables. Owing to this procedure, 73 patients are excluded altogether from the study. From a probabilistic point of view, patients contributing less visits are less likely to have any visit with completely observed history variables. This, in turn, is more likely to happen in patients who are enrolled in the study not immediately at the time of diagnosis, but later during their course of disease. Consequently, they have lower disease activity at enrolment. These observations explain the differences between included and excluded patients (Table 1). These patients represent a biased subset of the study, often contributing just a few visits, and we believe their exclusion is not problematic. Missing values in non-history variables are imputed in the Bayesian model, by assuming missingness at random and as a result we are able to retain a large number of patients and visits in the model.

In conclusion, the proposed methodology based on joint modelling of four correlated longitudinal measures of disease activity in JDM and exploiting a hurdle model approach, leads to a well-fitting model able to explain observed patterns over time, resulting in clinically meaningful inferences. Various clinical markers and symptoms of JDM are shown to be associated with disease activity. The clinical implications of these findings will be discussed in a clinical paper.

APPENDIX A: ADDITIONAL MODELS

We fit four additional models to the data and their performance is assessed and compared to that of the original model using the deviance information criterion (DIC).²¹ Results are reported in Table 3. Note that the original model makes use of the JAGS dinterval() distribution to deal with censored CK values, which has to be removed in order to calculate the DIC. Posterior inference is performed in the R software¹³ and JAGS, using the package rjags.¹⁸ In order to calculate DICs, two chains are required. Each chain is run for 15,000 iterations with a burn-in of 5000 and 2000 adaptation iterations. As before, we thin the chain every 10 iterations.

The DIC of the model presented in the paper is 4.1×10^6 (Table 3).

The first additional model we consider contains a time-dependent error term for all four outcomes in the model. Thus, instead of eliciting independent Gamma distributions with hyper-parameters (0.001, 0.001) for the observation precision $\tau^{(k)} = \frac{1}{\sigma_{(k)}^2}$, $k \in \{\text{CK, CMAS, MMT8, PGA}\}$, σ_k^2 is assumed to follow:

$$\log(\sigma_{(k),ij}^2) = \alpha_{\sigma}^{(k)} + \beta_{\sigma}^{(k)} \times t_{ij}$$

where t_{ij} is the time elapsed from diagnosis for individual i at visit j . Independent Normal(0, 1000) priors are specified for the coefficients $\alpha_{\sigma}^{(k)}$ and $\beta_{\sigma}^{(k)}$. The DIC of this model is 4.0×10^6 (Table 3). Even though this is slightly lower than the DIC of the original model, we judge this improvement too little to justify the increased complexity of the model with time-dependent error terms. Moreover, visual inspection of the out-of-sample predictions reveals that they do not improve in comparison with the original model (same accuracy and precision).

We also consider a model in which missing values are imputed conditionally on the sex of the patient and the time elapsed since diagnosis. This model, too, yields a slight improvement in DIC (4.0×10^6 , Table 3), but once again the reduction in DIC is judged too small to justify the increased complexity.

Furthermore, we investigate the effect of including covariate information in the model for the weights $w_{ij}^{(k)}, k \in \{\text{CMAS, MMT8, PGA}\}$. In this case, equation (2) becomes:

$$z_{ij}^{(k)} = \alpha^{(k)} + \eta_i^{(k)} + \gamma^{(k)} \times t_{ij} + X_{ij} \beta_w^{(k)}$$

where we added the term $X_{ij} \beta_w^{(k)}$. The $\beta_w^{(k)}$ are subject to the same Bayesian variable selection procedure as described for the original model. Thus, we define new coefficients $\tilde{\beta}_{w,p} = \delta_p \beta_{w,p}$ and specify independent Normal(0, $\frac{1}{\tau_{\beta}}$) priors for $\beta_{w,p}$. Here, δ_p and τ_{β} are defined as previously. The same δ_p is used for all instances of the covariate in the model, i.e., the linear part of CK and the linear parts and weights of CMAS, MMT8 and PGA. The DIC of this model is 7.2×10^6 , which is higher than the original model (Table 3).

Finally, we extend the original model by introducing random intercepts and random slopes for the effect of time since diagnosis in all components of the model, i.e., the linear

Table 3. Deviance information criterion (DIC) for four extensions of the model.

Model	Mean deviance	Penalty	DIC
TVE	2.1×10^4	4.0×10^6	4.0×10^6
CI	2.0×10^4	4.0×10^6	4.0×10^6
OM	2.2×10^4	4.0×10^6	4.1×10^6
CW	2.2×10^4	7.2×10^6	7.2×10^6
RIS	2.2×10^4	13.5×10^6	13.5×10^6

TVE: model with time dependent error terms; CI: model with imputation based on time from diagnosis and sex; OM: original model; CW: model that specifies a regression term for the weights of hurdle models for CMAS, MMT8 and PGA; RIS: model that includes a random intercept and slope for the coefficient of time from diagnosis for the mean and the weights of all four response variables.

term of the CK response as well as the linear terms and weights of CMAS, MMT8 and PGA models. Due to the significantly larger number of parameters, the DIC is much higher, 13.5×10^6 (Table 3).

APPENDIX B: JAGS CODE

This appendix includes the JAGS code of the model presented in this paper. For the sake of brevity, imputations are removed from the code presented here.

```

data {
  mu.ranint[1] <- 0
  mu.ranint[2] <- 0
  mu.ranint[3] <- 0
  mu.ranint[4] <- 0
}
model {
  # Imputations omitted
  # Random intercepts and random slope for time from diagnosis
  for (j in 1:J) {
    ranint[j, 1:4] ~ dmnorm(mu.ranint, Omega)
    ranslope.ck[j] ~ dnorm(0, tau.ranslope.ck)
  }

  # Model for each observational unit
  for (i in 1:N) {
    # X[, 1] contains time from diagnosis
    # X[, 2:nvar] all other predictors

    # CK
    mu.ck[i] <- alpha.ck + ranint[subj[i], 1] + (gamma.ck +
      ranslope.ck[subj[i]]) * X[i, 1] + inprod(X[i, 2:nvar], beta.ck)
    y.ck[i] ~ dnorm(mu.ck[i], tau.e.ck)
    censored[i] ~ dinterval(y.ck[i], cut[i])

    # CMAS
    probit(p.cmas[i]) <- zeta.cmas + ranint[subj[i], 2] +
      theta.cmas * X[i, 1]
    ind.cmas[i] ~ dbern(p.cmas[i])
    mu.cmas.temp[i] <- alpha.cmas + gamma.cmas * X[i, 1] +
      inprod(X[i, 2:nvar], beta.cmas)
    mu.cmas[i] <- ifelse(ind.cmas[i] == 0, sqrt(53), mu.cmas.temp[i])
    y.cmas[i] ~ dnorm(mu.cmas[i],
      tau.e.cmas[ind.cmas[i] + 1])T(0, sqrt(53))

    # MMT8
  }
}

```

```

probit(p.mmt[i]) <- zeta.mmt + ranint[subj[i], 3] +
  theta.mmt * X[i, 1]
ind.mmt[i] ~ dbern(p.mmt[i])
mu.mmt.temp[i] <- alpha.mmt + gamma.mmt * X[i, 1] +
  inprod(X[i, 2:nvar], beta.mmt)
mu.mmt[i] <- ifelse(ind.mmt[i] == 0, sqrt(80), mu.mmt.temp[i])
y.mmt[i] ~ dnorm(mu.mmt[i],
  tau.e.mmt[ind.mmt[i] + 1])T(0, sqrt(80))

# PGA
probit(p.pga[i]) <- zeta.pga + ranint[subj[i], 4] +
  theta.pga * X[i, 1]
ind.pga[i] ~ dbern(p.pga[i])
mu.pga.temp[i] <- alpha.pga + gamma.pga * X[i, 1] +
  inprod(X[i, 2:nvar], beta.pga)
mu.pga[i] <- ifelse(ind.pga[i] == 0, 0, mu.pga.temp[i])
y.pga[i] ~ dnorm(mu.pga[i],
  tau.e.pga[ind.pga[i] + 1])T(0, sqrt(10))
}

# Prior for Omega
# I is 4 x 4 identity matrix
Omega ~ dwish(R, 6)
R <- I * 60

# Priors for fixed intercepts and slopes
alpha.ck ~ dnorm(0, 0.001)
gamma.ck ~ dnorm(0, 0.001)
alpha.cmas ~ dnorm(0, 0.001)
gamma.cmas ~ dnorm(0, 0.001)
zeta.cmas ~ dnorm(0, 0.001)
theta.cmas ~ dnorm(0, 0.001)
alpha.mmt ~ dnorm(0, 0.001)
gamma.mmt ~ dnorm(0, 0.001)
zeta.mmt ~ dnorm(0, 0.001)
theta.mmt ~ dnorm(0, 0.001)
alpha.pga ~ dnorm(0, 0.001)
gamma.pga ~ dnorm(0, 0.001)
zeta.pga ~ dnorm(0, 0.001)
theta.pga ~ dnorm(0, 0.001)

for (i in 1:(nvar - 1)) {
  # Variable selection. Use same indicator for all four outcomes
  # (i.e. same variables are "in" and "out")
  ind[i] ~ dbern(pind[i])
  pind[i] ~ dbeta(0.1, 0.1)
  betaT.ck[i] ~ dnorm(0, taub)
}

```

```

betaT.cmas[i] ~ dnorm(0, taub)
betaT.mmt[i] ~ dnorm(0, taub)
betaT.pga[i] ~ dnorm(0, taub)
beta.ck[i] <- ind[i] * betaT.ck[i]
beta.cmas[i] <- ind[i] * betaT.cmas[i]
beta.mmt[i] <- ind[i] * betaT.mmt[i]
beta.pga[i] <- ind[i] * betaT.pga[i]
}

taub ~ dgamma(0.001, 0.001)

# Prior for tau of random slope
tau.ranslope.ck ~ dgamma(0.001, 0.001)

# Prior for error terms
# High precision in case the observation is estimated to be
# in the point mass. Non-informative prior otherwise.
tau.e.ck ~ dgamma(0.001, 0.001)
tau.e.cmas[1] <- 10000000
tau.e.cmas[2] ~ dgamma(0.001, 0.001)
tau.e.mmt[1] <- 10000000
tau.e.mmt[2] ~ dgamma(0.001, 0.001)
tau.e.pga[1] <- 10000000
tau.e.pga[2] ~ dgamma(0.001, 0.001)
}

```

ACKNOWLEDGEMENTS

The Juvenile Dermatomyositis Research Group would like to thank all of the patients and their families who contributed to the Juvenile Dermatomyositis Cohort & Biomarker Study & Repository. We thank all local research coordinators and principal investigators who have made this research possible. The members who contributed were as follows:

Dr Kate Armon, Mr Joe Ellis-Gage, Ms Holly Roper, Ms Vanja Briggs and Ms Joanna Watts (Norfolk and Norwich University Hospitals), Dr Liza McCann, Mr Ian Roberts, Dr Eileen Baildam, Ms Louise Hanna, Ms Olivia Lloyd and Susan Wadeson (The Royal Liverpool Childrens Hospital, Alder Hey, Liverpool), Dr Phil Riley and Ms Ann McGovern (Royal Manchester Childrens Hospital, Manchester), Dr Clive Ryder, Mrs. Janis Scott, Mrs. Beverley Thomas, Professor Taunton Southwood, Dr Eslam Al-Abadi (Birmingham Childrens Hospital, Birmingham), Dr Sue Wyatt, Mrs Gillian Jackson, Dr Tania Amin, Dr Mark Wood, Dr Vanessa VanRooyen and Ms Deborah Burton (Leeds General Infirmary, Leeds), Dr Joyce Davidson, Dr Janet Gardner-Medwin, Dr Neil Martin, Ms Sue Ferguson, Ms Liz Waxman and Mr Michael Browne (The Royal Hospital for Sick Children, Yorkhill, Glasgow), Dr Mark Friswell, Professor Helen Foster, Mrs Alison Swift, Dr Sharmila Jandial,

Ms Vicky Stevenson, Ms Debbie Wade, Dr Ethan Sen, Dr Eve Smith, Ms Lisa Qiao, Mr Stuart Watson and Ms Claire Duong (Great North Childrens Hospital, Newcastle), Dr Helen Venning, Dr Rangaraj Satyapal, Mrs Elizabeth Stretton, Ms Mary Jordan, Dr Ellen Mosley, Ms Anna Frost, Ms Lindsay Crate, Dr Kishore Warriar and Stefanie Stafford (Queens Medical Centre, Nottingham), Professor Lucy Wedderburn, Dr Clarissa Pilkington, Dr Nathan Hasson, Mrs Sue Maillard, Ms Elizabeth Halkon, Ms Virginia Brown, Ms Audrey Juggins, Dr Sally Smith, Mrs Sian Lunt, Ms Elli Enayat, Mrs Hemlata Varsani, Miss Laura Kassoumeri, Miss Laura Beard, Miss Katie Arnold, Mrs Yvonne Glackin, Ms Stephanie Simou, Dr Beverley Almeida, Dr Kiran Nistala, Dr Raquel Marques, Dr Claire Deakin, Ms Stefanie Dowle, Ms Charis Papadopoulou (Great Ormond Street Hospital, London), Dr Kevin Murray (Princess Margaret Hospital, Perth, Western Australia), Dr John Ioannou and Ms Linda Suffield (University College London Hospital, London), Dr Muthana Al-Obaidi, Ms Helen Lee, Ms Sam Leach, Ms Helen Smith, Dr Anne-Marie McMahon, Ms Heather Chisem and Ruth Kingshott (Sheffields Childrens Hospital, Sheffield); Dr Nick Wilkinson, Ms Emma Inness, Ms Eunice Kendall, Mr David Mayers, Ruth Etherton and Dr Kathryn Bailey (Oxford University Hospitals, Oxford); Dr Jacqui Clinch, Ms Natalie Fineman and Ms Helen Pluess-Hall (Bristol Royal Hospital for Children, Bristol); Ms Lindsay Vallance (Royal Aberdeen Childrens Hospital); Ms Louise Akeroyd (Bradford Teaching Hospitals); Dr Alice Leahy, Amy Collier, Rebecca Cutts, Dr Hans De Graaf, Dr Brian Davidson, Sarah Hartfree, Danny Pratt (University Hospital Southampton).

FUNDING

EHPvD received funding from the 7th Framework programme of the EU, SP3-People, support for training and career development for researchers (Marie Curie), Network for Initial Training (ITN), FP7-PEOPLE-2011-ITN, under the Marie Skłodowska-Curie grant agreement no. 289903. Funding for the UK JDM Cohort and Biomarker study has been provided by generous grants from the Wellcome Trust UK (085860), Action Medical Research UK (SP4252), The Myositis Support Group UK, Arthritis Research UK (14518, 20164), The Henry Smith Charity and Great Ormond Street Children's Charity (V1268), and the National Institute for Health Research (NIHR) Translational Research Collaboration (TRC) Rare Diseases. This research was supported by the NIHR Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and Institute of Child Health University College London (UCL). The JDM Cohort study is adopted onto the NIHR Comprehensive Research Network. The Arthritis Research UK Centre for Adolescent Rheumatology at UCL, UCL Hospital and GOSH is supported by grants from Arthritis Research UK (20164) and Great Ormond Street Children's Charity. This is a summary of independent research funded by the NIHRs Rare Diseases Translational Research

Collaboration. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

5

REFERENCE LIST

1. Ravelli A, Trail L, Ferrari C, Ruperto N, Pistorio A, Pilkington C et al. Long-term outcome and prognostic factors of juvenile dermatomyositis: a multinational, multicenter study of 490 patients. *Arthritis Care Res (Hoboken)* 2010; 62(1):63-72.
2. Nistala K, Wedderburn LR. Update in juvenile myositis. *Curr Opin Rheumatol* 2013; 25(6):742-746.
3. Rider LG, Katz JD, Jones OY. Developments in the classification and treatment of the juvenile idiopathic inflammatory myopathies. *Rheum Dis Clin North Am* 2013; 39(4):877-904.
4. Martin N, Krol P, Smith S, Murray K, Pilkington CA, Davidson JE et al. A national registry for juvenile dermatomyositis and other paediatric idiopathic inflammatory myopathies: 10 years' experience; the Juvenile Dermatomyositis National (UK and Ireland) Cohort Biomarker Study and Repository for Idiopathic Inflammatory Myopathies. *Rheumatology (Oxford)* 2011; 50(1):137-145.
5. Kim S, El-Hallak M, Dedeoglu F, Zurakowski D, Fuhlbrigge RC, Sundel RP. Complete and sustained remission of juvenile dermatomyositis resulting from aggressive treatment. *Arthritis Rheum* 2009; 60(6):1825-1830.
6. Shah M, Mamyrova G, Targoff IN, Huber AM, Malley JD, Rice MM et al. The clinical phenotypes of the juvenile idiopathic inflammatory myopathies. *Medicine (Baltimore)* 2013; 92(1):25-41.
7. Stringer E, Singh-Grewal D, Feldman BM. Predicting the course of juvenile dermatomyositis: significance of early clinical and laboratory features. *Arthritis Rheum* 2008; 58(11):3585-3592.
8. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med* 1975; 292(7):344-347.
9. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). *N Engl J Med* 1975; 292(8):403-407.
10. Lazarevic D, Pistorio A, Palmisani E, Miettunen P, Ravelli A, Pilkington C et al. The PRINTO criteria for clinically inactive disease in juvenile dermatomyositis. *Ann Rheum Dis* 2013; 72(5):686-693.
11. Lovell DJ, Lindsley CB, Rennebohm RM, Ballinger SH, Bowyer SL, Giannini EH et al. Development of validated disease activity and damage indices for the juvenile idiopathic inflammatory myopathies. II. The Childhood Myositis Assessment Scale (CMAS): a quantitative tool for the evaluation of muscle function. The Juvenile Dermatomyositis Disease Activity Collaborative Study Group. *Arthritis Rheum* 1999; 42(10):2213-2219.
12. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw* 2015; 67(1):1-48.
13. R core team. R: A language and environment for statistical computing. www R-project org [2015 [cited 2016];
14. Mullahy J. Specification and testing of some modified count data models. *J Econom* 1986; 33:341-365.
15. Li Q, Pan J, Belcher J. Bayesian inference for joint modelling of longitudinal continuous, binary and ordinal events. *Stat Methods Med Res* 2016; 25(6):2521-2540.
16. Kuo L, Mallick B. Variable selection for regression models. *Sankhya Ser B* 1998; 60:65-81.
17. O'Hara RB, Sillanpää MJ. A review of Bayesian variable selection methods: What, how and which. *Bayesian Anal* 2009; 4(1):85-117.
18. Plummer M. rjags: Bayesian Graphical Models using MCMC. R package version 4-6. <https://CRAN.R-project.org/package=rjags> [2016 [cited 2017 Mar. 17];
19. Brier GW. Verification of forecasts expressed in terms of probability. *Month Weather Rev* 1950; 78(1):1-3.
20. Steyerberg EW, Vickers AJ, Cook NR, Gerdts T, Gonen M, Obuchowski N et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010; 21(1):128-138.
21. Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Series B Stat Methodol* 2002; 64(4):583-639.

Chapter

6

CLINICAL SIGNS AND SYMPTOMS IN
A JOINT MODEL OF FOUR DISEASE
ACTIVITY PARAMETERS IN
JUVENILE DERMATOMYOSITIS:
A PROSPECTIVE, LONGITUDINAL,
MULTICENTER COHORT STUDY

E.H. Pieter van Dijkhuizen, Claire T. Deakin, Maria De Iorio
and Lucy R. Wedderburn on behalf of the JDRG

Submitted

ABSTRACT

Objective

To find clinical features most strongly associated with outcome variables in juvenile dermatomyositis (JDM) as a first step towards tailor-made treatment.

Methods

In a large, prospectively followed, multicenter cohort study of 340 JDM patients contributing multiple visits each, a Bayesian model of disease activity was developed, using the four continuous outcome variables creatine kinase (CK), childhood myositis assessment score (CMAS), manual muscle testing of 8 muscle groups (MMT8) and the physician's global assessment of disease activity (PGA). Covariates were clinical signs and symptoms. Dependencies among visits of the same patient were resolved by introducing subject-specific random effects.

Results

Myalgia and dysphonia were associated with worse disease activity according to all outcome variables. Periorbital rash, rash on the trunk, rash over large joints, nail fold changes and facial swelling were associated with higher PGA. Notably, periorbital rash was also associated with higher CK and lower CMAS and nail fold changes with lower CMAS. Contractures were associated with lower CMAS and MMT8 and higher PGA. Patients with higher CMAS exhibited a higher MMT8 as well. PGA had the highest probability among the four outcome variables of being abnormal even if the other three outcome variables were normal.

Conclusion

The signs and symptoms associated with disease activity could be used to stratify patients and adapt treatment plans to disease activity. The correlation between CMAS and MMT8 and the unique information captured by PGA implied that PGA should be maintained as an outcome variables, whereas CMAS and MMT8 might be simplified.

INTRODUCTION

The childhood inflammatory idiopathic myopathies (IIM) are a group of heterogeneous disorders, characterized by chronic skeletal muscle inflammation. Of these, juvenile dermatomyositis (JDM) is the most common, though still a rare disease with an incidence of about 1.9-4.1 per 1,000,000 children.¹ Its hallmark is muscle inflammation characterized by proximal muscle weakness, in concert with skin involvement presenting itself typically as Gottron's papules, heliotrope rash, malar rash or erythema overlying the extensor surfaces of the joints.² JDM potentially involves other internal organ systems as well, most notably the gastrointestinal and the respiratory tracts and cases with major organ involvement have a poor prognosis.² JDM is heterogeneous in nature, in terms of disease severity and various patterns of involvement of muscle, skin and internal organ systems. About 24-40% of patients experience a monocyclic course, whereas 50-60% have chronic disease activity.² The mortality rate is around 2-3%.²

Given the severity and burden of the disease and possible long-term complications, adequate treatment is of utmost importance.^{2,3} A recent trial has shown that the combination of prednisone and methotrexate has the best potential to induce disease remission in new-onset JDM.⁴ Ideally, treatment is tailored to the patient, in such a way that patients with high disease activity, at risk to develop serious sequelae of the disease, receive early and aggressive treatment, whereas those with less severe forms of the disease get milder therapy. Previous studies have revealed some clinical factors associated with a worse prognosis.^{1,5,6} However, in these reports disease activity was either taken dichotomously at a single point in time, or analyzed as time to remission.

The aim of the current study was to find clinical signs and symptoms associated with higher disease activity as measured by four widely-used continuous outcome variables, being the serum level of creatine kinase (CK), the childhood myositis assessment scale (CMAS), manual muscle testing of 8 muscle groups (MMT8) and the physician's global assessment of disease activity (PGA), assessed longitudinally in a large, multicenter cohort of JDM patients. Such associations with high disease activity could be used in follow up studies to predict disease outcome in JDM patients.

PATIENTS AND METHODS

Patients were retrieved from the ongoing UK Juvenile Dermatomyositis Cohort and Biomarker Study, which started recruitment of JDM cases across the UK in 2000.⁷ Patients were enrolled at diagnosis or shortly thereafter, and followed up approximately every 3 months for 2 years and subsequently at least annually. At each visit, data were collected regarding signs and symptoms of the disease, such as skin manifestations, muscular involvement and symptoms of involvement of other organ systems (see supplementary

data for full list). Blood was drawn for routine laboratory testing. Furthermore, data was collected regarding treatment and disease activity according to the four selected outcome variables, i.e., CK, CMAS,⁸ MMT8 and PGA. These outcomes were selected because they have been validated as a set of parameters able to classify JDM patients as active or inactive.⁹ Furthermore, they are widely used and readily available in routine daily care. However, rather than applying previously published criteria for inactive disease and dichotomizing patients as active or inactive, we modelled the four parameters in a continuous way, thus taking full advantage of the information they contain.

Ethical approval was obtained by the multicenter ethical review board covering all participating institutions. All participants provided written informed consent, or age-appropriate assent with parental consent. The study was performed according to the declaration of Helsinki and good clinical practice guidelines.

At the time of analysis 4,122 visits of 469 patients were available. Of these, 413 patients contributing 3,881 visits met the inclusion criteria of having probable or definite JDM according to Bohan and Peter criteria.^{10;11}

Statistical analysis

The data were analyzed using a Bayesian approach. Details of the analysis have been described elsewhere.¹² Briefly, a mixed effect regression model was fitted, by specifying a joint model for the four clinical outcome variables. These outcome parameters were modelled as continuous variables. This approach allowed us to account for the correlation among disease activity measures and, therefore, better exploit the information contained in the variables, compared to an analysis which considers the four outcomes as independent and treats the responses as dichotomous. In the analysis, CK values were log-transformed so that the distribution of CK was closer to a normal. On the other hand, CMAS, MMT8 and PGA were square root transformed, as they are non-negative variables potentially assuming value zero. More interestingly, these three variables showed an excess of the best possible clinical value for that parameter for visits in disease remission and a long tail towards the pathological end of the scale. These distributional characteristics must be accounted for in the analysis to avoid bias in the estimates. As such we model CMAS, MMT8 and PGA using an approach similar to hurdle models,¹³ which assumes a mixture of a point mass at the non-pathological values and a continuous distribution. This implies that the outcome variable will be equal to the non-pathological value with a certain probability p , which is also object of inference, and with probability $1-p$ is a draw from a continuous distribution, in our case a truncated normal distribution. We model p , the probability of the non-pathological value, using probit regression. This step is therefore similar to a regression model with binary outcome data. For the continuous component of the mixture, we assume that the mean of the truncated normal distribution is a linear function of the covariates, as in standard linear regression.

This approach allowed the model to estimate many more visits to be in disease remission than a standard linear regression would do, thus mirroring the observed distribution of the outcome parameters.

To make maximal use of the information contained in the dataset, all visits of all patients were analyzed simultaneously. Temporal correlation between multiple visits of the same patient was accounted for using a subject-specific random intercept and, in the case of CK level, a subject-specific random slope for the time since diagnosis, in a way similar to a mixed model in a repeated measurements analysis.¹⁴ Furthermore, due to correlations between the four outcome parameters, the four subject-specific random intercepts (one for each outcome variable) were modelled by specifying a multivariate normal distribution as random effect distribution. The covariance matrix of this distribution was estimated from the data.¹⁵ This allowed the assessment of the correlation among the outcome parameters (i.e., the correlation which measures whether one outcome parameter tended to increase if another parameter increased as well within the same individual over time). Missing values in the covariates were imputed in the Bayesian model, by specifying an appropriate model for them. We did not impute missing values for the history variables, since this would amount to imputing a participant's recollection, the accuracy of which is doubtful. Visits with any unobserved history variables were therefore excluded from the analysis.

The aim of the model was to find all clinical features associated with disease activity. Therefore, all clinical signs and symptoms and treatment variables, as well as the time elapsed since diagnosis were eligible to be included as independent variables in the model. A variable pre-selection was performed by selecting 50% of the variables using univariate linear mixed models. All treatment covariates and the time elapsed since diagnosis were included in the model, regardless of their performance in the univariate analysis. Bayesian variable selection on the pre-selected covariates was performed.¹⁶ This approach yields the sparsest set of predictors that is still able to estimate the outcome parameters accurately. A predictor variable would either be included for all four outcomes variables, or excluded.

The goodness of fit of the model was assessed visually by checking the fit of the model for patients with more than 10 visits, by looking at 95% Bayesian credible intervals (CI) for the disease trajectory over time. Furthermore, the ability of the model to predict values above or below the validated cut-off points (CK ≤ 150 , CMAS ≥ 48 , MMT8 ≥ 78 and PGA ≤ 0.2),⁹ was assessed by calculating the scaled Brier score. This score has a range from 0 to 100% and has an interpretation similar to Pearson's R^2 statistic.¹⁷ The out of sample prediction ability of the model was tested by randomly selecting five fully observed patients and leaving them out while fitting the model. The predicted values for the four outcomes over time were then compared to the observed values. Statistical analysis was performed in R 3.2.2 (R foundation for statistical computing, Vienna, Austria) and JAGS, using the package rjags.¹⁸

RESULTS

Due to exclusion of visits with missing data in history variables, 340 of 413 patients were included in the analysis. Of these, the majority (69.4%) was female, the median age at diagnosis was 7.4 years (1st quartile-3rd quartile 4.5-10.5) and disease duration since the onset of the first symptoms was short (median of 0.3 years, 1st quartile-3rd quartile 0.2-0.6). Patients contributing more visits to the study were more likely to have at least one visit without missing values in the history variables and were therefore more likely to be included. Excluded patients had shorter duration of follow up in the study, longer period of time after diagnosis before enrolment, and less active disease (Table 1). The proportion of missing data in the analyzed data set was 7.3%.

The goodness of fit of the model was good as evidenced by a plot of the observed values versus the predicted values and 95% Bayesian CI (Figure 1). Over all visits and outcome parameters, only 1.8% of observed values were outside of the CIs. The scaled Brier score was 49%, 42%, 63% and 80% of the maximally obtainable Brier score for CK, CMAS, MMT8 and PGA, respectively. The visual check of the out of sample predictions showed accurate predictions of the four outcomes. Only in two cases involving CMAS the predicted value was above the cut-off point for inactive disease (48 points), whereas the observed value was below, but in both these cases the observed value was close to the cut-off point (46 and 47 respectively). The precision of the predictions was modest as evidenced by wide prediction CIs, owing to uncertainties in predictions, parameter estimation and missing value imputation.

Table 1. Baseline table

Parameter	Included N = 340	Excluded N = 73
Female, n (%)	236 (69.4)	54 (74.0)
Age at diagnosis, y	7.4 [4.5, 10.5]	7.3 [4.1, 11.1]
Disease duration at diagnosis, y	0.3 [0.2, 0.6]	0.3 [0.2, 1.0]
Time after diagnosis at enrolment, y	0.2 [0.1, 1.1]	2.3 [0.4, 5.4]
Duration of follow up, y	4.1 [1.6, 7.1]	1.2 [0.1, 2.6]
Disease activity at enrolment:		
CK, U/L	103 [64, 440]	98 [45, 256]
CMAS, points	41 [21, 50]	46 [37, 52]
MMT8, points	65 [45, 80]	80 [64, 80]
PGA, cm	3 [1.3, 6.0]	2.3 [0.5, 4.0]

Values are the median [1st quartile, 3rd quartile], except where indicated otherwise.

Abbreviations: CK, creatine kinase; cm, centimeter; CMAS, childhood myositis assessment scale; MMT8, manual muscle testing of 8 muscle groups; PGA, physician's global assessment of disease activity; U/L, units per liter; y, year.

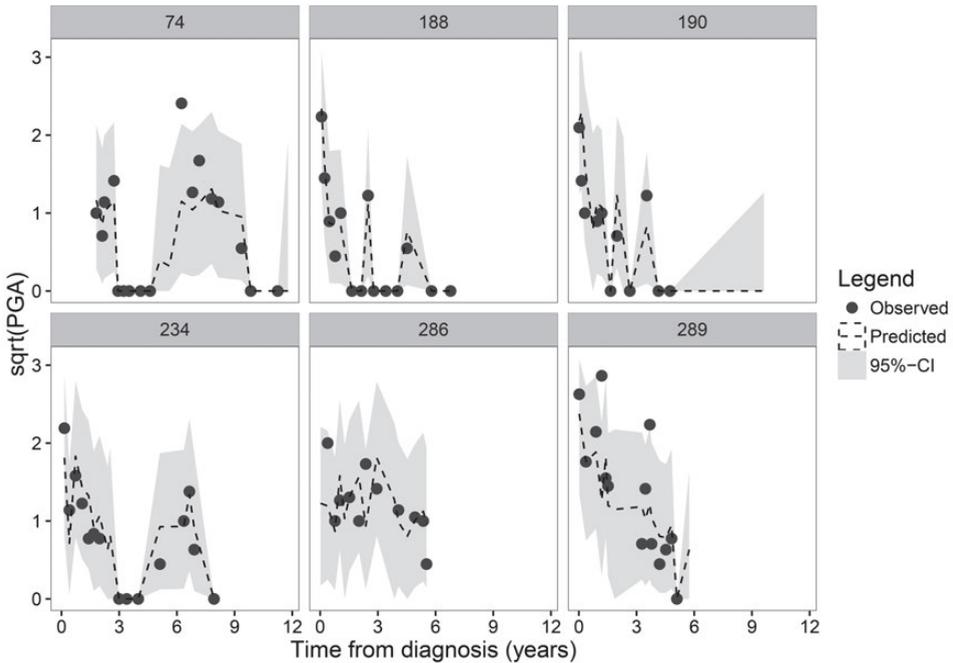


Figure 1. Goodness of fit of the physician's global assessment of disease activity. Observed values of the (square root transformed) parameter (dark grey dots) for six randomly selected individuals were plotted against the predicted values by the model (dashed line) and the 95% CI (light grey area), showing that the predicted values corresponded well to observed patterns over time. Goodness of fit of the other outcome parameters was similar (not shown). **Abbreviations:** CI, credible interval; sqrt, square root.

Estimates of regression coefficients of the continuous component of the model showed that CMAS, MMT8 and PGA tended to normalize over time, whereas hardly any influence of time on CK was noted (Figure 2). Many muscular symptoms, such as myalgia and dysphonia were associated with higher disease activity. Cutaneous symptoms such as periorbital rash, rash on the trunk, rash over large joints, nail fold changes and facial swelling were associated with higher disease activity and, notably, were in some cases associated with more disease activity according to strictly muscular outcome parameters (e.g., periorbital rash was associated with higher CK values and lower CMAS and nail fold changes was associated with lower CMAS).

The presence of contractures was associated with lower CMAS and MMT8 values and higher PGA, whereas CK was not affected. An association was found between calcinosis and higher CMAS and PGA values. Joint swelling was associated with lower CMAS, lower CK and higher PGA and the presence of hematuria was accompanied by markedly elevated CK levels. Of all signs and symptoms at baseline, only arthritis was associated with increased CMAS and MMT8.

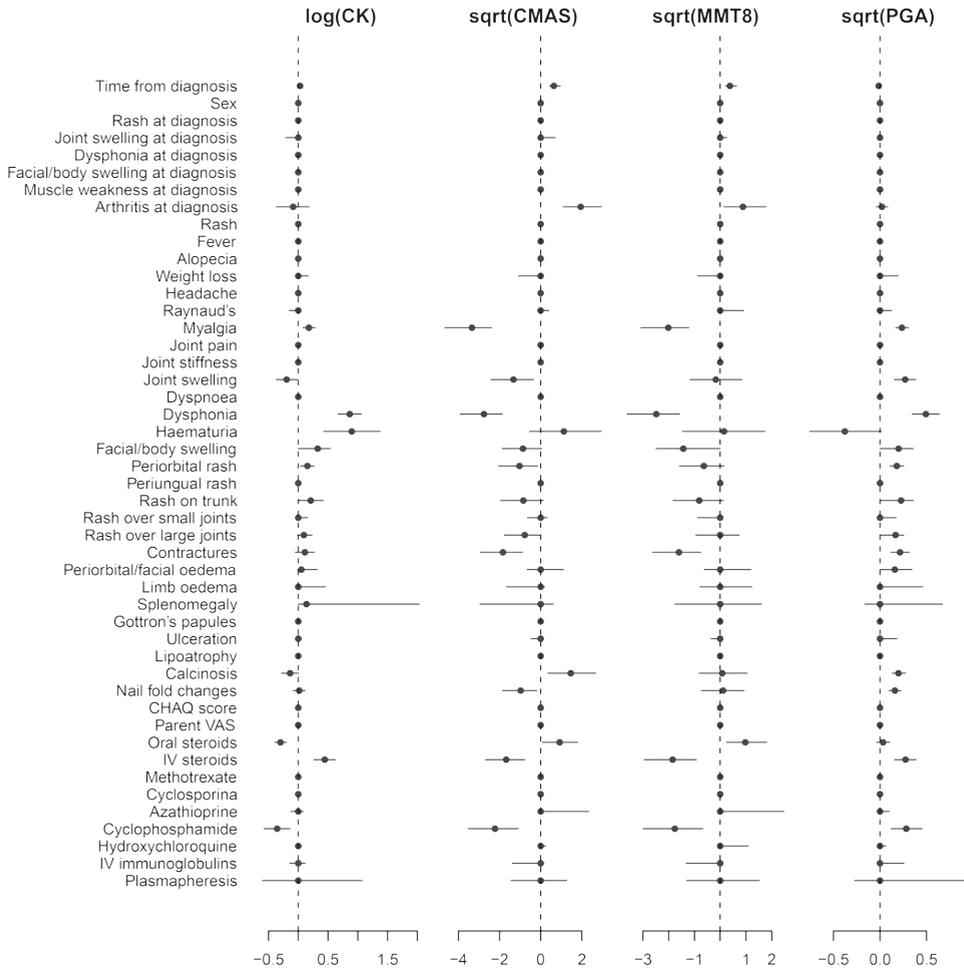


Figure 2. Regression coefficients (dots) with 95% credible interval (horizontal lines) of the fixed effects of time elapsed since diagnosis and all covariates in the model, for all outcomes. Regression coefficients to the left of the vertical dashed line indicate the parameter is associated with lower values of the corresponding outcome measurement, and conversely for regression coefficients to the right of the dashed line. Some credible intervals are collapsed to 0 (e.g., association between sex and log(CK)), due to there not being any association at all.

The use of both cyclophosphamide and IV steroids was associated with higher disease activity, though cyclophosphamide was also associated with lower CK levels. Conversely, oral steroids were associated with lower disease activity.

Correlations between the subject-specific random intercepts showed that patients with a higher CMAS tended to have a higher MMT8 as well (correlation coefficient 0.54, 95% Bayesian CI 0.42-0.65). A weak correlation was found between MMT8 and PGA (0.23, 95% CI 0.08-0.38) and CMAS and PGA (0.17, 95% CI 0.01-0.32). All correlations

between the subject-specific random intercept for CK and the other three parameters were low (<0.07).

Of visits during which three outcome variables showed normal values, the median (1st, 3rd quartile) probability that the predicted value for PGA remained abnormal (i.e. > 0.2) was estimated to be 52% (<1 , 78%). The median (1st, 3rd quartile) probability was 6% (2, 20%), 1% (<1 , 8%) and 1% (<1 , 17%), for CK, CMAS and MMT8 respectively.

DISCUSSION

This study identified clinical signs and symptoms associated with four outcome parameters taken continuously and longitudinally in a large, multicenter cohort of JDM patients. This approach not only allowed to estimate the associations of the various signs and symptoms with disease activity, but also enabled to assess correlations among the outcome parameters, in order to have a clearer understanding of disease activity by accounting jointly for all the outcome values. The results showed as expected that dysphonia, already known to be a marker of severe disease activity,² was associated with higher disease activity. Hematuria was also associated with markedly elevated CK levels. Hematuria was measured by urine dipstick, which gives a positive result in the case of hematuria or myoglobinuria. In the former case, hematuria is an indication of severe systemic (i.e., renal) disease. In the latter case, myoglobinuria is due to severe muscular involvement, leading to rhabdomyolysis.

Cutaneous symptoms were associated with PGA, the only parameter taking account of cutaneous disease activity, but, interestingly, periorbital rash and nail fold changes were also associated with lower CMAS values and higher CK values, implying that cutaneous and muscular disease were correlated. It could therefore be hypothesized that the return of cutaneous symptoms in a patient in disease remission signals an imminent muscular relapse. This finding lends support to the expert-based opinion that ongoing skin disease reflects ongoing systemic disease activity.¹⁹ Interestingly, Gottron's papules, pathognomonic for JDM, were not associated with any outcome parameter.

The association between contractures and lower CMAS and MMT8 values might be due to difficulty in performing these instruments in the presence of contractures,²⁰ especially given the absence of association with CK. Likewise, the association between calcinosis and higher CMAS levels might indicate that this phenomenon occurred in a late stage of disease and persisted in disease remission. Yet, the results showed that physicians gave higher PGA scores to patients with calcinosis.

Arthritis at diagnosis was associated with increased CMAS and MMT8, potentially due to a subset of patients with an overlap between JDM and juvenile idiopathic arthritis (JIA) with less severe muscular inflammation.² Joint swelling at the visit was associated

with a higher PGA and lower CMAS, possibly indicating difficulty in executing the tasks of the CMAS in the presence of arthritis.

The substantial correlation between the subject-specific random intercepts for CMAS and MMT8 indicated that patients with a higher CMAS tended to have a higher MMT8, as reported previously.²⁰ This implied that information captured by CMAS was also partially captured by MMT8 and vice versa. Conversely, in many visits where CK, CMAS and MMT8 were normal, the model estimated PGA to be abnormal, implying that PGA captures aspects of disease that are not measured by the other outcome parameters. This was consistent with the observation that PGA was the only outcome parameter taking account of cutaneous and other organ involvement.^{21,22} Similar estimates for the other outcome parameters in our model were much lower, indicating that these variables captured information already conveyed by the other outcome measures. In conclusion, our observations supported the previously made proposal to increase the weight of PGA in the evaluation of JDM disease activity, to account for cutaneous disease activity.²² Moreover, our results suggested that the three muscular disease activity measures, mainly CMAS and MMT8, could be shortened or even summarized into a single instrument, thus saving precious time during busy clinics, whilst retaining the same level of information.

The results of our study were in line with results obtained previously.^{1;5;6} In a large cohort of 490 JDM patients, analyzing the dichotomized outcome parameter at one time point, a mean of 7.7 years after diagnosis, an association of CMAS with dysphagia and dysphonia was found.¹ This study also found an increased probability of having a CMAS score of 52 points for patients with cutaneous symptoms at onset, whereas in our study cutaneous manifestations (though not at onset) were associated with a lower CMAS score.¹ In a Canadian cohort of 84 patients, the persistence of skin rash, especially Gottron's papules, three months after diagnosis was associated with a longer time to remission.⁵ However, given that disease remission was defined in this study as absence of skin rash (including Gottron's papules), myositis and arthritis, this finding might be tautological.⁵ Gottron's papules were not associated with any of the four outcome parameters in our study. In a retrospective study of 61 JDM patients, a lower skin disease activity score (DAS) at baseline was also associated with a monocyclic disease course.⁶

Next to clinical factors, the antibody patterns in the childhood IIM have been studied extensively and were found to determine different subsets of the disease, characterized among others by the distribution of skin rash, contractures, dysphonia, dysphagia, and the outcome of the disease.^{23;24} Furthermore, prediction of outcome could be aided by muscle biopsy findings.^{25;26}

The goal of our study was to find clinical signs and symptoms associated with four frequently used disease activity parameters.^{9;22} As a consequence, no other outcome parameters, mainly cutaneous disease activity measurements, were considered. Future work may amend this, however, this would entail that cutaneous symptoms that form part of the outcome measurement are not available as predictors anymore.

Hardly any features at baseline were associated with disease activity at follow up. This was most probably because the model included also signs and symptoms during follow up, the associations of which overwhelmed associations with baseline variables. Future work may investigate the predictive ability of these signs and symptoms in the assessment of disease activity.

In conclusion, the associations between clinical signs and symptoms and four continuous disease activity measurements in this large, multicenter cohort of JDM patients who were followed longitudinally open up possibilities to personalize treatment plans in JDM, by offering more aggressive treatment to patients with signs and symptoms associated with higher disease activity. Follow up studies may attempt to predict future disease activity using the associations found in the current study. Furthermore, the associations highlight interesting patterns, such as the association between skin disease and muscular outcome measures. Finally, the correlations between the subject-specific random effects for each of the four outcome measurements and the probabilities for each outcome parameter of having an abnormal value during visits where the other three parameters were normal, as estimated by the model, provided insight into unique information captured by each parameter and might be helpful in the determination of a parsimonious set of outcome parameters in JDM, for example by summarizing the CMAS and MMT8 in a single instrument and increasing the importance of PGA in the evaluation of disease remission.

ACKNOWLEDGEMENTS

The Juvenile Dermatomyositis Research Group would like to thank all of the patients and their families who contributed to the Juvenile Dermatomyositis Cohort & Biomarker Study & Repository. We thank all local research coordinators and principal investigators who have made this research possible. The JDRG members were as follows:

Dr Kate Armon, Mr Joe Ellis-Gage, Ms Holly Roper, Ms Vanja Briggs and Ms Joanna Watts (Norfolk and Norwich University Hospitals), Dr Liza McCann, Mr Ian Roberts, Dr Eileen Baidam, Ms Louise Hanna, Ms Olivia Lloyd and Susan Wadeson (The Royal Liverpool Children's Hospital, Alder Hey, Liverpool), Dr Phil Riley and Ms Ann McGovern (Royal Manchester Children's Hospital, Manchester), Dr Clive Ryder, Mrs. Janis Scott, Mrs. Beverley Thomas, Professor Taunton Southwood, Dr Eslam Al-Abadi (Birmingham Children's Hospital, Birmingham), Dr Sue Wyatt, Mrs Gillian Jackson, Dr Tania Amin, Dr Mark Wood, Dr Tania Amin, Dr Vanessa VanRooyen and Ms Deborah Burton (Leeds General Infirmary, Leeds), Dr Joyce Davidson, Dr Janet Gardner-Medwin, Dr Neil Martin, Ms Sue Ferguson, Ms Liz Waxman and Mr Michael Browne (The Royal Hospital for Sick Children, Yorkhill, Glasgow), Dr Mark Friswell, Professor Helen Foster, Mrs Alison Swift, Dr Sharmila Jandial, Ms Vicky Stevenson, Ms Debbie Wade, Dr Ethan Sen, Dr Eve Smith, Ms Lisa Qiao, Mr Stuart Watson and Ms Claire Duong (Great North Children's Hospital,

Newcastle), Dr Helen Venning, Dr Rangaraj Satyapal, Mrs Elizabeth Stretton, Ms Mary Jordan, Dr Ellen Mosley, Ms Anna Frost, Ms Lindsay Crate, Dr Kishore Warriar and Stefanie Stafford (Queens Medical Centre, Nottingham), Professor Lucy Wedderburn, Dr Clarissa Pilkington, Dr Nathan Hasson, Mrs Sue Maillard, Ms Elizabeth Halkon, Ms Virginia Brown, Ms Audrey Juggins, Dr Sally Smith, Mrs Sian Lunt, Ms Elli Enayat, Mrs Hemlata Varsani, Miss Laura Kassoumeri, Miss Laura Beard, Miss Katie Arnold, Mrs Yvonne Glackin, Ms Stephanie Simou, Dr Beverley Almeida, Dr Kiran Nistala, Dr Raquel Marques, Dr Claire Deakin, Ms Stefanie Dowle, Dr Charalampia Papadopoulou, Dr Shireena Yasin, Mrs Cerise Johnson-Moore, Ms Emily Robinson (Great Ormond Street Hospital, London), Dr Kevin Murray (Princess Margaret Hospital, Perth, Western Australia) Dr John Ioannou and Ms Linda Suffield (University College London Hospital, London) Dr Muthana Al-Obaidi, Ms Helen Lee, Ms Sam Leach, Ms Helen Smith, Dr Anne-Marie McMahon, Ms Heather Chisem and Ruth Kingshott (Sheffield's Children's Hospital, Sheffield); Dr Nick Wilkinson, Ms Emma Inness, Ms Eunice Kendall, Mr David Mayers, Ruth Etherton, Danielle Miller and Dr Kathryn Bailey (Oxford University Hospitals, Oxford); Dr Jacqui Clinch, Ms Natalie Fineman and Ms Helen Pluess-Hall (Bristol Royal Hospital for Children, Bristol); Ms Lindsay Vallance (Royal Aberdeen Children's Hospital); Ms Louise Akeroyd (Bradford Teaching Hospitals); Dr Alice Leahy, Amy Collier, Rebecca Cutts, Emma Macleod, Dr Hans De Graaf, Dr Brian Davidson, Sarah Hartfree, Danny Pratt (University Hospital Southampton)

FUNDING

EHPvD received funding from the 7th Framework programme of the EU, SP3-People, support for training and career development for researchers (Marie Curie), Network for Initial Training (ITN), FP7-PEOPLE-2011-ITN, under the Marie Skłodowska-Curie grant agreement No 289903. CTD was supported by a fellowship from The Myositis Association (TMA). Funding for the UK JDM Cohort and Biomarker study has been provided by generous grants from the Wellcome Trust UK [085860], Action Medical Research UK [SP4252], The Myositis Support Group UK, Arthritis Research UK [14518, 20164], The Henry Smith Charity and Great Ormond Street Children's Charity [V1268], The Myositis Association, The Medical Research Council and the National Institute for Health Research (NIHR) Translational Research Collaboration (TRC) Rare Diseases. This research was supported by the NIHR Great Ormond Street Hospital for Children NHS Foundation Trust Biomedical Research Centre (NIHR-GOSH BRC). The JDM Cohort study is adopted onto the NIHR Comprehensive Research Network. The Arthritis Research UK Centre for Adolescent Rheumatology at UCL, UCL Hospital and GOSH is supported by grants from Arthritis Research UK [20164] and Great Ormond Street Children's Charity. This is a summary of independent research funded by the NIHR's Rare Diseases Translational Research Collaboration. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

SUPPLEMENTARY MATERIAL

List of available covariate data

Demographics

Sex, age at diagnosis.

History at diagnosis

Rash, weakness, myalgia, dyspnoea, dysphonia, dysphagia, joint pain, joint stiffness, joint swelling, fever, alopecia, weight loss, fatigue, mouth ulcers, headache, irritability, chest pain, abdominal pain, diarrhoea, melena, haematuria, facial or body swelling.

Physical examination at diagnosis

Weakness, arthritis, Gottron's papules, ulceration, oedema.

Symptoms at visit

Rash, weakness, myalgia, dyspnoea, dysphonia, dysphagia, joint pain, joint stiffness, joint swelling, Raynaud's phenomenon, fever, alopecia, weight loss, fatigue, mouth ulcers, headache, irritability, chest pain, abdominal pain, diarrhoea, melena, haematuria, facial or body swelling.

Physical examination at visit

Periorbital rash, periungual rash, rash on trunk, rash over small joints, rash over large joints, Gottron's papules, nail fold changes, ulceration, calcinosis, lipoatrophy, arthritis, contractures, oedema, periorbital or facial oedema, limb oedema, trunk oedema, hepatomegaly, splenomegaly.

Measurements at visit

Childhood health assessment questionnaire, parent global assessment of well-being, parent assessment of pain.

Treatment

Methotrexate, oral steroids, intravenous (IV) steroids, cyclosporine, azathioprine, cyclophosphamide, hydroxychloroquine, IV immunoglobulins, plasmapheresis.

REFERENCE LIST

1. Ravelli A, Trail L, Ferrari C, Ruperto N, Pistorio A, Pilkington C et al. Long-term outcome and prognostic factors of juvenile dermatomyositis: a multinational, multicenter study of 490 patients. *Arthritis Care Res (Hoboken)* 2010; 62(1):63-72.
2. Rider LG, Katz JD, Jones OY. Developments in the classification and treatment of the juvenile idiopathic inflammatory myopathies. *Rheum Dis Clin North Am* 2013; 39(4):877-904.
3. Kim S, El-Hallak M, Dedeoglu F, Zurakowski D, Fuhlbrigge RC, Sundel RP. Complete and sustained remission of juvenile dermatomyositis resulting from aggressive treatment. *Arthritis Rheum* 2009; 60(6):1825-1830.
4. Ruperto N, Pistorio A, Oliveira S, Zulian F, Cuttica R, Ravelli A et al. Prednisone versus prednisone plus ciclosporin versus prednisone plus methotrexate in new-onset juvenile dermatomyositis: a randomised trial. *Lancet* 2016; 387(10019):671-678.
5. Stringer E, Singh-Grewal D, Feldman BM. Predicting the course of juvenile dermatomyositis: significance of early clinical and laboratory features. *Arthritis Rheum* 2008; 58(11):3585-3592.
6. Christen-Zaech S, Seshadri R, Sundberg J, Paller AS, Pachman LM. Persistent association of nailfold capillaroscopy changes and skin involvement over thirty-six months with duration of untreated disease in patients with juvenile dermatomyositis. *Arthritis Rheum* 2008; 58(2):571-576.
7. Martin N, Krol P, Smith S, Murray K, Pilkington CA, Davidson JE et al. A national registry for juvenile dermatomyositis and other paediatric idiopathic inflammatory myopathies: 10 years' experience; the Juvenile Dermatomyositis National (UK and Ireland) Cohort Biomarker Study and Repository for Idiopathic Inflammatory Myopathies. *Rheumatology (Oxford)* 2011; 50(1):137-145.
8. Lovell DJ, Lindsley CB, Rennebohm RM, Ballinger SH, Bowyer SL, Giannini EH et al. Development of validated disease activity and damage indices for the juvenile idiopathic inflammatory myopathies. II. The Childhood Myositis Assessment Scale (CMAS): a quantitative tool for the evaluation of muscle function. The Juvenile Dermatomyositis Disease Activity Collaborative Study Group. *Arthritis Rheum* 1999; 42(10):2213-2219.
9. Lazarevic D, Pistorio A, Palmisani E, Miettunen P, Ravelli A, Pilkington C et al. The PRINTO criteria for clinically inactive disease in juvenile dermatomyositis. *Ann Rheum Dis* 2013; 72(5):686-693.
10. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). *N Engl J Med* 1975; 292(8):403-407.
11. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med* 1975; 292(7):344-347.
12. Van Dijkhuizen EHP, Deakin CT, Wedderburn LR, De Iorio M. Modelling disease activity in juvenile dermatomyositis: A Bayesian approach. *Stat Methods Med Res* 2017;962280217713233.
13. Mullahy J. Specification and testing of some modified count data models. *J Econom* 1986; 33:341-365.
14. Pinheiro JC, Bates DM. Mixed-effects models in S and S-PLUS. New York: Springer-Verlag; 2000.
15. Li Q, Pan J, Belcher J. Bayesian inference for joint modelling of longitudinal continuous, binary and ordinal events. *Stat Methods Med Res* 2016; 25(6):2521-2540.
16. Kuo L, Mallick B. Variable selection for regression models. *Sankhya Ser B* 1998; 60:65-81.
17. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010; 21(1):128-138.
18. Plummer M. rjags: Bayesian Graphical Models using MCMC. R package version 4-6. <https://CRAN.R-project.org/package=rjags> [2016 [cited 2017 Mar. 17];

19. Bellutti Enders F, Bader-Meunier B, Baildam E, Constantin T, Dolezalova P, Feldman BM et al. Consensus-based recommendations for the management of juvenile dermatomyositis. *Ann Rheum Dis* 2017; 76(2):329-340.
20. Rider LG, Werth VP, Huber AM, Alexanderson H, Rao AP, Ruperto N et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis: Physician and Patient/Parent Global Activity, Manual Muscle Testing (MMT), Health Assessment Questionnaire (HAQ)/Childhood Health Assessment Questionnaire (C-HAQ), Childhood Myositis Assessment Scale (CMAS), Myositis Disease Activity Assessment Tool (MDAAT), Disease Activity Score (DAS), Short Form 36 (SF-36), Child Health Questionnaire (CHQ), physician global damage, Myositis Damage Index (MDI), Quantitative Muscle Testing (QMT), Myositis Functional Index-2 (FI-2), Myositis Activities Profile (MAP), Inclusion Body Myositis Functional Rating Scale (IBMFRS), Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI), Cutaneous Assessment Tool (CAT), Dermatomyositis Skin Severity Index (DSSI), Skindex, and Dermatology Life Quality Index (DLQI). *Arthritis Care Res (Hoboken)* 2011; 63 Suppl 11:S118-S157.
21. Nistala K, Wedderburn LR. Update in juvenile myositis. *Curr Opin Rheumatol* 2013; 25(6):742-746.
22. Almeida B, Campanilho-Marques R, Arnold K, Pilkington CA, Wedderburn LR, Nistala K. Analysis of Published Criteria for Clinically Inactive Disease in a Large Juvenile Dermatomyositis Cohort Shows That Skin Disease Is Underestimated. *Arthritis Rheumatol* 2015; 67(9):2495-2502.
23. Rider LG, Shah M, Mamyrova G, Huber AM, Rice MM, Targoff IN et al. The myositis autoantibody phenotypes of the juvenile idiopathic inflammatory myopathies. *Medicine (Baltimore)* 2013; 92(4):223-243.
24. Shah M, Mamyrova G, Targoff IN, Huber AM, Malley JD, Rice MM et al. The clinical phenotypes of the juvenile idiopathic inflammatory myopathies. *Medicine (Baltimore)* 2013; 92(1):25-41.
25. Deakin CT, Yasin SA, Simou S, Arnold KA, Tansley SL, Betteridge ZE et al. Muscle Biopsy Findings in Combination With Myositis-Specific Autoantibodies Aid Prediction of Outcomes in Juvenile Dermatomyositis. *Arthritis Rheumatol* 2016; 68(11):2806-2816.
26. Miles L, Bove KE, Lovell D, Wargula JC, Bukulmez H, Shao M et al. Predictability of the clinical course of juvenile dermatomyositis based on initial muscle biopsy: a retrospective study of 72 patients. *Arthritis Rheum* 2007; 57(7):1183-1191.

Part

METHOTREXATE



Chapter

7

PREDICTION OF METHOTREXATE EFFICACY AND ADVERSE EVENTS IN PATIENTS WITH JUVENILE IDIOPATHIC ARTHRITIS: A SYSTEMATIC LITERATURE REVIEW

E.H. Pieter van Dijkhuizen and Nico M. Wulffraat

ABSTRACT

Background

Methotrexate (MTX) is the cornerstone disease-modifying anti-rheumatic drug in juvenile idiopathic arthritis (JIA). In JIA, it is important to start effective treatment early to avoid long-term sequelae, such as joint damage. To accomplish this goal, it is crucial to know beforehand who is going to respond well to MTX. In addition, MTX adverse effects such as MTX intolerance occur frequently, potentially hindering its efficacy. To avoid inefficacy of an otherwise effective drug, the physician should be timely aware of these adverse events. Consequently, to optimise treatment of JIA patients with MTX, predictors for efficacy and adverse events should be used in daily clinical practice. The aim of this study was to summarise the existing knowledge about such predictors.

Methods

A systematic literature search was performed in PubMed, Embase and The Cochrane Library, and 1,331 articles were identified. These were selected based on their relevance to the topic and critically appraised according to pre-defined criteria. Predictors for MTX efficacy and adverse events were extracted from the literature and tabulated.

Results

Twenty articles were selected. The overall quality of the studies was good. For MTX efficacy, candidate predictors were antinuclear antibody positivity, the childhood health assessment questionnaire score, the myeloid-related protein 8/14 level, long-chain MTX polyglutamates, bilateral wrist involvement and some single nucleotide polymorphisms (SNPs) in the adenosine triphosphate binding cassette and solute carrier transporter gene families. For MTX adverse events, potential predictors were alanine aminotransferase and thrombocyte level and two SNPs in the γ -glutamyl hydrolase and methylenetetrahydrofolate reductase genes. However, validation of most predictors in independent cohorts was still lacking.

Conclusions

Interesting candidate predictors were found, especially for MTX efficacy. However, most of these were not validated. This should be the goal of future efforts. A clinically relevant way to validate the predictors is by means of creating a clinical prediction model.

INTRODUCTION

Juvenile Idiopathic Arthritis (JIA) is the most common childhood rheumatologic disorder, with a prevalence of 16-150 per 100,000 children. It is characterised by chronic arthritis of unknown aetiology, lasting at least 6 weeks, with an onset before 16 years of age.¹ JIA is a heterogeneous group of disorders, whose manifestations range from relatively mild inflammation of a single joint, to severe involvement of multiple joints lasting into adulthood and leading to structural joint damage and incapacity. These long-term sequelae should be avoided and it is thought that early and effective therapy in the so-called window of opportunity is crucial in doing so.²⁻⁴

The most widely used disease-modifying anti-rheumatic drug (DMARD) in the treatment of JIA is methotrexate (MTX), which has been used since more than 25 years. It is an inexpensive and safe drug and is beneficial in around 70% of JIA patients.^{5,6} Other treatment options include intra-articular joint injections or the more potent biologicals for MTX or corticosteroid resistant cases. It is still impossible to predict the individual prognosis and hence the treatment requirements at the onset of the disease,⁷ leading to the current step-up approach of starting MTX and adding a biological if the patient does not respond sufficiently well to MTX monotherapy. However, given the abovementioned goal to start effective treatment immediately in order to prevent joint damage and the fact that MTX monotherapy is completely ineffective in around 30% of patients, it is essential to know beforehand who is going to respond well to MTX and who is not. The latter group may then be prescribed a biological from the outset.

Next to drug effectiveness, its side effects should be taken into account. It has been shown previously that MTX despite being safe frequently causes transient elevation of liver enzymes and potentially also cytopenias, for which periodic evaluation of blood counts and liver function tests are advised.^{6,8} Perhaps more importantly, gastrointestinal side effects and MTX intolerance occur frequently.⁹⁻¹² MTX intolerance has been shown to influence the quality of life of patients negatively.¹³ Furthermore, these adverse effects potentially cause non-compliance and hence ineffectiveness of an otherwise effective drug,^{9,12,14,15} interfering with the goal to induce early disease remission. To avoid this problem, the risk of occurrence of these adverse effects should be known early, in order for the physician to intervene timely.

Therefore, to optimise treatment of JIA patients, it is necessary to predict the probability of response as well as the risk of developing adverse events. This systematic literature review aims to find and summarise studies, which assessed factors capable of doing so.

METHODS

On 20 April 2014, a systematic literature search was performed in PubMed, Embase and The Cochrane Library, without any publication date or language constraints. Using

the algorithm in table 1 to retrieve all papers regarding JIA and MTX, 1,331 articles were identified (figure 1). These were then screened for applicability to the research subject, the identification of predictors of MTX efficacy and adverse effects (outcome) in JIA patients (domain). Based on the title and the abstract, 45 articles were selected for full-text screening (figure 1). To ensure that all relevant articles had been found, references of selected articles were screened to identify any missed papers.

Selected articles were critically appraised, using predefined criteria (table 2). Studies that were selected, aimed to find predictors for MTX efficacy or side effects within 6 months after the start of therapy in JIA patients, using standardised outcome criteria.

The assessed predictors in the selected studies were summarised in tabular form. Because of the high number of studies and predictors, it was decided to show the direction of the effect of each predictor only, instead of providing an odds ratio and 95% confidence interval. A cut point of $P < 0.05$ was defined to denote significance. Even though some of the assessed studies aimed at constructing a prediction model, for which this cut point is not important, we report significant predictors ($P < 0.05$) only.

RESULTS

After full-text screening, 20 original research papers and 3 reviews were selected (figure 1), of which the former were critically appraised (table 2). The overall relevance and validity of the selected papers was good, leading to only a few papers being excluded from the analysis. No study described whether all patients were eligible to receive the same treatment, but we knew this was the case in our own studies and assumed it was the case in all other studies. Hardly any article described if the physician and researchers were

Table 1. Search strategy^a

Search algorithm	PubMed ^b	Embase ^b	Cochrane ^c
#1 "juvenile idiopathic arthritis" OR "juvenile chronic arthritis" OR "juvenile rheumatoid arthritis" OR "juvenile rheumatic arthritis" OR "childhood arthritis" OR "juvenile arthritis" OR JIA OR JCA OR JRA	7,844	10,906	296
#2 methotrexate OR MTX OR "disease-modifying antirheumatic drug" OR "disease-modifying antirheumatic drugs" OR "disease-modifying anti rheumatic drug" OR "disease-modifying anti rheumatic drugs" OR DMARD OR DMARDs	34,919	50,157	5,251
#3 #1 AND #2	662	1,229	64

^a Search performed on 11 April 2014

^b In PubMed and Embase terms were searched in title and abstract only

^c In The Cochrane Library terms were searched in title, abstract and keywords only

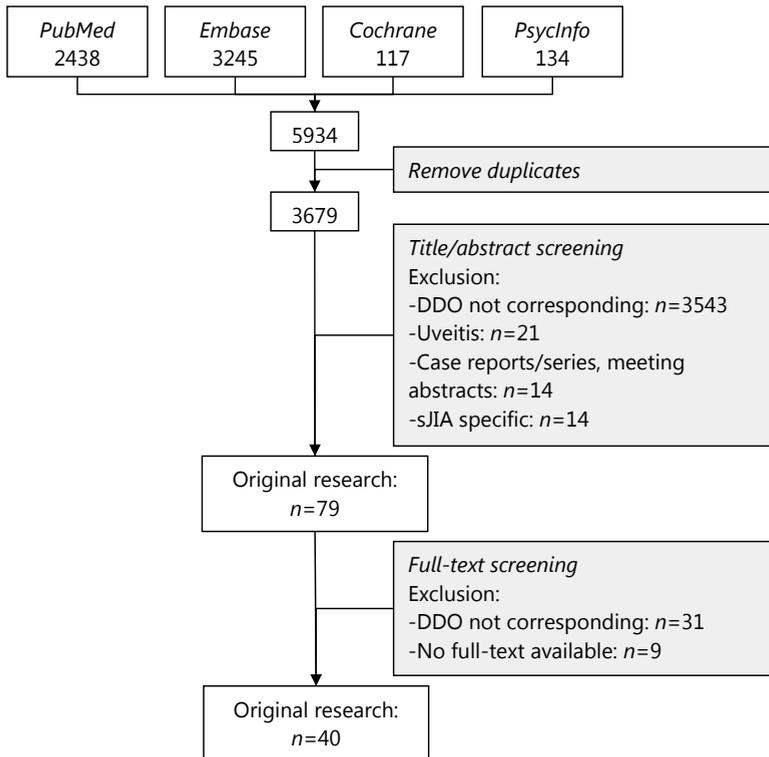


Figure 1. Flow chart of the article selection procedure. **Abbreviations:** DDO, domain, determinant and outcome.

blinded for the outcome at the time of predictor determinations. Follow up was only short term (<1 year) in almost all studies.

MTX efficacy

Fifteen studies assessed MTX efficacy, of which some used a derivation cohort and an independent replication cohort. Since these were independent cohorts, results obtained in these cohorts were reported as if obtained in separate studies. In most of the studies, MTX was started within a median time of 1.5 years after disease onset. The most often used outcome criteria were the American College of Rheumatology (ACR) response criteria. Follow up ranged from 6 months to 1 year, but was as much as 7.3 years in one study (table 3).

The results of these studies are shown in supplementary table 1. Demographics, as well as JIA categories, were analysed extensively and were not predictive in almost all studies. Disease activity parameters showed inconsistent results in general, but the childhood health assessment questionnaire (CHAQ) score was a potential predictor. The same

Table 2. Critical appraisal

Reference	Design	Relevance			Validity				
		Dom	Det	Out	Blind	Rec	SoC	Loss	Mis
Outcome: MTX efficacy									
17	Retrospective (validation prospective)	+	+	+/-	+	+	+	+	+
22	Mixed retrospective and prospective	+	+	+/-	+	+	+	+	+
28	Prospective	+	+	+/-	+	+	+	+	?
29	Prospective	+	+	+/-	+/-	+	?	?	+
30	Cross-sectional	+	+	+/-	+/-	+	?	?	+
10	Retrospective	+	+	+/-	?	+	?	+	+
31	Retrospective	+	+	+/-	?	+	?	?	+
16	Mixed retrospective and prospective	+	+/-	+/-	?	+	?	+	+
32	Prospective (validation unknown)	+	+/-	+/-	?	+	?	?	?
33	Prospective (validation unknown)	+	+/-	+/-	?	+	?	?	?
34a	Retrospective	+/-	+/-	+	?	+	?	+/-	+
23	Prospective	+/-	+	+/-	?	+	?	+	+
35	Prospective	+/-	+	+/-	?	+	?	+	+/-
36	Prospective	+/-	+	+/-	?	+	?	+	?
37	Retrospective	+/-	+/-	+	?	+	?	+/-	+
34b	Retrospective	+/-	+/-	+/-	?	+	?	+/-	+
38	Retrospective	+	+	-	?	+	?	+	+
39	Cross-sectional	+	+	-	+/-	+	?	?	?
40	Retrospective	+	+/-	-	+/-	+	?	?	?
41	Retrospective	+/-	+	-	?	+	?	?	+
Outcome: MTX adverse effects									
19	Prospective	+	+	+/-	+	+	+	+	+
28	Prospective	+	+	+/-	+	+	+	?	?
10	Retrospective	+	+	+/-	?	+	?	+	+
31	Retrospective	+	+	+/-	?	+	?	?	+
30	Cross-sectional	+	+	+/-	+/-	+	?	?	+
39	Cross-sectional	+	+	+/-	+/-	+	?	?	?
40	Retrospective	+	+/-	+/-	+/-	+	?	?	?
42	Cross-sectional	-	+	+	?	-	?	+	+

Abbreviations: Blind, blinding; Det, determinant; Dom, domain; Loss, loss to follow up; Mis, missing predictors; Out, outcome; Rec, recruitment; Ref, reference number; SoC, standardization of care.

Shaded articles were excluded for analysis.

^a Predictors after 6 months for outcome after 5 years; ^b Predictors at baseline for outcome after 6 months.

Criteria: **Domain:** + Children with confirmed JIA, according to currently valid ILAR criteria, starting MTX +/- Children with JCA/JRA according to previously valid criteria, or children with JIA and additional criteria (e.g. hospitalized, specific categories only), starting MTX - Children without JIA/JCA/JRA, or no MTX; **Determinant:** + Prediction model or single predictors corrected for confounding in multivariable analysis +/- Single predictors in univariate analysis - No predictors; **Outcome:** + Efficacy: Any standardized outcome measurement, follow up >1 year. Adverse effects: Any outcome measurement, follow up >1 year +/- follow-up <1 year - Efficacy: No use of standardized outcome criteria; **Blinding:** + Both patient and physician blinded (or not applicable in case of objective measurements) +/- Patient or physician not blinded - Not blinded; **Recruitment:** + Predictors determined at time of start of MTX or <6 months (or time of determination does not matter as in genetic evaluations, gender, age at onset, etc.) +/- Predictors determined more than 6 months after start of MTX, but <1 year - Predictors determined after 1 year, or completely at random; **Standardization of care:** + All participants

Table 2. (continued)

treated according to standards of care - No standardized care; **Loss to follow up (missing outcome):** + <20% and unselective loss to follow up; or >20%, unselective and solved with a statistically valid method (imputation) +/- >20% (not imputed) but unselective loss to follow up - Selective loss to follow up; **Missing predictors:** + <20% and unselective; or >20%, unselective and solved with a statistically valid method (imputation) +/- >20% (not imputed) but unselective - Selective missing predictors.

held true for the physician's global assessment (PGA), although less convincingly so. The involvement of individual joints was assessed in too few studies to be conclusive, but bilateral wrist involvement was a potential predictor. Among laboratory data, positive antinuclear antibody (ANA) was a predictor of better response in three studies. Other interesting predictors could be long-chain MTX polyglutamates (PGs), the myeloid-related protein (MRP) 8/14 (also known as S100A8/A9), the pro-inflammatory molecule osteopontin, or even the haemoglobin level, although these were assessed in only one study each (supplementary table 1).

Next to these predictors, many single-nucleotide polymorphisms (SNPs) were analysed. These were SNPs in genes involved in the MTX metabolic pathway and in genes with altered post-treatment gene expression. Moreover, recently a genome-wide analysis study (GWAS) was published.¹⁶ Of the latter study, only gene regions showing association with MTX response could be reported in this review.

Overall, no unequivocal predictive SNP has been found yet, because many were assessed in only one study, or were predictive in one study and showed no effect in others. However, some SNPs (rs1045642, rs35592 and rs4793665) in the B1, C1 and C3 members of the adenosine triphosphate binding cassette (ABC) transporter family, and others (rs3763980 and rs1051266) in the 16A7 and 19A1 members of the solute carrier (SLC) transporter family were interesting. Furthermore, the gene regions associated in the GWAS study were promising predictors (supplementary table 1).

In all, many potential predictors for MTX efficacy were assessed, yielding some interesting candidates, which, however, were analysed in too few studies to draw a firm conclusion yet.

MTX adverse events

Seven of the selected studies assessed MTX adverse events (table 3). The assessed outcome varied from overall adverse events to single adverse events such as liver toxicity, gastrointestinal complaints or MTX intolerance measured with the Methotrexate Intolerance Severity Score (MISS).⁹ None of the articles focused on (serious) infections. Since the outcome MTX adverse events is a composite of all these outcomes, all studies were included. Follow up ranged from 6 months to a mean of 58.2 months.

Many predictors were evaluated in only one or two studies, making the results inconclusive. However, interesting predictors were the alanine aminotransferase (ALT)

Table 3. Characteristics of included studies

Reference	Design	Country of origin	N	Inclusion criteria	Outcome ^a	Follow up
28 ^b	Prospective	The Netherlands	113	JIA, starting MTX	1k, 2a, 2d, 2f	1 y
29	Prospective	UK	87	JIA, starting MTX	1b, 1j	6 mo
22 ^c	Retrospective and prospective	The Netherlands	287	JIA, starting MTX	1c	1 y
10	Retrospective	Germany	411	JIA, starting MTX	1a, 1b, 1c, 2i	1 y
17 (deriv)	Retrospective	The Netherlands	183	JIA, starting MTX	1e	1 y
17 (rep)	Prospective	The Netherlands	104	JIA, starting MTX	1e	1 y
32 (deriv) ^d	Prospective	UK	197	JIA, starting MTX	1d	6 mo
32 (rep) ^d	Unknown	USA	210	JIA, starting MTX	1g	6 mo
39	Cross-sectional	Japan	92	JIA, at least 3 mo MTX	2e	Mean 58.2 mo ^e
33 (deriv) ^d	Prospective	UK	197	JIA, starting MTX	1d	6 mo
33 (rep) ^d	Unknown	USA	210	JIA, starting MTX	1g	6 mo
30	Cross-sectional	Czech Republic	69	JIA, at least 3 mo MTX	1i, 2d, 2f, 2g, 2h	Median 1.3-1.4 y ^e
36 ^f	Prospective	Multinational (PRINTO)	563	RF negative polyarticular course JIA, starting MTX	1a, 1c	6 mo
23	Prospective	Italy	60	JIA, ≥2 active joints in oligo persistent, ≥5 active joints in other categories	1b	1 y
35 ^f	Prospective	Multinational (PRINTO)	521	RF negative polyarticular course JIA, starting MTX	1l	6 mo
34	Retrospective	Italy	125	Polyarticular JIA, starting MTX	1f, 1i	6 mo, 5 y
40	Retrospective	Germany	58	JIA, at least 3 mo MTX	2d, 2i	Mean 48 months
31	Retrospective	Italy	80	JIA, at least 6 mo MTX	1a, 2c, 2g	Efficacy: 6 mo Toxicity: median 6-9 mo
37	Retrospective	USA	49	JRA, starting MTX	1h	Mean 2.6 y (range 1.0-7.3 y)
19 ^b	Prospective	The Netherlands	152	JIA, starting MTX	2b	1 y

Table 3. (continued)

Reference	Design	Country of origin	N	Inclusion criteria	Outcome ^a	Follow up
16	Retrospective and prospective	Czech Republic, UK, The Netherlands	694	JIA, starting MTX	1f	6 mo

Abbreviations: ACR30/50/70, American College of Rheumatology pediatric 30-, 50 or 70 response criteria, respectively; AE, adverse event; ALT, alanine aminotransferase; AST, aspartate aminotransferase; CHQ, child health questionnaire; deriv, derivation cohort; GI, gastrointestinal; HRQOL, health-related quality of life; JADAS, juvenile arthritis disease activity score; JIA, juvenile idiopathic arthritis; min, minutes; MISS, methotrexate intolerance severity score; mo, months; MTX, methotrexate; NR, non-response; PHS, physical component summary score; PsS, psychosocial component summary score; RA, rheumatoid arthritis; rep, replication cohort; RF, rheumatoid factor; ULN, upper limit of normal; y, years.

^a **1a:** Achievement of ACR30; **1b:** Achievement of ACR50; **1c:** Achievement of ACR70; **1d:** Achievement of ACR70 vs. non-achievement of ACR30; **1e:** Achievement of ACR70 in 2/3 visits; **1f:** NR vs. ACR30 vs. ACR50 vs. ACR70; **1g:** >70% improvement in joint count vs. <30%; **1h:** Adapted ACR criteria for RA: morning stiffness <15 min, no fatigue, no joint swelling, no joint pain for 2 consecutive months; **1i:** Clinical inactive disease on MTX monotherapy according to Wallace criteria; **1j:** JADAS-10; **1k:** JADAS-27; **1l:** HRQOL: CHQ PHS \geq 30 and PsS \geq 30; **2a:** MISS: intolerant (score >6); **2b:** MISS: intolerant (score >6) after 6 and/or 12 months; **2c:** ALT/AST >ULN; **2d:** ALT/AST >2 ULN; **2e:** ALT >5 ULN; **2f:** Bone marrow suppression (any cytopenia); **2g:** GI toxicity; **2h:** Other (alopecia, headaches, behavioural changes, nodulosis); **2i:** Any AE; ^b This is the same cohort as the replication cohort of ¹⁷, but different outcome and/or predictors; ^c This cohort is the derivation and replication cohort of ¹⁷ together, but uses a slightly different outcome and different predictors; ^d These are the same cohorts, but they use different predictors; ^e Time after start of MTX; ^f These are the same cohorts, but they use different outcome measurements.

and thrombocyte level, as well as a SNP (rs1800909) in the γ -glutamyl hydrolase (*GGH*) gene, involved in the breakdown of MTX PGs, and another SNP (rs1801133) in the methylenetetrahydrofolate reductase (*MTHFR*) gene, involved in the folate metabolism. Finally, the polyarticular categories could potentially pose a risk to develop MTX side effects (supplementary table 2).

DISCUSSION

This systematic literature review aimed to find predictors for MTX efficacy and adverse events in JIA patients. For MTX efficacy, many candidate predictors were investigated, and some interesting results were found, such as ANA positivity, the CHAQ score, the MRP 8/14 (S100A8/A9) level, long-chain MTX PGs, bilateral wrist involvement, osteopontin level, haemoglobin, some SNPs in the *ABC* and *SLC* transporter gene families and several gene regions elucidated in the recently published GWAS. Most of these variables have not yet been validated in independent cohorts. Therefore, future efforts should be directed at validating these candidate predictors. A clinically relevant way to do so consists in combining these predictors into a prediction model and assessing the prognostic accuracy of the model. Thus far, only one prediction model for MTX efficacy in JIA patients has been developed, containing the erythrocyte sedimentation rate and four SNPs in genes involved in the MTX metabolic pathway.¹⁷ This model could be improved using the abovementioned candidate predictors. The advantage of this method is twofold: statistically, the selection of predictors for the model in the independent cohort will be literature-driven, instead of data-driven, leading to a reduction in the so-called optimism.¹⁸ Clinically, physicians will have an easy-to-use tool at hand to determine the probability of MTX efficacy in individual patients, allowing them to start MTX in patients with a high probability of responding and to initiate biologicals in those with a low probability of responding. The clinical benefit of this approach should ideally be estimated in a randomised clinical trial.

Regarding MTX adverse events, the results were less clear. Although some interesting candidate predictors were found, such as ALT and thrombocyte level and two SNPs in the *GGH* and *MTHFR* genes, here too, validation of these was lacking. Furthermore, it seemed questionable if these predictors would be sufficient to predict the individual patients' risk of developing MTX adverse effects. Consequently, future efforts should be directed both at validating existing candidate predictors and at finding new predictors. These too should be combined in a clinical prediction model, an example of which is presented in the accompanying paper.¹⁹ Such a tool could be used to monitor high-risk patients more closely, intervening as soon as adverse events occur, for example by lowering the dose, stopping MTX temporarily or prescribing other drugs. On the other hand, low-risk patients could be saved the burden of frequent checks.

Recently, a review was published about genetic predictors of MTX efficacy and toxicity in rheumatoid arthritis.²⁰ SNPs investigated in five or more independent studies were

considered (n=4). Only *ATIC* rs2372536 showed a potential association of the minor allele with toxicity. This SNP did not show an association with adverse events in our review. Conversely, whereas we found an association of *SLC19A1* rs1051266 with efficacy (though in a single study), results about this SNP were inconsistent in the adult review. Thus, the results of the adult and paediatric review were quite different. This might be due to incomparability of children and adults in this respect, maybe because of a different metabolism of MTX.²¹ On the other hand, it shows there is still much to do in the field of MTX prediction.

Most SNPs investigated in this review were located in genes involved in the MTX metabolic pathway. In short, MTX enters the cell via the members of the SLC protein family. It becomes polyglutamated by FPGS, causing its cellular retention. Depolyglutamation is brought about by GGH. MTX is pumped out of the cell by members of the ABC transporter family. Intracellular MTX-PGs exert a range of actions. First, they inhibit DHFR and influence MTHFR, enzymes in the folate pathway involved in polyamine synthesis. Secondly, they inhibit TYMS, an enzyme in the pyrimidine synthesis pathway. Finally, by blocking *ATIC*, MTX-PGs stimulate the production of adenosine, an anti-inflammatory agent. Other important enzymes in this pathway are ITPA and AMPD1, which themselves are not influenced by MTX.^{20;22} It can be hypothesised that SNPs in any of these genes cause increased or decreased sensibility to the actions of MTX-PGs and hence lead to altered MTX efficacy.

Other potential predictors included the MRP8/14 (S100A8/A9) level, a danger signal and activator of toll-like receptor 4, which in turn plays a role in the innate immune response in inflammatory conditions. MRP8/14 was earlier shown to predict disease flare or continuation of remission after withdrawal of MTX in children who were in remission.⁵ Another candidate predictor, osteopontin, is expressed by natural killer cells and activated T cells, and plays a role in the production of pro-inflammatory cytokines. It is overexpressed in synovial T cells in patients with rheumatoid arthritis, demonstrating its role in inflammatory arthritis.²³ The theoretical background of these candidate predictors may increase the likelihood that they really have something to do with MTX efficacy, however, it should be kept in mind that the found associations do not prove a causal effect.

The most frequently used outcome criteria with respect to MTX efficacy were the American College of Rheumatology (ACR) response criteria.²⁴ These criteria, though validated, have their limitations in practical use. For example, a patient with 69% improvement in all core set criteria will not be an ACR70 responder, whereas someone with 70% improvement in three core set criteria and worsening up to 29% in the remainder will be an ACR70 responder, despite the fact that the former obviously responds better than the latter. Hence, if one takes the achievement of ACR70 response status as an outcome measurement, considerable misclassification may occur. Alternatively, the percentage of change of the juvenile arthritis disease activity score (JADAS) could be used as an outcome measurement to assess MTX efficacy.^{25;26} Because this is a continuous

and composite outcome, the risk of misclassification may be reduced. Recently, the ACR criteria and the JADAS were compared, showing excellent ability of the JADAS to classify the ACR response categories.²⁷ To answer the question which outcome measurement should be preferred, both measurements should be compared to a reference standard, such as a panel of experts.

The studies that were analysed in this review were generally of good quality. They recruited patients at the start of MTX and followed them prospectively. However, many articles did not describe whether the studies were blinded, an absence of which could lead to biases. On a review level, due to our extensive search strategy, it is likely that we found all pertinent papers. However, some negative results might not have been published, leading to reporting and publication bias. The results found in different studies were often quite variable. This may be due to heterogeneity of patient groups, differences in sample size and the absence or presence of linkage disequilibrium between the tested SNP and the actual polymorphism which causes the altered MTX response.²⁰ Furthermore, other patient factors may confound the observed relationship, reason why it is important to perform a multivariate analysis. This was not done in all studies. Finally, the authors of this review were (co-)authors of some of the included papers, causing them to know more about the study design and potentially biasing them to be more lenient towards their own studies.

In conclusion, this systematic review of the literature with respect to predictors for MTX efficacy and adverse events shows that many interesting candidates were found. However, validation of these potential predictors is still lacking in many cases. Therefore, future efforts should be directed at validating these candidate predictors, potentially by means of a clinical prediction model. For the outcome adverse events, next to validating existing candidates, more candidate predictors should be investigated.

FUNDING

EHPvD was supported by funding from the 7th Framework programme of the EU, SP3-People, support for training and career development for researchers (Marie Curie), Network for Initial Training (ITN), FP7-PEOPLE-2011-ITN, under the Marie Skłodowska-Curie grant agreement No 289903. The project was furthermore supported by the SHARE project, EAHC grant number 2011 1202 [grant to NMW]. The sponsors had no role in the study planning, design, management or data analysis.

SUPPLEMENTARY MATERIAL

Supplementary table 1. Results for outcome MTX efficacy^a

Baseline predictors (within 6 months after MTX start)	Number of studies					References
	--	< ^b	NS	S ^b	> ^b	
Demographics						
Gender: female			8			17;29-31;34-36
Higher age at onset			7			17;29;31;34-36
Higher age at MTX start			8			17;29-31;35-37
Longer disease duration before MTX start	1		7			17;29;31;34-37
Higher body surface area			1			30
JIA category^c						
Oligoarticular persistent			1			37
Oligoarticular extended			3		1	31;34;35;37
Systemic			7			17;31;34-37
Polyarticular, rheumatoid factor negative	0/1		6			17;31;34-37
Polyarticular, rheumatoid factor positive	0/1		6			17;31;34-37
Psoriatic			2			17
Enthesitis-related arthritis			2			17
Undifferentiated			2			17
Disease activity						
Higher active joint count	1/0	5	1/0	1/0		17;29-31;34-36
Higher swollen joint count	1/0		1/0			35;36
Higher limited joint count		6			1	17;29;30;34-36
Higher painful joint count		1	1/0			35;36
Higher PGA		4		1/0	1	17;29;34-36
Higher parent/patient GA	1		2		1	29;34-36
Higher pain VAS	1		1			35;36
Higher CHAQ score	2	1/0	2			29;31;34-36
Functional status			1			37
Higher CHQ PhS score			1			36
Higher CHQ PsS score			1			36
Achievement ACR70 at 6 months					0/1	34
Hand or wrist involvement		1/0				31
Wrist: right and left activity	1					36
Wrist: ≥1 active		1/0				36
MCP 1: right and left activity		1/0				36
MCP 1: ≥1 active		1/0				36
MCP 2: ≥1 active		1/0				36
MCP 3: ≥1 active		1/0				36
PIP 1: right and left activity		1/0				36
PIP 2: right and left activity		1/0				36
Knee: ≥1 active					1/0	36

Supplementary table 1. (continued)

Baseline predictors (within 6 months after MTX start)	Number of studies					References
	--	< ^b	NS	S ^b	> ^b +	
Laboratory						
ANA positive		2		0/1	2	17;35-37
Rheumatoid factor positive		1				37
Higher ESR	1	5		2/1		17;23;29-31;34-36
Higher CRP		1				29
Higher haemoglobin					1	23
Higher MRP8/14					1	29
Higher IL-1 β		1				29
Higher IL-2				1/0		29
Higher IL-6		1				29
Higher IL-10		1				29
Higher IL-12		1				29
Higher IL-18		1				29
Higher IL-22		1				29
Higher TNF- α		1				29
Higher IFN- γ		1				29
Higher osteopontin level	1					23
Medication						
Higher MTX start dose		6				17;29;34;36;37
Subcutaneous route of administration		3				29;36;43
Start with folic acid		2				17
Taking corticosteroids		2				31;36
Taking NSAIDs		1				36
MTX-polyglutamates						
Higher MTX-PG1		1				28
Higher MTX-PG2		1				28
Higher MTX-PG3					1	28
Higher MTX-PG4					1	28
Higher MTX-PG5					1	28
Higher MTX-PG3-5					1	28
Higher total MTX-PG		1				28
Imaging						
Presence of radiologic lesions		2				31;37
Higher Poznanski score		1				34
Genetics^d						
<i>MTHFR</i> rs7538516, minor allele		1				32
<i>MTHFR</i> rs4846047, minor allele		1				32
<i>MTHFR</i> rs4846049, minor allele		1				32
<i>MTHFR</i> rs2274976, minor allele		1				32
<i>MTHFR</i> rs3818762, minor allele		1				32
<i>MTHFR</i> rs12121543, minor allele		1				32

Supplementary table 1. (continued)

Baseline predictors (within 6 months after MTX start)	Number of studies					References	
	--	< ^b	NS	S ^b	> ^b		+
<i>MTHFR</i> rs1801133, dominant model			3				17;30
<i>MTHFR</i> rs1801133, homozygous variant vs. wild type			1				30
<i>MTHFR</i> rs1801133, heterozygous vs. wild type			1				30
<i>MTHFR</i> rs1801133, minor allele			2				30;32
<i>MTHFR</i> rs1801131, dominant model			2		0/1		17;30
<i>MTHFR</i> rs1801131, homozygous variant vs. wild type			1				30
<i>MTHFR</i> rs1801131, heterozygous vs. wild type			1				30
<i>MTHFR</i> rs1801131, minor allele			1				30
<i>MTHFR</i> rs17367504, minor allele			1				32
<i>MTHFR</i> rs7553194, minor allele			1				32
<i>MTRR</i> rs1801394, dominant model			1		1/0		17
<i>AMPD1</i> rs6658815, minor allele			1				32
<i>AMPD1</i> rs6679869, minor allele			1				32
<i>AMPD1</i> rs2268699, minor allele			1				32
<i>AMPD1</i> rs2336363, minor allele			1				32
<i>AMPD1</i> rs2268701, minor allele			1				32
<i>AMPD1</i> rs17602729, dominant model			2				17
<i>AMPD1</i> rs17602729, minor allele			1				32
<i>AMPD1</i> rs11587596, minor allele			1				32
<i>ATIC</i> rs2372536, dominant model			2				17
<i>ATIC</i> rs12995526, minor allele		0/1	1				32
<i>ATIC</i> rs12477799, minor allele			1				32
<i>ATIC</i> rs16853826, minor allele			1				32
<i>ATIC</i> rs4673990, minor allele		0/1	1				32
<i>ATIC</i> rs4672768, minor allele			1				32
<i>ATIC</i> rs10498036, minor allele			1				32
<i>ABCB1</i> rs1128503, recessive model			3				17;22
<i>ABCB1</i> rs1045642, recessive model			1		1/0	1	17;22
<i>ABCB1</i> rs2032582, recessive model			1				22
<i>ABCB1</i> rs1128503/rs32032582/rs1045642 haplotype AAA, recessive model			1				22
<i>ABCB1</i> rs1128503/rs32032582/rs1045642 haplotype GCG, recessive model			1/0				22
<i>ABCC1</i> rs35592, dominant model			1			1	17
<i>ABCC1</i> rs35592, recessive model					1/0		22
<i>ABCC1</i> rs3784862, dominant model			2				17
<i>ABCC1</i> rs3784862, recessive model			1				22
<i>ABCC2</i> rs4148396, recessive model			1				22
<i>ABCC2</i> rs717620, dominant model			2				17
<i>ABCC2</i> rs717620, recessive model			1				22

Supplementary table 1. (continued)

Baseline predictors (within 6 months after MTX start)	Number of studies					References
	--	< ^b	NS	S ^b	> ^b	
<i>ABCC2</i> rs4148396/rs717620 haplotype TC, recessive model			1			22
<i>ABCC2</i> rs4148396/rs717620 haplotype TT, recessive model			1			22
<i>ABCC3</i> rs4793665, dominant model			2			17
<i>ABCC3</i> rs4793665, recessive model					1	22
<i>ABCC3</i> rs3785911, dominant model			2			17
<i>ABCC3</i> rs3785911, recessive model			1			22
<i>ABCC5</i> rs2139560, dominant model			2			17
<i>ABCC5</i> rs2139560, recessive model			1			22
<i>ABCG2</i> rs13120400, dominant model			2			17
<i>ABCG2</i> rs13120400, recessive model		1/0				22
<i>ABCG2</i> rs13120400/rs2231142 haplotype CC, recessive model			1			22
<i>ABCG2</i> rs13120400/rs2231142 haplotype CG, recessive model		1/0				22
<i>ABCG2</i> rs13120400/rs2231142 haplotype TG, recessive model					1/0	22
<i>ABCG2</i> rs2728124, minor allele			1			32
<i>ABCG2</i> rs2231164, minor allele			1			32
<i>ABCG2</i> rs12505410, minor allele			1			32
<i>ABCG2</i> rs2622621, minor allele			1			32
<i>ABCG2</i> rs2199936, minor allele			1			32
<i>ABCG2</i> rs1564481, minor allele			1			32
<i>ABCG2</i> rs3114018, minor allele			1			32
<i>ABCG2</i> rs3109823, minor allele			1			32
<i>ABCG2</i> rs6857600, minor allele			1			32
<i>ABCG2</i> rs2622626, minor allele			1			32
<i>ABCG2</i> rs17731799, minor allele			1			32
<i>DHFR</i> rs1222809, minor allele			1			32
<i>DHFR</i> rs12517451, minor allele			1			32
<i>DHFR</i> rs1650723, minor allele			1			32
<i>DHFR</i> rs1643657, minor allele			1			32
<i>DHFR</i> rs11951910, minor allele			1			32
<i>DHFR</i> rs10072026, minor allele			1			32
<i>DHFR</i> rs380691, minor allele			1			32
<i>GGH</i> rs11545078, minor allele			1			32
<i>GGH</i> rs12335094, minor allele			1			32
<i>GGH</i> rs16930092, minor allele			1			32
<i>GGH</i> rs3780130, minor allele			1			32
<i>GGH</i> rs10957267, minor allele			1			32
<i>GGH</i> rs17194931, minor allele			1			32

Supplementary table 1. (continued)

Baseline predictors (within 6 months after MTX start)	Number of studies					References
	--	< ^b	NS	S ^b	> ^b	
<i>GGH</i> rs719235, minor allele			1			32
<i>GGH</i> rs10106587, dominant model			2			17
<i>GGH</i> rs10106587, recessive model			1			22
<i>GGH</i> rs3758149, dominant model			2			17
<i>GGH</i> rs3758149, recessive model			1			22
<i>GGH</i> rs3758149, minor allele			1			32
<i>GGH</i> rs10106587/rs3758149 haplotype AA, recessive model			1			22
<i>GGH</i> rs10106587/rs3758149 haplotype CG, recessive model			1			22
<i>GGH</i> rs10106587/rs3758149 haplotype AG, recessive model			1			22
<i>GGH</i> rs6998134, minor allele			1			32
<i>FPGS</i> rs4451422, recessive model			1			22
<i>FPGS</i> rs4451422, minor allele			1			32
<i>FPGS</i> rs10819309, minor allele			1			32
<i>MTHFD1</i> rs2983733, minor allele			1			32
<i>MTHFD1</i> rs1956545, minor allele			1			32
<i>MTHFD1</i> rs8011839, minor allele			1			32
<i>MTHFD1</i> rs8003379, minor allele			1			32
<i>MTHFD1</i> rs1950902, minor allele			1			32
<i>MTHFD1</i> rs17101851, minor allele			1			32
<i>MTHFD1</i> rs3783726, minor allele			1			32
<i>MTHFD1</i> rs2236224, minor allele			1			32
<i>MTHFD1</i> rs11629135, minor allele			1			32
<i>MTHFD1</i> rs3818239, minor allele			1			32
<i>MTHFD1</i> rs1256146, minor allele			1			32
<i>MTHFD1</i> rs11627387, minor allele			1			32
<i>MTHFD1</i> rs745686, minor allele			1			32
<i>MTHFD1</i> rs3742609, minor allele			1			32
<i>SHMT1</i> rs1979276, minor allele			1			32
<i>SHMT1</i> rs2168781, minor allele			1			32
<i>SHMT1</i> rs11868708, minor allele			1			32
<i>SHMT1</i> rs2273027, minor allele			1			32
<i>SHMT1</i> rs2273026, minor allele			1			32
<i>SHMT1</i> rs9901160, minor allele			1			32
<i>SHMT1</i> rs8065874, minor allele			1			32
<i>TYMS</i> rs9966612, minor allele			1			32
<i>TYMS</i> rs11664283, minor allele			1			32
<i>TYMS</i> rs2853741, minor allele			1			32
<i>TYMS</i> rs2853533, minor allele			1			32
<i>TYMS</i> rs2612095, minor allele			1			32

Supplementary table 1. (continued)

Baseline predictors (within 6 months after MTX start)	Number of studies					References
	--	< ^b	NS	S ^b	> ^b	
<i>TYMS</i> rs2847150, minor allele			1			32
<i>TYMS</i> rs9948583, minor allele			1			32
<i>TYMS</i> rs2298582, minor allele			1			32
<i>TYMS</i> rs2741186, minor allele			1			32
<i>TYMS</i> rs7239738, minor allele			1			32
<i>ITPA</i> rs6051639, minor allele			1			32
<i>ITPA</i> rs2295553, minor allele		0/1				32
<i>ITPA</i> rs6051644, minor allele			1			32
<i>ITPA</i> rs6084305, minor allele			1			32
<i>ITPA</i> rs1127354, minor allele			1			32
<i>ITPA</i> rs4815576, minor allele			1			32
<i>ITPA</i> rs6037506, minor allele			1			32
<i>ITPA</i> rs6139035, minor allele			1			32
<i>ITPA</i> rs13830, minor allele			1			32
<i>ITPA</i> rs6051655, minor allele			1			32
<i>ITPA</i> rs3810560, minor allele			1			32
<i>SLC16A7</i> rs17122830, minor allele			2			33
<i>SLC16A7</i> rs1497474, minor allele			2			33
<i>SLC16A7</i> rs10877327, minor allele			1			33
<i>SLC16A7</i> rs7976956, minor allele			1			33
<i>SLC16A7</i> rs1000708, minor allele			1			33
<i>SLC16A7</i> rs7971953, minor allele			2			33
<i>SLC16A7</i> rs12231740, minor allele			2			33
<i>SLC16A7</i> rs2711669, minor allele			2			33
<i>SLC16A7</i> rs2706301, minor allele			1			33
<i>SLC16A7</i> rs10877333, minor allele			1	0/1		33
<i>SLC16A7</i> rs2711655, minor allele		0/1				33
<i>SLC16A7</i> rs12718000, minor allele			1			33
<i>SLC16A7</i> rs3763980, minor allele		0/2				33
<i>SLC16A7</i> rs10784000, minor allele			1			33
<i>SLC16A7</i> rs2706301/ rs10877333/ rs2711655/ rs12718000/ rs3763980/ rs10784000 haplotype GTGGAA		0/1				33
<i>SLC16A7</i> rs2706301/ rs10877333/ rs2711655/ rs12718000/ rs3763980/ rs10784000 haplotype CTAGTG			1			33
<i>SLC16A7</i> rs2706301/ rs10877333/ rs2711655/ rs12718000/ rs3763980/ rs10784000 haplotype GGAGTA				0/1		33
<i>SLC16A7</i> rs2706301/ rs10877333/ rs2711655/ rs12718000/ rs3763980/ rs10784000 haplotype CTAGTA			1			33

Supplementary table 1. (continued)

Baseline predictors (within 6 months after MTX start)	Number of studies					References
	--	< ^b	NS	S ^b	> ^b	
<i>SLC16A7</i> rs2706301/ rs10877333/ rs2711655/ rs12718000/ rs3763980/ rs10784000 haplotype GTGATA			1			33
<i>SLC16A7</i> rs2706301/ rs10877333/ rs2711655/ rs12718000/ rs3763980/ rs10784000 haplotype GTGGTA			1			33
<i>SLC19A1</i> rs1051266, recessive model	1					22
<i>SLC19A1</i> rs2274808, minor allele			1			32
<i>SLC19A1</i> rs9977268, minor allele			1			32
<i>SLC19A1</i> rs11702425, minor allele			1			32
<i>SLC19A1</i> rs1556329, minor allele			1			32
<i>SLC19A1</i> rs2236475, minor allele			1			32
<i>SLC19A1</i> rs2236479, minor allele			1			32
<i>SLC19A1</i> rs7279445, minor allele			1			32
<i>SLC19A1</i> rs3753019, minor allele			1			32
<i>SLC19A1</i> rs10483080, minor allele			1			32
<i>SLC19A1</i> rs2838950, minor allele			1			32
<i>SLC19A1</i> rs2838951, minor allele			1			32
<i>SLC19A1</i> rs12483377, minor allele			1			32
<i>SLC19A1</i> rs7499, minor allele			1			32
<i>SLC19A1</i> rs17004785, minor allele			1			32
<i>SLC19A1</i> rs2838956, minor allele			1			32
<i>SLC19A1</i> rs3788205, minor allele			1			32
<i>SLC46A1</i> rs2239907, dominant model			2			17
<i>SLC46A1</i> rs2239907, recessive model			1			22
<i>FOLR2</i> rs514933, recessive model			1			22
<i>ADORA2A</i> rs5751876, dominant model			2			17
<i>ADORA2A</i> rs1041748, minor allele			1			32
<i>ADORA2A</i> rs5751846, minor allele			1			32
<i>ADORA2A</i> rs6004146, minor allele			1			32
<i>ADORA2A</i> rs5751862, minor allele			1			32
<i>ADORA2A</i> rs4822488, minor allele			1			32
<i>ADORA2A</i> rs3761422, minor allele			1			32
<i>ADORA2A</i> rs11704811, minor allele			1			32
<i>ADORA2A</i> rs2236624, minor allele			1			32
<i>THAP6</i> rs3853187, minor allele			1			33
<i>THAP6</i> rs1841934, minor allele			1			33
<i>THAP6</i> rs2126854, minor allele			1			33
<i>THAP6</i> rs9307834, minor allele			1			33
<i>THAP6</i> rs6535523, minor allele			1			33
<i>THAP6</i> rs12649508, minor allele			1			33
<i>BAT1</i> rs3132454, minor allele			1			33

Supplementary table 1. (continued)

Baseline predictors (within 6 months after MTX start)	Number of studies					References
	--	< ^b	NS	S ^b	> ^b	
<i>BAT1</i> rs9267464, minor allele			1			33
<i>BAT1</i> rs3093993, minor allele			1			33
<i>BAT1</i> rs3093992, minor allele			1			33
<i>BAT1</i> rs3093988, minor allele			1			33
<i>BAT1</i> rs2734574, minor allele			1			33
<i>BAT1</i> rs2259435, minor allele			1			33
<i>BAT1</i> rs3130055, minor allele			1			33
<i>BAT1</i> rs3219190, minor allele			1			33
<i>BAT1</i> rs2734583, minor allele			1			33
<i>BAT1</i> rs2239709, minor allele			1			33
<i>BAT1</i> rs2844509, minor allele			1			33
<i>BAT1</i> rs2239705, minor allele			1			33
<i>BAT1</i> rs2071591, minor allele			1			33
<i>BAT1</i> rs2523500, minor allele			1			33
<i>BAT1</i> rs6916921, minor allele			1			33
<i>BAT1</i> rs6929796, minor allele			1			33
<i>ZEB1</i> rs11008463, minor allele			1			33
<i>ZEB1</i> rs2839657, minor allele			1			33
<i>ZEB1</i> rs172683, minor allele			1			33
<i>MALAT1</i> rs600231, minor allele			1			33
<i>MALAT1</i> rs3200401, minor allele			1			33
<i>NFATC2IP</i> rs12931589, minor allele			1			33
<i>NFATC2IP</i> rs11150675, minor allele			1			33
<i>NFATC2IP</i> rs7192056, minor allele			1			33
Region <i>LMX1A, PBX1</i>				0/1		16
Region <i>CDH6</i>				0/1		16
Region <i>CFTR, CTTNBP2</i>				0/1		16
Region <i>CSMD1</i>				0/1		16
Region <i>SNX16</i>				0/1		16
Region <i>PVT1, ADCY8</i>				0/1		16
Region <i>ZMIZ1</i>				0/1		16
Region <i>ANGPTL5, KIAA1377</i>				0/1		16
Region <i>ANKS1B, ANO4, ARL1, SPIC</i>				0/1		16
Region <i>CMKLR1</i>				0/1		16
Region <i>GABRB3</i>				0/1		16
Region <i>TGIF1</i>				0/1		16
Region <i>CYTH4</i>				0/1		16
Region <i>CACNA1I</i>				0/1		16

Supplementary table 1. (continued)

Abbreviations: *ABCB1* to *ABCG2*, members of the adenosine triphosphate binding cassette (ABC) transporters; *ACR70*, American College of Rheumatology paediatric 70 response criterion; *ADCY8*, adenylate cyclase 8; *ADORA2A*, adenosine A2A receptor; *ALT*, alanine aminotransferase; *AMPD1*, adenosine monophosphate deaminase; *ANA*, antinuclear antibodies; *ANGPTL5*, angiopoietin-like 5; *ANKS1B*, ankyrin repeat and sterile alpha motif domain containing 1B; *ANO4*, anoctamin 4; *ARL1*, ADP-ribosylation factor-like 1; *AST*, aspartate aminotransferase; *ATIC*, 5-aminoimidazole-4-carboxamide ribonucleotide transformylase; *BAT1*, HLA-B associated transcript 1; *CACNA1I*, calcium-channel, voltage-dependent, T type, alpha 1I subunit; *CDH6*, cadherin 6; *CFTR*, cystic fibrosis transmembrane conductance regulator; *CHAQ*, childhood health assessment questionnaire; *CHQ*, child health questionnaire; *CMKLR1*, chemokine-like receptor 1; *CRP*, C-reactive protein; *CSMD1*, CUB and Sushi multiple domains 1; *CTTNBP2*, cortactin-binding protein 2; *CYTH4*, cytohesin 4; *DHFR*, dihydrofolate reductase; *ESR*, erythrocyte sedimentation rate; *FOLR*, folate receptor; *FPGS*, folic acid synthetase; *GA*, global assessment; *GABRB3*, gamma-aminobutyric acid A receptor, beta 3; *GGH*, γ -glutamyl hydrolase; *HLA*, human leukocyte antigen; *IFN*, interferon; *IL*, interleukin; *ITPA*, inosine triphosphatase; *JADAS*, juvenile arthritis disease activity score; *JIA*, juvenile idiopathic arthritis; *LMX1A*, LIM homeobox transcription factor 1 alpha; *MALAT1*, metastasis associated lung adenocarcinoma transcript 1; *MCP*, metacarpophalangeal joint; *MRP*, myeloid-related protein; *MTHFD1*, methylenetetrahydrofolate dehydrogenase; *MTHFR*, methylenetetrahydrofolate reductase; *MTRR*, methionine synthase reductase; *MTX*, methotrexate; *MTX-PG*, methotrexate polyglutamate; *NFATC2IP*, nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2 interacting protein; *NSAIDs*, non-steroidal anti-inflammatory drugs; *PBX1*, pre-B-cell leukemia homeobox 1; *PGA*, physician's global assessment; *PhS*, physical component summary score; *PIP*, proximal interphalangeal joint; *PsS*, psychosocial component summary score; *PVT1*, Pvt1 oncogene; *SHMT1*, serine hydroxymethyltransferase; *SLC*, solute carrier; *SNP*, single nucleotide polymorphism; *SNX16*, sorting nexin 16; *SPIC*, Spi-C transcription factor; *TGIF1*, transforming growth factor beta-induced factor homeobox 1; *THAP6*, THAP domain containing 6; *TNF*, tumor necrosis factor; *TYMS*, thymidylate synthase; *VAS*, visual analog score; *ZEB1*, zinc finger-enhancer protein 1, *ZMIZ1*, zinc-finger, MIZ-type containing 1.

^a Values are the number of studies which found that particular result for the respective predictors; ^b The first number in this column refers to the number of studies which performed a multivariate analysis disconfirming the univariate finding (studies in which the multivariate analysis confirmed the univariate finding, are shown in the columns labelled -- and +, respectively). The second number in this column refers to the number of studies which did not perform a multivariate analysis; ^c In studies that evaluated onset categories oligoarticular, polyarticular and systemic, results have been duplicated for oligoarthritis persistent and extended, and polyarthritis RF positive and negative respectively; ^d SNPs were analyzed in different ways, denoted in the table as follows: minor allele: allelic comparison. Odds ratios shown in the source study are allelic odds ratios, indicating the increase in odds per minor allele. Dominant model: homozygous variant and heterozygous subjects vs. wild type. Recessive model: homozygous variant vs. heterozygous and wild type. Haplotype: recessive model: subjects homozygous for the shown haplotype versus all other subjects.

Symbols used:

-- lower chance of favourable outcome in multivariate analysis ($p \leq 0.05$)

< lower chance of favourable outcome in univariate analysis ($p \leq 0.05$)

NS: not significant ($p > 0.05$)

S: significant in univariate analysis. Direction of the effect not shown, or different SNPs in the same region with opposed effects ($p \leq 0.05$)

> higher chance of favourable outcome in univariate analysis ($p \leq 0.05$)

+ higher chance of favourable outcome in multivariate analysis ($p \leq 0.05$)

Supplementary table 2. Results for outcome absence of MTX adverse events^a

Baseline predictors (within 6 months after MTX start)	Number of studies					References
	--	< ^b	NS	> ^b	+	
Demographics						
Gender: female			3			19;31;39
Higher age at onset			3			19;31;39
Higher age at MTX start			2			19;31
Longer disease duration before MTX start			3			19;31;39
JIA category^c						
Oligoarticular persistent						
Oligoarticular extended			1			31
Systemic			2			19;31
Polyarticular, rheumatoid factor negative	1		1			19;31
Polyarticular, rheumatoid factor positive	1		1			19;31
Psoriatic			1			19
Enthesitis-related arthritis			1			19
Undifferentiated			1			19
Disease activity						
Higher active joint count			2			19;31
Higher limited joint count			1			19
Higher PGA			1			19
Higher parent/patient GA			1			19
Higher pain VAS			1			19
Higher CHAQ score			2			19;31
Higher JADAS-27				1/0		19
Hand or wrist involvement			1			31
Uveitis present			1			19
Laboratory						
ANA positive			1			19
Rheumatoid factor positive			1			19
Higher ESR			2			19;31
Higher CRP			1			19
Higher hemoglobin			1			19
Higher leukocyte count			1			19
Higher thrombocyte count			1		1	19;31
Higher ALT level					1	19
Higher AST level			1			19
Higher creatinine level			1			19
Medication						
Subcutaneous route of administration			1			43
Restart of MTX			1			19
Start with folic acid			1			39
Taking corticosteroids			2			31;39
Taking NSAIDs			1			19

Supplementary table 2. (continued)

Baseline predictors (within 6 months after MTX start)	Number of studies					References
	--	< ^b	NS	> ^b	+	
MTX-polyglutamates						
Higher MTX-PG1			1			28
Higher MTX-PG2			1			28
Higher MTX-PG3			1			28
Higher MTX-PG4			1			28
Higher MTX-PG5			1			28
Higher MTX-PG3-5			1			28
Higher total MTX-PG			1			28
Imaging						
Presence of radiologic lesions			1			31
Genetics^d						
HLA-B27 positive			1			19
<i>MTHFR</i> rs1801133, dominant model		0/1	1			39;40
<i>MTHFR</i> rs1801133, recessive model		1/0	2			19;30;39
<i>MTHFR</i> rs1801133, homozygous variant vs. wild type		1/0				30
<i>MTHFR</i> rs1801133, heterozygous vs. wild type		0/1	1			30;40
<i>MTHFR</i> rs1801133, minor allele	1		1			30;39
<i>MTHFR</i> rs1801131, dominant model			4			19;30;39;40
<i>MTHFR</i> rs1801131, recessive model			1			39
<i>MTHFR</i> rs1801131, homozygous variant vs. wild type			1			30
<i>MTHFR</i> rs1801131, heterozygous vs. wild type			2			30;40
<i>MTHFR</i> rs1801131, minor allele			2			30;39
<i>MTRR</i> rs1801394, dominant model			1			19
<i>AMPD1</i> rs17602729, dominant model			1			19
<i>ATIC</i> rs2372536, dominant model			2			19;39
<i>ATIC</i> rs2372536, recessive model			1			39
<i>ATIC</i> rs2372536, minor allele			1			39
<i>ABCB1</i> rs1128503, recessive model			1			19
<i>ABCB1</i> rs1045642, recessive model			1			19
<i>ABCB1</i> rs2032582, dominant model			1			19
<i>ABCC1</i> rs35592, dominant model			1			19
<i>ABCC1</i> rs3784862, dominant model			1			19
<i>ABCC2</i> rs4148396, recessive model			1			19
<i>ABCC2</i> rs717620, dominant model			1			19
<i>ABCC3</i> rs4793665, dominant model			1			19
<i>ABCC3</i> rs3785911, dominant model			1			19
<i>ABCC4</i> rs868853, dominant model			1			19
<i>ABCC4</i> rs2274407, dominant model			1			19
<i>ABCC5</i> rs2139560, dominant model			1			19
<i>ABCG2</i> rs13120400, dominant model			1			19
<i>ABCG2</i> rs2231142, dominant model			2			19;39
<i>ABCG2</i> rs2231142, recessive model			1			39

Supplementary table 2. (continued)

Baseline predictors (within 6 months after MTX start)	Number of studies					References
	--	< ^b	NS	> ^b	+	
<i>ABCG2</i> rs2231142, minor allele			1			39
<i>GGH</i> rs1800909, dominant model			1			39
<i>GGH</i> rs1800909, recessive model					1	39
<i>GGH</i> rs1800909, minor allele			1			39
<i>GGH</i> rs11545078, recessive model			1			39
<i>GGH</i> rs11545078, minor allele			1			39
<i>GGH</i> rs10106587, dominant model			1			19
<i>GGH</i> rs3758149, dominant model			1			19
<i>FPGS</i> rs10106, dominant model			1			39
<i>FPGS</i> rs10106, recessive model			1			39
<i>FPGS</i> rs10106, minor allele			1			39
<i>FPGS</i> rs4451422, dominant model			1			19
<i>ITPA</i> rs1127354, recessive model			1			19
<i>SLC19A1</i> rs1051266, dominant model			1			39
<i>SLC19A1</i> rs1051266, recessive model			2			19;39
<i>SLC19A1</i> rs1051266, minor allele			1			39
<i>SLC46A1</i> rs2239907, dominant model			1			19
<i>ADORA2A</i> rs5751876, recessive model			1			19

For declaration of symbols, footnotes and abbreviations, see supplementary table 1.

REFERENCE LIST

1. Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. *Lancet* 2011; 377(9783):2138-2149.
2. Boers M. Understanding the window of opportunity concept in early rheumatoid arthritis. *Arthritis Rheum* 2003; 48(7):1771-1774.
3. Tynjala P, Vahasalo P, Tarkiainen M, Kroger L, Aalto K, Malin M et al. Aggressive combination drug therapy in very early polyarticular juvenile idiopathic arthritis (ACUTE-JIA): a multicentre randomised open-label clinical trial. *Ann Rheum Dis* 2011; 70(9):1605-1612.
4. Wallace CA, Giannini EH, Spalding SJ, Hashkes PJ, O'Neil KM, Zeff AS et al. Trial of early aggressive therapy in polyarticular juvenile idiopathic arthritis. *Arthritis Rheum* 2012; 64(6):2012-2021.
5. Foell D, Wulffraat N, Wedderburn LR, Wittkowski H, Frosch M, Gerss J et al. Methotrexate withdrawal at 6 vs 12 months in juvenile idiopathic arthritis in remission: a randomized clinical trial. *JAMA* 2010; 303(13):1266-1273.
6. Ruperto N, Murray KJ, Gerloni V, Wulffraat N, deOliveira SK, Falcini F et al. A randomized trial of parenteral methotrexate comparing an intermediate dose with a higher dose in children with juvenile idiopathic arthritis who failed to respond to standard doses of methotrexate. *Arthritis Rheum* 2004; 50(7):2191-2201.
7. Van Dijkhuizen EHP, Wulffraat NM. Early predictors of prognosis in juvenile idiopathic arthritis: a systematic literature review. *Ann Rheum Dis* 2015; 74(11):1996-2005.
8. Ortiz-Alvarez O, Morishita K, Avery G, Green J, Petty RE, Tucker LB et al. Guidelines for blood test monitoring of methotrexate toxicity in juvenile idiopathic arthritis. *J Rheumatol* 2004; 31(12):2501-2506.
9. Bulatovic M, Heijstek MW, Verkaaik M, Van Dijkhuizen EHP, Armbrust W, Hoppenreijns EP et al. High prevalence of methotrexate intolerance in juvenile idiopathic arthritis: development and validation of a methotrexate intolerance severity score. *Arthritis Rheum* 2011; 63(7):2007-2013.
10. Klein A, Kaul I, Foeldvari I, Ganser G, Urban A, Horneff G. Efficacy and safety of oral and parenteral methotrexate therapy in children with juvenile idiopathic arthritis: an observational study with patients from the German Methotrexate Registry. *Arthritis Care Res (Hoboken)* 2012; 64(9):1349-1356.
11. Murray KJ, Lovell DJ. Advanced therapy for juvenile arthritis. *Best Pract Res Clin Rheumatol* 2002; 16(3):361-378.
12. van der Meer A, Wulffraat NM, Prakken BJ, Gijsbers B, Rademaker CM, Sinnema G. Psychological side effects of MTX treatment in juvenile idiopathic arthritis: a pilot study. *Clin Exp Rheumatol* 2007; 25(3):480-485.
13. Brunner HI, Johnson AL, Barron AC, Passo MH, Griffin TA, Graham TB et al. Gastrointestinal symptoms and their association with health-related quality of life of children with juvenile rheumatoid arthritis: validation of a gastrointestinal symptom questionnaire. *J Clin Rheumatol* 2005; 11(4):194-204.
14. Beukelman T, Patkar NM, Saag KG, Tolleson-Rinehart S, Cron RQ, DeWitt EM et al. 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: initiation and safety monitoring of therapeutic agents for the treatment of arthritis and systemic features. *Arthritis Care Res (Hoboken)* 2011; 63(4):465-482.
15. Horneff G. Update on biologicals for treatment of juvenile idiopathic arthritis. *Expert Opin Biol Ther* 2013.
16. Cobb J, Cule E, Moncrieffe H, Hinks A, Ursu S, Patrick F et al. Genome-wide data reveal novel genes for methotrexate response in a large cohort of juvenile idiopathic arthritis cases. *Pharmacogenomics J* 2014.
17. Bulatovic M, Heijstek MW, Van Dijkhuizen EHP, Wulffraat NM, Pluijm SM, de Jonge R. Prediction of clinical non-

- response to methotrexate treatment in juvenile idiopathic arthritis. *Ann Rheum Dis* 2012; 71(9):1484-1489.
18. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York: Springer; 2009.
 19. Van Dijkhuizen EHP, Bulatovic Calasan M, Pluijm SMF, De Rotte MCFJ, Vastert SJ, Kamphuis SMM et al. Prediction of Methotrexate Intolerance in Juvenile Idiopathic Arthritis: a prospective, observational cohort study. *Pediatr Rheumatol Online J* 2015; 13:5.
 20. Malik F, Ranganathan P. Methotrexate pharmacogenetics in rheumatoid arthritis: a status report. *Pharmacogenomics* 2013; 14(3):305-314.
 21. Albertioni F, Flato B, Seideman P, Beck O, Vinje O, Peterson C et al. Methotrexate in juvenile rheumatoid arthritis. Evidence of age dependent pharmacokinetics. *Eur J Clin Pharmacol* 1995; 47(6):507-511.
 22. de Rotte MC, Bulatovic M, Heijstek MW, Jansen G, Heil SG, van Schaik RH et al. ABCB1 and ABCC3 gene polymorphisms are associated with first-year response to methotrexate in juvenile idiopathic arthritis. *J Rheumatol* 2012; 39(10):2032-2040.
 23. Masi L, Ricci L, Zulian F, Del Monte F, Simonini G, Capannini S et al. Serum osteopontin as a predictive marker of responsiveness to methotrexate in juvenile idiopathic arthritis. *J Rheumatol* 2009; 36(10):2308-2313.
 24. Ruperto N, Ravelli A, Falcini F, Lepore L, De SR, Zulian F et al. Performance of the preliminary definition of improvement in juvenile chronic arthritis patients treated with methotrexate. Italian Pediatric Rheumatology Study Group. *Ann Rheum Dis* 1998; 57(1):38-41.
 25. Consolaro A, Ruperto N, Bazso A, Pistorio A, Magni-Manzoni S, Filocamo G et al. Development and validation of a composite disease activity score for juvenile idiopathic arthritis. *Arthritis Rheum* 2009; 61(5):658-666.
 26. Horneff G, Becker I. Definition of improvement in juvenile idiopathic arthritis using the Juvenile Arthritis Disease Activity Score. *Rheumatology (Oxford)* 2014.
 27. Ringold S, Bittner R, Neogi T, Wallace CA, Singer NG. Performance of rheumatoid arthritis disease activity measures and juvenile arthritis disease activity scores in polyarticular-course juvenile idiopathic arthritis: Analysis of their ability to classify the American College of Rheumatology pediatric measures of response and the preliminary criteria for flare and inactive disease. *Arthritis Care Res (Hoboken)* 2010; 62(8):1095-1102.
 28. Bulatovic Calasan M, den Boer E, de Rotte MC, Vastert SJ, Kamphuis S, de Jonge R et al. Methotrexate polyglutamates in erythrocytes are associated with lower disease activity in juvenile idiopathic arthritis patients. *Ann Rheum Dis* 2013.
 29. Moncrieffe H, Ursu S, Holzinger D, Patrick F, Kassoumeri L, Wade A et al. A subgroup of juvenile idiopathic arthritis patients who respond well to methotrexate are identified by the serum biomarker MRP8/14 protein. *Rheumatology (Oxford)* 2013; 52(8):1467-1476.
 30. Tukova J, Chladek J, Hroch M, Nemcova D, Hoza J, Dolezalova P. 677TT genotype is associated with elevated risk of methotrexate (MTX) toxicity in juvenile idiopathic arthritis: treatment outcome, erythrocyte concentrations of MTX and folates, and MTHFR polymorphisms. *J Rheumatol* 2010; 37(10):2180-2186.
 31. Ravelli A, Viola S, Migliavacca D, Ruperto N, Pistorio A, Martini A. The extended oligoarticular subtype is the best predictor of methotrexate efficacy in juvenile idiopathic arthritis. *J Pediatr* 1999; 135(3):316-320.
 32. Hinks A, Moncrieffe H, Martin P, Ursu S, Lal S, Kassoumeri L et al. Association of the 5-aminoimidazole-4-carboxamide ribonucleotide transformylase gene with response to methotrexate in juvenile idiopathic arthritis. *Ann Rheum Dis* 2011; 70(8):1395-1400.

33. Moncrieffe H, Hinks A, Ursu S, Kassoumeri L, Etheridge A, Hubank M et al. Generation of novel pharmacogenomic candidates in response to methotrexate in juvenile idiopathic arthritis: correlation between gene expression and genotype. *Pharmacogenet Genomics* 2010; 20(11):665-676.
34. Bartoli M, Taro M, Magni-Manzoni S, Pistorio A, Traverso F, Viola S et al. The magnitude of early response to methotrexate therapy predicts long-term outcome of patients with juvenile idiopathic arthritis. *Ann Rheum Dis* 2008; 67(3):370-374.
35. Cespedes-Cruz A, Gutierrez-Suarez R, Pistorio A, Ravelli A, Loy A, Murray KJ et al. Methotrexate improves the health-related quality of life of children with juvenile idiopathic arthritis. *Ann Rheum Dis* 2008; 67(3):309-314.
36. Vilca I, Munitis PG, Pistorio A, Ravelli A, Buoncompagni A, Bica B et al. Predictors of poor response to methotrexate in polyarticular-course juvenile idiopathic arthritis: analysis of the PRINTO methotrexate trial. *Ann Rheum Dis* 2010; 69(8):1479-1483.
37. Wallace CA, Sherry DD, Mellins ED, Aiken RP. Predicting remission in juvenile rheumatoid arthritis with methotrexate treatment. *J Rheumatol* 1993; 20(1):118-122.
38. Albers HM, Wessels JA, van der Straaten RJ, Brinkman DM, Suijlekom-Smit LW, Kamphuis SS et al. Time to treatment as an important factor for the response to methotrexate in juvenile idiopathic arthritis. *Arthritis Rheum* 2009; 61(1):46-51.
39. Yanagimachi M, Naruto T, Hara T, Kikuchi M, Hara R, Miyamae T et al. Influence of polymorphisms within the methotrexate pathway genes on the toxicity and efficacy of methotrexate in patients with juvenile idiopathic arthritis. *Br J Clin Pharmacol* 2011; 71(2):237-243.
40. Schmeling H, Biber D, Heins S, Horneff G. Influence of methylenetetrahydrofolate reductase polymorphisms on efficacy and toxicity of methotrexate in patients with juvenile idiopathic arthritis. *J Rheumatol* 2005; 32(9):1832-1836.
41. Ravelli A, Ramenghi B, Di Fuccia G, Ruperto N, Zonta L, Martini A. Factors associated with response to methotrexate in systemic-onset juvenile chronic arthritis. *Acta Paediatr* 1994; 83(4):428-432.
42. Patil P, Rawcliffe C, Olaleye A, Moore S, Fox A, Sen D et al. Comparison of methotrexate induced nausea and vomiting between adolescent and adult patients with inflammatory arthritis. *Rheumatology* 2012; 51:iii107-iii108.
43. Klein A, Kaul I, Foeldvari I, Ganser G, Urban A, Horneff G. Efficacy and safety of oral and parenteral methotrexate therapy in children with juvenile idiopathic arthritis: an observational study with patients from the German Methotrexate Registry. *Arthritis Care Res (Hoboken)* 2012; 64(9):1349-1356.

Chapter

8

HIGH PREVALENCE OF METHOTREXATE INTOLERANCE IN JUVENILE IDIOPATHIC ARTHRITIS: DEVELOPMENT AND VALIDATION OF A METHOTREXATE INTOLERANCE SEVERITY SCORE

Maja Bulatović, Marloes W. Heijstek,
Marleen Verkaaik, E.H. Pieter van Dijkhuizen,
Wineke Armbrust, Esther P.A. Hoppenreijts,
Sylvia Kamphuis, Wietse Kuis,
Toine C.G. Egberts, Gerben Sinnema,
Carin M.A. Rademaker and Nico M. Wulffraat

ABSTRACT

Objective

To design and validate a new questionnaire for identifying patients with methotrexate (MTX) intolerance, and to determine the prevalence of MTX intolerance in patients with juvenile idiopathic arthritis (JIA) using this questionnaire.

Methods

The MTX Intolerance Severity Score (MISS) questionnaire was constructed, consisting of 5 domains: stomach ache, nausea, vomiting, sore mouth, and behavioral symptoms. The domains each consisted of 3 questions pertaining to the presence of a symptom upon, prior to (anticipatory), and when thinking of (associative) MTX intake. The MISS questionnaire was validated in 86 patients by determining its discriminative power between patients with and those without MTX intolerance, identified as such by a gold standard (physician's opinion). Using the MISS questionnaire, the prevalence of MTX intolerance was determined in 297 JIA patients.

Results

The MISS questionnaire discriminated well between MTX-intolerant and MTX-tolerant patients. A cutoff score of 6 yielded the best sensitivity (88%) and specificity (80%). MTX intolerance was found in 150 (50.5%) of 297 patients. Of 220 patients receiving oral MTX, 98 (44.5%) experienced MTX intolerance, whereas 67.5% of 77 patients receiving parenteral MTX experienced intolerance to the drug ($P = 0.001$).

Conclusion

Our findings indicate that the MISS questionnaire is a highly sensitive and specific tool for the diagnosis of MTX intolerance, and that there is a high prevalence of MTX intolerance among JIA patients. The prevalence of intolerance in patients receiving parenteral MTX exceeds that in patients receiving oral MTX. The frequent occurrence of anticipatory and associative symptoms suggests that classic conditioning plays an important role in MTX intolerance.



INTRODUCTION

Juvenile idiopathic arthritis (JIA), an autoimmune disease characterized by chronic arthritis in one or more joints, is one of the most common chronic diseases in childhood, with a reported prevalence of between 16 and 159 per 100,000.¹ It encompasses various subtypes whose severity and clinical course differ.¹

Methotrexate (MTX) is the first-choice disease-modifying anti-rheumatic drug (DMARD) for the treatment of JIA. It is an effective drug that induces disease remission in >70% of patients.^{2,3} Serious adverse effects, such as hepatotoxicity and bone marrow suppression, are infrequent and usually transient if MTX is stopped.⁴ However, gastrointestinal adverse effects, which include nausea, abdominal pain, vomiting, or diarrhea, are common during MTX treatment.^{5,6} Folic acid supplementation is one of the accepted strategies to treat and prevent the development of these adverse effects.⁷⁻¹⁰ Despite folic acid use, many JIA patients experience gastrointestinal adverse effects after MTX intake. JIA patients also develop anticipatory and associative gastrointestinal adverse effects occurring before MTX intake and when thinking of MTX as well as behavioral symptoms, such as restlessness and crying, when taking MTX.⁶ These adverse effects arise as a conditioned response to the above-mentioned physical symptoms experienced after MTX intake.

Anticipatory nausea and vomiting, related to strongly emetogenic chemotherapy treatments, are well-known conditioned responses in cancer patients, with a reported frequency of 30%.^{11;12} In contrast to cancer patients, this problem has remained largely unrecognized in JIA patients. Although these adverse effects, in concert with behavioral symptoms, could lead to refusal and premature discontinuation of an otherwise efficacious and safe drug, their type and frequency are poorly elucidated.

The objectives of this study were to design and validate a new questionnaire for MTX-related gastrointestinal and behavioral symptoms, termed MTX intolerance, and to use this questionnaire to determine the prevalence of MTX intolerance in patients with JIA.

PATIENTS AND METHODS

Study design, study population, and data collection

We performed a cross-sectional descriptive study in 4 University Medical Centers (UMC) in The Netherlands that have pediatric rheumatology departments (Utrecht, Nijmegen, Groningen, and Rotterdam). The MTX intolerance study was approved by the local medical ethics committees, and it was performed under good clinical practice conditions. The study population consisted of patients between 2 and 18 years of age, with a confirmed JIA diagnosis according to the International League of Associations for Rheumatology criteria,¹³ who were seen in the outpatient clinic between August 2007 and June 2009. All patients had been receiving either oral or parenteral MTX for at least

3 months at the time of inclusion. Patients with a history of noncompliance to earlier treatments unrelated to MTX were excluded. Written informed consent was obtained from all patients and/or their parents.

Demographic data, such as sex, JIA subtype, and age at disease onset, were obtained from medical records. Information on duration of MTX use, MTX route of administration (oral or parenteral), and MTX dosage (mg/m²/week) was obtained. The physician's global assessment of disease activity (on a 0-10 scale) at the time the questionnaire was completed was documented. We acquired information on the concomitant use of folic acid, nonsteroidal anti-inflammatory drugs (NSAIDs), oral steroids (mg/kg), and other DMARDs, such as anti-tumor necrosis factor α therapy and anti-interleukin-1 receptor blockade.

MTX Intolerance Severity Score

To determine the prevalence of MTX intolerance, we designed the MTX Intolerance Severity Score (MISS) questionnaire. MTX intolerance included gastrointestinal adverse effects and behavioral symptoms occurring after MTX intake, before MTX intake (anticipatory symptoms), and/or when thinking of taking the medication (associative symptoms). Behavioral symptoms were included in the MISS questionnaire since they often develop in response to MTX-induced gastrointestinal symptoms or the anticipation thereof. The MISS questionnaire was constructed by 2 physicians (MWH and NMW) and a psychologist (GS), based on their extensive clinical experience with JIA patients with MTX intolerance. The initial MISS questionnaire contained 16 items divided into 5 domains, namely, abdominal pain (stomach ache), nausea, vomiting, oral pain (sore mouth), and behavioral symptoms. The first 4 domains contained 3 items each, pertaining to adverse effects experienced after MTX intake as well as anticipatory and associative adverse effects. The fifth domain, on behavioral symptoms, included restlessness, crying, irritability, and refusal of MTX. On each item a patient could score 0 (no symptoms), 1 point (mild symptoms), 2 points (moderate symptoms), or 3 points (severe symptoms).

Validation of the MISS questionnaire

The initial MISS questionnaire was validated in JIA patients from UMC Utrecht by determining its ability to discriminate between patients with and those without MTX intolerance as established by the gold standard. Since there were no available instruments designed to measure MTX intolerance, a physician's opinion on the presence or absence of MTX intolerance was considered the gold standard. The physician indicated that a patient was intolerant to MTX if the patient or the patient's parents confirmed the following during a short interview at the outpatient clinic: the presence of gastrointestinal symptoms before and after MTX intake, a persistent nature of these symptoms, and/or a negative

effect of these symptoms in the days after MTX intake. The physician was blinded with regard to the results of the questionnaire, and the patients were blinded with regard to the physician's opinion.

To analyze whether the MISS questionnaire could significantly ($P \leq 0.05$) discriminate between patients with and those without MTX intolerance, we compared the scores on each item of the questionnaire between the 2 groups, using the Mann-Whitney U test. Next, we removed the items that were not significantly discriminative between patients with and those without MTX intolerance. Additionally, to evaluate the reliability of the MISS questionnaire, its homogeneity was assessed by item-analysis using Cronbach's alpha. Items that lowered the homogeneity of the MISS questionnaire were removed to ensure optimal homogeneity. Subsequently, we determined whether the total score on the modified MISS questionnaire could discriminate between patients with and those without MTX intolerance. To evaluate the discriminant validity of the modified questionnaire, the area under the receiver operating characteristic (ROC) curve and the 95% confidence interval (95% CI) were computed. Next, we determined the optimal cutoff score on the MISS questionnaire for classifying patients as either tolerant or intolerant to MTX. Sensitivity, specificity, the sum of sensitivity and specificity, positive predictive value (PPV), and negative predictive value (NPV) were computed for 10 cutoff values for the MISS questionnaire. The cutoff value with the maximum score for the sum of sensitivity and specificity was considered the optimal cutoff point.

Prevalence of MTX intolerance

We determined the prevalence of MTX intolerance in JIA patients from all 4 UMCs, using the validated MISS questionnaire with the optimal cutoff score. Furthermore, scores on the MISS questionnaire and the prevalence of MTX intolerance in patients receiving oral MTX were compared with those in patients receiving parenteral MTX, by chi-square test. The Mann-Whitney U test was used to determine whether the total MISS and the scores for the separate domains of the MISS, as well as MTX dose, JIA subtype, physician's global assessment of disease activity, disease duration, and duration of MTX use, differed between MTX-intolerant patients who were receiving oral MTX and those who were receiving parenteral MTX.

To evaluate associations of MTX intolerance with clinically relevant covariates, such as the route of MTX administration, MTX dose, co-medication (oral steroids, NSAIDs, anti-emetics, biologic agents, or other DMARDs), JIA subtype, age, physician's global assessment of disease activity, disease duration, and duration of MTX use, we performed a multivariate logistic regression using a P value of >0.10 as a removal criterion and a P value of ≤ 0.05 as an inclusion criterion. Odds ratios (ORs) and 95% CIs were calculated to express these associations. Statistical analyses were carried out using SPSS, version 15.0.1.

RESULTS

The MISS questionnaire discriminates between patients with and those without MTX intolerance

For the purpose of validation of the MISS questionnaire, 89 patients attending the UMC Utrecht outpatient clinic completed the questionnaire. Three patients did not meet the inclusion criteria and were therefore excluded. According to the gold standard (physician's opinion), 29.1% of the patients experienced MTX intolerance.

All items of the MISS questionnaire were discriminative between patients with and those without MTX intolerance as determined by the gold standard (data not shown), with the exception of "vomiting when thinking of MTX" (associative vomiting) and the entire domain of oral pain. These 4 items also reduced the homogeneity of the MISS questionnaire, and removal of these items resulted in a high Cronbach's alpha of 0.919, which confirmed the reliability of the questionnaire. Upon exclusion of the 4 non-discriminative items (i.e., associative vomiting and the oral pain domain), the total score of the modified 12-item MISS questionnaire remained discriminative between MTX-intolerant patients (median score 13 [interquartile range 7.5–18.5]) and MTX-tolerant patients (median score 1 [interquartile range 0-4]). The modified MISS questionnaire had a minimum possible score of 0 points and a maximum possible score of 36 points. The modified 12-item MISS questionnaire satisfies the criteria of feasibility, face validity, and content validity since it is short and easy to complete by patients or parents, easy to interpret by physicians, and contains all relevant aspects of MTX intolerance as established by a consensus of expert rheumatologists (WA, EPAH, SK, WK, and NMW). This questionnaire was used for further analysis and will be referred to as the MISS questionnaire (see Supplementary table 1).

The area under the ROC curve was 0.90 (95% CI 0.83–0.96), indicating that 90% of the patients were classified correctly with the MISS questionnaire. Based on the scores obtained, we expected to find the most discriminative value between 2 and 11 points. Table 1 shows sensitivity, specificity, PPV, and NPV for the scores between 2 and 11 points. Cutoff scores of 5 and 6 showed the highest sum score of sensitivity (88%) and specificity (80%), namely, 168%. At these cutoff points, 88% of the patients diagnosed as having MTX intolerance according to the gold standard would also be identified as intolerant to MTX according to the MISS questionnaire. Furthermore, 80% of the patients who were tolerant to MTX according to the gold standard would indeed be identified as MTX tolerant according to the MISS questionnaire. At these cutoff points, 65% of patients who were intolerant to MTX (PPV) and 94% who were tolerant to MTX (NPV) would be correctly diagnosed. The diagnostic parameters of the cutoff scores 5 and 6 were equal, since there were no patients in the validation cohort with a score of 5. Since MTX-intolerant patients scored high on gastrointestinal symptoms before MTX intake and when thinking of MTX intake, we defined MTX intolerance as a total score of ≥ 6 with at least 1 point on anticipatory and/or associative and/or behavioral symptoms.

Table 1. Sensitivity, specificity, PPV, and NPV for various cutoff scores on the MISS questionnaire*

Cutoff score	Sensitivity	Specificity	Σ†	PPV	NPV
2	100	59	159	50	100
3	92	67	159	53	95
4	92	72	164	58	96
5	88	80	168‡	65	94
6	88	80	168‡	65	94
7	84	81	164	64	92
8	76	82	158	63	89
9	64	84	148	62	85
10	60	90	150	71	85
11	60	90	150	71	85

* Values are the %. The Methotrexate Intolerance Severity Score (MISS) questionnaire was validated in a cohort of 86 patients. PPV = positive predictive value; NPV = negative predictive value.

† Sum of sensitivity and specificity.

‡ Optimal sum of sensitivity and specificity.

High prevalence of MTX intolerance in JIA patients

Two hundred ninety-seven patients completed the MISS questionnaire during outpatient visits. All 86 patients in the validation cohort were included in the larger cohort. There were no significant differences between the validation cohort patients and the JIA patients included in this cohort with respect to sex distribution, JIA subtype, or MTX dose (data not shown). Table 2 shows the baseline characteristics of the patients.

We found that 150 (50.5%) of the patients had MTX intolerance according to the above-mentioned definition (Table 3). Furthermore, the frequency of gastrointestinal adverse events in each domain was significantly higher in patients diagnosed as having MTX intolerance compared to patients without MTX intolerance. The percentages of patients diagnosed as having MTX intolerance who had nausea and behavioral symptoms were 91.3% and 88.7%, respectively. Twenty-eight (18.7%) of the MTX-intolerant patients experienced anticipatory vomiting, whereas this symptom did not occur in patients considered to be tolerant to MTX (Table 3). Twenty-three (56.1%) of the 41 patients who were taking anti-emetics were intolerant to MTX.

Patients with MTX intolerance were receiving a slightly higher MTX dosage than patients without MTX intolerance (10.9 mg/m²/week and 9.8 mg/m²/week, respectively [P = 0.002]). Furthermore, MTX-intolerant patients had longer disease durations and received MTX for longer periods than did MTX-tolerant patients (median disease duration 4.3 years in MTX-intolerant patients versus 3.0 years in MTX-tolerant patients [P = 0.026] and median duration of MTX use 2.0 years in MTX-intolerant patients versus 1.2 years in MTX-tolerant patients [P = 0.001]). Patients with MTX intolerance were also somewhat younger than those without MTX intolerance (median age 11 years versus 12 years

Table 2. Baseline characteristics of the 297 patients at the time of completing the MISS questionnaire*

Sex, female	204 (68.7)
JIA subtype	
Oligoarticular JIA, persistent	79 (26.6)
Oligoarticular JIA, extended	50 (16.8)
Polyarticular JIA, RF positive	20 (6.7)
Polyarticular JIA, RF negative	95 (32)
Systemic-onset JIA	31 (10.4)
Enthesitis-related JIA	5 (1.7)
Psoriatic arthritis	17 (5.7)
Disease characteristics	
ANA positive†	141 (50.0)
RF positive‡	21 (9.9)
HLA-B27 positive§	21 (22.3)
Chronic iridocyclitis	52 (17.5)
Age, mean ± SD years	10.9 ± 3.9
JIA duration, mean ± SD years	4.9 ± 4.1
Duration of MTX use, median (IQR) years	1.6 (0.6-3.6)
Disease activity, toxicity, and medication use	
Physician's global assessment of disease activity, median (IQR) (0-10 scale)	0.5 (0-2)
MTX dose (mg/m ² /week), median (IQR)¶	10.2 (8.1-12.8)
Oral route of MTX administration	220 (74.1)
Folic acid	295 (99.3)
NSAIDs	189 (63.6)
Oral steroids	27 (9.1)
Antiemetics	41 (13.8)
Other DMARDs#	49 (16.5)

* Except where indicated otherwise, values are the number (%) of patients. MISS = Methotrexate Intolerance Severity Score; JIA = juvenile idiopathic arthritis; IQR = interquartile range; NSAIDs = nonsteroidal anti-inflammatory drugs.

† Antinuclear antibody (ANA) status was determined in 282 patients.

‡ Rheumatoid factor (RF) status was determined in 212 patients.

§ HLA-B27 status was determined in 94 patients.

¶ Data on methotrexate (MTX) dose (mg/m²) were available for 292 patients.

Of the 49 patients taking other disease-modifying anti-rheumatic drugs (DMARDs), 32 were taking etanercept, 9 were taking anakinra, 4 were taking adalimumab, and 4 were taking infliximab.

[P = 0.015]). JIA subtype and physician's global assessment of disease activity did not differ significantly between the 2 groups.

Higher prevalence of MTX intolerance in patients receiving parenteral MTX than in patients receiving oral MTX

We determined the prevalence of MTX intolerance in patients receiving oral MTX and in patients receiving parenteral MTX. Table 3 shows that the prevalence of MTX intolerance

Table 3. Overall and per domain prevalence of MTX-related gastrointestinal symptoms in all patients and by route of MTX administration*

	Tolerance to MTX	Intolerance to MTX	Oral MTX	Parenteral MTX
Total	147 (49.5)	150 (50.5)	220 (74.1)	77 (25.9)
Cutoff score ≥ 6	0 (0)	150 (100)	98 (44.5)	52 (67.5) [†]
Abdominal pain	41 (27.9)	111 (74.0)	72 (73.5)	39 (75.0)
After MTX	34 (23.1)	100 (66.7)	66 (67.3)	34 (65.4)
Anticipatory and associative	12 (8.2)	85 (56.7)	54 (55.1)	31 (59.6)
Nausea	54 (36.7)	137 (91.3)	90 (91.8)	47 (90.4)
After MTX	43 (29.3)	126 (84.0)	83 (84.7)	43 (82.7)
Anticipatory and associative	22 (15.0)	116 (77.3)	75 (76.5)	41 (78.8)
Vomiting	7 (4.8)	74 (49.3)	45 (45.9)	29 (55.8)
After MTX	7 (4.8)	73 (48.7)	45 (45.9)	28 (53.8)
Anticipatory	0 (0)	28 (18.7)	14 (14.3)	14 (26.9)
Behavioral symptoms	39 (26.5)	133 (88.7)	82 (83.7)	51 (98.1) [‡]
Restlessness	22 (15.0)	116 (77.3)	69 (70.4)	47 (90.4) [§]
Crying	2 (1.4)	69 (46.0)	38 (38.8)	31 (59.6) [¶]
Irritability	22 (15.0)	105 (70.0)	67 (68.4)	38 (73.1)
Refusal of MTX	22 (15.0)	116 (77.3)	69 (70.4)	47 (90.4) [§]

* Values are the number (%) of patients. Intolerance to methotrexate (MTX) was defined as a score of ≥ 6 on the MTX Intolerance Severity Score questionnaire. For tolerance and intolerance, the percentages for each symptom are out of the total number of patients; for oral MTX and parenteral MTX, the percentages for each symptom are out of the number of patients with a cutoff score of ≥ 6 . Anticipatory refers to before MTX intake, and associative refers to when thinking of MTX.

[†] P = 0.001 versus oral MTX, by chi-square test.

[‡] P = 0.008 versus oral MTX, by chi-square test.

[§] P = 0.005 versus oral MTX, by chi-square test.

[¶] P = 0.015 versus oral MTX, by chi-square test.

was significantly higher in patients who were receiving parenteral MTX than in patients who were receiving oral MTX (67.5% and 44.5%, respectively [P = 0.001]). Significantly more patients who were receiving parenteral MTX than those who were receiving oral MTX refused to take the drug (P = 0.005) and exhibited restlessness (P = 0.005) as well as crying (P = 0.015) when taking MTX (Table 3). More patients who were receiving parenteral MTX than patients who were receiving oral MTX experienced vomiting after administration of MTX and anticipatory vomiting as well as anticipatory and associative abdominal pain and nausea, though these differences were not statistically significant. Although the prevalence of MTX intolerance was higher in patients receiving parenteral MTX, the severity of MTX intolerance did not differ between the 2 groups; the median score on the MISS questionnaire was 12 points for both patients receiving oral MTX and patients receiving parenteral MTX.

In The Netherlands, it is common clinical practice to initially treat patients with oral MTX; those patients in need of higher dosages to reach clinical efficacy are treated with

parenteral MTX. As expected, patients receiving parenteral MTX received a higher median MTX dosage (13.5 mg/m²/week) than patients receiving oral MTX (9.6 mg/m²/week) ($P < 0.001$). JIA subtype, physician's global assessment of disease activity, duration of JIA, and duration of MTX use did not differ between the 2 groups. In addition, the median age of MTX-intolerant patients who were receiving oral MTX (10.7 years) was similar to that of MTX-intolerant patients who were receiving parenteral MTX (10.0 years).

In a multivariate logistic regression analysis, MTX intolerance was associated with the route of MTX administration only. Patients receiving parenteral MTX had higher odds of having MTX intolerance than patients receiving oral MTX (OR 1.9 [95% CI 1.01-3.58], $P = 0.046$). Moreover, patients receiving a higher MTX dose showed a trend toward increased odds of having MTX intolerance (OR 1.08 [95% CI 1.00-1.16], $P = 0.051$). Significant associations were not found between MTX intolerance and other covariates, such as age, co-medication, JIA subtype, physician's global assessment of disease activity, disease duration, and duration of MTX use.

8

DISCUSSION

We designed and validated the MISS questionnaire. Using this tool we determined that the prevalence of MTX intolerance reached 50.5% in JIA patients.

To date, the only instrument used to determine gastrointestinal symptoms in JIA patients is the Gastrointestinal Symptom Scale for Kids (GISSK).¹⁴ The GISSK assesses the presence and the severity of gastrointestinal symptoms, such as heartburn, bloating, nausea, vomiting, stomach ache, stool consistency, and loss of appetite. In contrast to the GISSK, which screens for gastrointestinal side effects in JIA patients in general and irrespective of treatment, the MISS questionnaire assesses gastrointestinal adverse effects related to MTX use specifically. Furthermore, unlike the GISSK, the MISS questionnaire considers the occurrence of MTX-related anticipatory and associative gastrointestinal adverse effects and behavioral symptoms. The MISS questionnaire is therefore a more appropriate tool to determine the frequency of such symptoms related to MTX intake.

We found a notably high prevalence of MTX intolerance. Several studies have demonstrated that gastrointestinal adverse effects occur frequently in JIA patients receiving MTX.^{5;14;15} Three previous studies determined the prevalence of MTX-related gastrointestinal adverse effects, which ranged from 10% to 21% depending on the type of gastrointestinal symptoms.^{3;16;17} The substantially higher prevalence of MTX intolerance in our study compared to the aforementioned findings could have several explanations. We used a specific definition to determine the prevalence of MTX intolerance, which included gastrointestinal symptoms before and after MTX intake as well as behavioral symptoms. Previously, these frequently occurring symptoms were not taken into account, which could have resulted in the higher prevalence of MTX intolerance in our study. Moreover,

we determined the prevalence of MTX intolerance with a structured questionnaire. This could have led the patients and their parents to consider the occurrence of MTX-related symptoms more thoroughly, leading to a higher prevalence.

The prevalence of MTX intolerance was 23% higher in patients receiving parenteral MTX than in patients receiving oral MTX. The higher prevalence of MTX intolerance in patients receiving parenteral MTX originates from a higher prevalence of behavioral symptoms as well as anticipatory and associative abdominal pain, nausea, and vomiting in these patients, although the differences in the prevalence of these gastrointestinal effects were not great. In addition to the aversion to MTX, fear of needles likely contributed to a higher prevalence of these adverse effects and in turn to a higher prevalence of MTX intolerance in patients receiving parenteral MTX.¹⁸ However, the prevalence of these adverse effects in the absence of needles, namely in patients receiving oral MTX, is also very high. This shows that aversion to MTX plays an important role in the induction of anticipatory and associative gastrointestinal adverse effects and behavioral problems for both routes of administration.

The frequent occurrence of the above-mentioned symptoms before MTX intake supports the notion that classic conditioning mechanisms play an important role in the development of MTX intolerance. In classic conditioning terms, an unconditioned stimulus (i.e., MTX) produces an unconditioned response, such as abdominal pain, nausea, or vomiting, once MTX is administered. At that moment, many potential conditioned stimuli are present. The most commonly reported conditioned stimuli at our outpatient clinic are the yellow color of the pill or injection fluid, and the liquid with which MTX is administered, such as orange juice or water. After a number of weeks of taking MTX, the conditioned stimuli mentioned above lead to the conditioned response of anticipatory and associative adverse effects. It is thought that these pretreatment adverse effects occur as a result of stimulation of the higher centers of the central nervous system if a patient experiences the conditioned stimuli.¹⁹ Therefore, cognitive-behavioral therapy may be beneficial in treating JIA patients with MTX intolerance.⁶

Anticipatory and associative adverse effects are clinically not very evident. Consequently, these symptoms cannot easily be detected by physician assessment only, but can be detected with the MISS questionnaire. Therefore, using the MISS questionnaire is advantageous, since it allows early detection of MTX intolerance symptoms, which could create a window of opportunity for the treatment of MTX intolerance before the conditioned response becomes so severe that it negatively affects the patient's quality of life or leads to early cessation of an otherwise effective MTX treatment.

Interestingly, post-treatment gastrointestinal adverse effects were equally prevalent in patients receiving oral MTX and those receiving parenteral MTX. This suggests that MTX, regardless of the route of administration, uses the same mechanism to induce post-treatment gastrointestinal symptoms. Nausea and vomiting after treatment with cytostatic agents occur as a consequence of stimulation of the chemoreceptor trigger zone, which

in turn stimulates the vomiting center in the medulla.²⁰ It is, however, unknown whether low-dose MTX treatment in JIA can stimulate the chemoreceptor trigger zone, thereby causing nausea and vomiting in these patients.

The findings of the present study do not indicate which clinical variables (i.e., MTX dose, age) are associated with the development of MTX intolerance and how many patients discontinue MTX due to intolerance. Large prospective trials following up JIA patients from the start of MTX treatment are needed to answer such questions. We are presently gathering a large prospective JIA cohort in order to address such issues.

To our knowledge, this is the first study to demonstrate a high prevalence of MTX intolerance in JIA patients, using a newly designed MISS questionnaire. Since the extent of MTX intolerance in JIA patients often remains unrecognized in clinical practice, we suggest that clinicians use the MISS questionnaire in their assessment of the frequency, severity, and type of MTX intolerance symptoms. We showed that anticipatory and associative gastrointestinal adverse effects as well as behavioral symptoms occurred frequently, strongly suggesting that classic conditioning plays a central role in the development of MTX intolerance. Therefore, we are currently performing a randomized controlled trial to compare the effect of cognitive-behavioral therapy on MTX intolerance with the effect of a switch to parenteral MTX or a continuation of oral MTX with a concomitant antiemetic.

ACKNOWLEDGMENT

We thank R. K. Stellato for contributing to the preparation of this manuscript.

FUNDING

Supported by MEDAC GmbH, Germany (unrestricted grant to Dr. Wulffraat) and the Dutch Arthritis Association (grant NR 07-01-402 to Dr. Wulffraat).

SUPPLEMENTARY MATERIAL

Supplementary table 1. Methotrexate intolerance severity score – MISS.

	NO complaints	COMPLAINTS (score 1-3 points)		
		Mild	Moderate	Severe
		0	1	2
STOMACHACHE				
· My child has a stomachache after taking MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
· My child has a stomachache several hours to one day before taking MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
· My child has a stomachache when thinking of MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
NAUSEA				
· My child is nauseous after taking MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
· My child is nauseous several hours to one day before taking MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
· My child is nauseous when thinking of MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
VOMITING				
· My child vomits after taking MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
· My child vomits hours to one day before taking MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
BEHAVIORAL COMPLAINTS				
· My child is restless when taking MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
· My child cries when taking MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
· My child is irritable when taking MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
· My child refuses to take MTX	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



REFERENCE LIST

1. Ravelli A, Martini A. Juvenile idiopathic arthritis. *Lancet* 2007; 369(9563):767-778.
2. Foell D, Wulfraat N, Wedderburn LR, Wittkowski H, Frosch M, Gerss J et al. Methotrexate withdrawal at 6 vs 12 months in juvenile idiopathic arthritis in remission: a randomized clinical trial. *JAMA* 2010; 303(13):1266-1273.
3. Ruperto N, Murray KJ, Gerloni V, Wulfraat N, de Oliveira SK, Falcini F et al. A randomized trial of parenteral methotrexate comparing an intermediate dose with a higher dose in children with juvenile idiopathic arthritis who failed to respond to standard doses of methotrexate. *Arthritis Rheum* 2004; 50(7):2191-2201.
4. Ortiz-Alvarez O, Morishita K, Avery G, Green J, Petty RE, Tucker LB et al. Guidelines for blood test monitoring of methotrexate toxicity in juvenile idiopathic arthritis. *J Rheumatol* 2004; 31(12):2501-2506.
5. Murray KJ, Lovell DJ. Advanced therapy for juvenile arthritis. *Best Pract Res Clin Rheumatol* 2002; 16(3):361-378.
6. van der Meer A, Wulfraat NM, Prakken BJ, Gijbbers B, Rademaker CM, Sinnema G. Psychological side effects of MTX treatment in juvenile idiopathic arthritis: a pilot study. *Clin Exp Rheumatol* 2007; 25(3):480-485.
7. Alarcon GS, Morgan SL. Folinic acid to prevent side effects of methotrexate in juvenile rheumatoid arthritis. *J Rheumatol* 1996; 23(12):2184-2185.
8. Ravelli A, Migliavacca D, Viola S, Ruperto N, Pistorio A, Martini A. Efficacy of folinic acid in reducing methotrexate toxicity in juvenile idiopathic arthritis. *Clin Exp Rheumatol* 1999; 17(5):625-627.
9. van Ede AE, Laan RF, Rood MJ, Huizinga TW, van de Laar MA, van Denderen CJ et al. Effect of folic or folinic acid supplementation on the toxicity and efficacy of methotrexate in rheumatoid arthritis: a forty-eight week, multicenter, randomized, double-blind, placebo-controlled study. *Arthritis Rheum* 2001; 44(7):1515-1524.
10. Visser K, Katchamart W, Loza E, Martinez-Lopez JA, Salliot C, Trudeau J et al. Multinational evidence-based recommendations for the use of methotrexate in rheumatic disorders with a focus on rheumatoid arthritis: integrating systematic literature research and expert opinion of a broad international panel of rheumatologists in the 3E Initiative. *Ann Rheum Dis* 2009; 68(7):1086-1093.
11. Figueroa-Moseley C, Jean-Pierre P, Roscoe JA, Ryan JL, Kohli S, Palesh OG et al. Behavioral interventions in treating anticipatory nausea and vomiting. *J Natl Compr Canc Netw* 2007; 5(1):44-50.
12. Dolgin MJ, Katz ER, McGinty K, Siegel SE. Anticipatory nausea and vomiting in pediatric cancer patients. *Pediatrics* 1985; 75(3):547-552.
13. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004; 31(2):390-392.
14. Brunner HI, Johnson AL, Barron AC, Passo MH, Griffin TA, Graham TB et al. Gastrointestinal symptoms and their association with health-related quality of life of children with juvenile rheumatoid arthritis: validation of a gastrointestinal symptom questionnaire. *J Clin Rheumatol* 2005; 11(4):194-204.
15. Hashkes PJ, Laxer RM. Medical treatment of juvenile idiopathic arthritis. *JAMA* 2005; 294(13):1671-1684.
16. Ramanan AV, Whitworth P, Baildam EM. Use of methotrexate in juvenile idiopathic arthritis. *Arch Dis Child* 2003; 88(3):197-200.
17. Alsufyani K, Ortiz-Alvarez O, Cabral DA, Tucker LB, Petty RE, Malleson PN. The role of subcutaneous administration of methotrexate in children with juvenile idiopathic arthritis who have failed oral methotrexate. *J Rheumatol* 2004; 31(1):179-182.
18. Duff AJ. Incorporating psychological approaches into routine

- paediatric venepuncture. *Arch Dis Child* 2003; 88(10):931-937.
19. Hornby PJ. Central neurocircuitry associated with emesis. *Am J Med* 2001; 111 Suppl 8A:106S-112S.
20. Veyrat-Follet C, Farinotti R, Palmer JL. Physiology of chemotherapy-induced emesis and antiemetic therapy. Predictive models for evaluation of new compounds. *Drugs* 1997; 53(2):206-234.



Chapter

9

METHOTREXATE INTOLERANCE IN ORAL AND SUBCUTANEOUS ADMINISTRATION IN PATIENTS WITH JUVENILE IDIOPATHIC ARTHRITIS: A CROSS-SECTIONAL, OBSERVATIONAL STUDY

E. H. Pieter van Dijkhuizen, Juliëtte N. Pouw, Andrea Scheuern,
Boris Hügle, Sven Hardt, Gerd Ganser,
Jasmin Beate Kümmerle-Deschner, Gerd Horneff,
Dirk Holzinger, Maja Bulatović Čalasan and Nico M. Wulffraat

ABSTRACT

Objective

Methotrexate (MTX) is the cornerstone disease-modifying anti-rheumatic drug (DMARD) in juvenile idiopathic arthritis (JIA). In Dutch patients, MTX intolerance occurred frequently and was associated with subcutaneous (SC) administration. The aim of this study was to assess the prevalence of MTX intolerance and its association with the route of administration in a German cohort of JIA patients.

Methods

A cross-sectional study of JIA patients on MTX was performed. Primary outcome was MTX intolerance, which was determined using the validated Methotrexate Intolerance Severity Score (MISS) questionnaire. The prevalence of gastrointestinal adverse effects and MTX intolerance was compared between patients on MTX SC and PO.

Results

Of 179 JIA patients on MTX, 73 (40.8%) were intolerant. The odds of MTX intolerance were higher in patients using MTX exclusively SC compared to exclusively PO (adjusted odds ratio 3.37 [95% confidence interval 1.19-10.0]). There was strong evidence that the former experienced more behavioural complaints (76.1% versus 47.4%, $P = 0.001$) and weak evidence that they experienced more abdominal pain after MTX intake (43.5% versus 27.4%, $P = 0.056$).

Conclusion

The prevalence of MTX intolerance was high and exclusively SC administration of MTX was associated with MTX intolerance and behavioural adverse effects. The prevalence of gastrointestinal adverse effects was at least as high as in patients on MTX PO. The frequently held assumption that SC causes fewer side effects than PO seems unwarranted. Definite answers about the differences between SC and PO administration with respect to safety and efficacy should be obtained by randomised trials.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is the most common chronic rheumatic disease in childhood with a prevalence between 16 and 150 per 100,000.¹ It is defined as chronic joint inflammation of unknown aetiology, lasting at least 6 weeks, with the onset before 16 years of age. In JIA, methotrexate (MTX) is the most frequently used disease-modifying anti-rheumatic drug (DMARD)^{2,3} because it is beneficial in around 70% of the patients,^{2,4} safe and relatively inexpensive.^{5,6} It is usually administered orally (PO) or subcutaneously (SC) in a dose of 10-20 mg/m²/week.^{2,7} MTX can give rise to adverse effects including gastrointestinal complaints, such as abdominal pain, nausea and vomiting.^{3,8-12} It is thought that these originate from intestinal serotonin release or by stimulation of the chemoreceptor trigger zone (CTZ). This zone is located in the medulla oblongata and communicates with the area postrema (the vomiting centre) in the medulla. However, the exact mechanism remains unknown.^{13,14}

These gastrointestinal complaints after MTX intake in turn can lead to complaints occurring prior to MTX intake (anticipatory complaints) or when thinking of MTX (associative complaints), via what is thought to be a classical conditioning response.^{3,10,11,15} These symptoms, occurring not only after but also before MTX intake, are termed MTX intolerance. It has been observed in about 50% of the JIA patients on MTX.¹⁵ MTX intolerance influences the quality of life of JIA patients negatively.¹⁶ Moreover, up to three-quarters of MTX intolerant patients are reluctant to take MTX, potentially leading to non-compliance, inefficacy and switch to costly biologicals.^{11,15,17}

Previously it has been assumed that MTX SC causes fewer adverse effects.^{8,10,18} Furthermore, because MTX SC has a higher drug bioavailability and a greater absorption,¹⁸⁻²¹ it has been supposed that it is more efficacious, compared to MTX PO.^{10,21} However, both assumptions have been challenged recently. Firstly, in a Dutch cohort, the prevalence of MTX intolerance in patients receiving MTX SC exceeded the prevalence in patients on MTX PO significantly (67.5% vs. 44.5%, $P = 0.001$).¹⁵ Secondly, in a German observational cohort, no differences between MTX PO and SC were found in terms of efficacy or toxicity. Furthermore, MTX SC was significantly more frequently discontinued because of adverse effects.²² MTX SC may therefore not be superior to MTX PO regarding adverse effects and efficacy. These findings are clinically relevant, since they may guide physicians in their choice of the route of administration of MTX and may spare JIA patients unnecessary injections.

To date, the prevalence of MTX intolerance and its related factors have only been assessed in Dutch patients. Furthermore, the finding that MTX intolerance was more frequent in patients taking MTX SC was quite unexpected. Consequently, the question of what route of administration to choose when starting MTX in daily clinical practice has not yet been settled unequivocally. The aim of this study was to determine the prevalence of MTX intolerance in a cohort of German JIA patients and to determine whether this

prevalence was associated with the route of administration of MTX, to aid physicians in the choice of route of administration.

PATIENTS AND METHODS

Study design and population

We performed a cross-sectional multicentre study in five hospitals in Germany that have paediatric rheumatology departments (Garmisch-Partenkirchen, Sendenhorst, Tübingen, Sankt Augustin and Münster). Centres were gradually added, as the number of participants remained low. In one centre (Sankt Augustin), patients of one paediatric rheumatologist were enrolled (GH). The case mix of patients among rheumatologists in this centre was comparable. The study was approved by the local medical ethics committees, and it was performed according to good clinical practice regulations and the declaration of Helsinki. Written informed consent was obtained from all the patients and/or their parents.

All patients with a confirmed JIA diagnosis according to ILAR criteria,²³ aged between 2 and 18 years and using MTX for at least three months, either orally or subcutaneously, were eligible. The cohort contained 190 patients, who were included between August 2009 and April 2013.

MTX intolerance

To determine whether the patients experienced intolerance, they completed the validated Methotrexate Intolerance Severity Score (MISS) questionnaire, either at the outpatient clinic, or at home (Sendenhorst).¹⁵ The MISS consists of 12 questions distributed over four domains, being abdominal pain, nausea, vomiting and behavioural symptoms. The first three domains each assess experiencing symptoms after intake of MTX, anticipatory (before intake) and/or associative (when thinking of MTX) complaints. The behavioural domain assesses crying, irritability, restlessness and refusal to take MTX. The items can be assigned 0 (no symptoms), 1 (mild), 2 (moderate) or 3 (severe) points. The MISS was calculated as the sum of the questionnaire, while blank questions were assigned 0 points. The score could range from 0 to 36. A patient was considered intolerant if she had a score above the validated cut point of 6 points in concert with at least one associative, anticipatory or behavioural symptom.¹⁵

Data collection

Patient characteristics were collected from the medical records at the time of completion of the questionnaire (Table 1). For five patients, the dose of MTX in mg/wk was converted to mg/m²/wk using the mean body surface area for their age and gender. The juvenile

Table 1. Baseline characteristics by MTX intolerance

Characteristics	Tolerant N=106 (59.2)	Intolerant N=73 (40.8)	P value*
Sex, female	69 (65.1)	49 (67.1)	0.78
Age (years), mean \pm SD	10.3 \pm 4.8	11.0 \pm 4.1	0.30 [†]
JIA subtype			
Oligoarticular, persistent	24 (22.6)	20 (27.4)	0.47
Oligoarticular, extended	14 (13.2)	18 (24.7)	0.049
Polyarticular	41 (38.7)	19 (26.0)	0.08
Systemic	3 (2.8)	2 (2.7)	0.99 [#]
Enthesitis-related	12 (11.3)	5 (6.8)	0.44 [#]
Psoriatic arthritis	8 (7.5)	5 (6.8)	0.99 [#]
Undifferentiated	4 (3.8)	4 (5.5)	0.72 [#]
Disease characteristics			
ANA positive	65 (61.9)	42 (61.8)	0.95
RF positive	7 (8.4)	2 (3.4)	0.31 [#]
HLA-B27 positive	17 (17.5)	9 (15.0)	0.83 [#]
Disease duration (years), mean \pm SD	3.2 \pm 3.3	5.2 \pm 4.1	<0.001 [†]
AJC, median (IQR)	0.0 (0.0-1.0)	0.0 (0.0-2.0)	0.078
JLM, median (IQR)	1.0 (0.0-2.0)	1.5 (0.0-3.0)	0.034
ESR (mm/h), median (IQR)	8.0 (4.0-14.0)	6.0 (4.0-12.0)	0.22
PGA, median (IQR)	1.0 (0.0-3.0)	2.0 (1.0-3.0)	0.32
Parent/patient global assessment, median (IQR)	2.0 (0.0-3.0)	2.0 (0.0-3.0)	0.79
JADAS-10, median (IQR)	4.0 (2.0-9.0)	4.0 (1.6-7.5)	0.91
MTX use			
Route of administration, exclusively PO	67 (63.2)	28 (38.4)	0.001
Route of administration, exclusively SC	26 (24.5)	20 (27.4)	0.67
Duration use (months), median (IQR)	14.5 (6.0-23.0)	21.0 (10.0-31.0)	0.006
Dose (mg/m ² /wk), median (IQR)	12.0 (10.0-14.0)	11.4 (9.8-13.2)	0.30
Additional medication			
Steroids	15 (21.7)	4 (10.3)	0.19 [#]
NSAIDs	33 (47.8)	18 (46.2)	0.87
Folic acid	39 (56.5)	21 (53.8)	0.79
Other DMARDs	14 (20.3)	12 (30.8)	0.22
MTX intolerance: MISS, median (IQR)	1.0 (0.0-3.0)	11.0 (8.0-16.0)	<0.001

Abbreviations: AJC, active joint count; ANA, antinuclear antibodies; DMARDs, disease-modifying anti-rheumatic drugs; ESR, erythrocyte sedimentation rate; HLA-B27, human leukocyte antigen type B27; IQR, interquartile range; JADAS, juvenile arthritis disease activity score; JIA, juvenile idiopathic arthritis; JLM, joints limited in movement; MISS, Methotrexate Intolerance Severity Score; MTX, methotrexate; NSAIDs, non-steroidal anti-inflammatory drugs; PGA, physician global assessment; PO, oral administration; RF, rheumatoid factor; SC, subcutaneous administration; SD, standard deviation.

Values are the number (%) of patients of non-missing data, except where indicated otherwise. ANA status was determined in 174 patients, RF in 141, HLA-B27 in 157, AJC in 174, JLM in 161, ESR in 147, PGA in 122, Parent/patient global assessment in 89, JADAS-10 in 87 and additional medication in 108 patients.

* Continuous variables were tested with the Mann-Whitney U test and categorical variables with Pearson's chi-square test, except where indicated otherwise.

[#] Fisher's exact test.

[†] Student's t test.

arthritis disease activity score (JADAS)-10 was calculated from the number of active joints, PGA, ESR and the parent/patient assessment of wellbeing.²⁴

Statistical analysis

To determine factors associated with intolerance, the patient characteristics were tested in a univariate analysis for differences between intolerant and tolerant patients. Route of administration was categorised as exclusively PO, exclusively SC, switch from PO to SC and switch from SC to PO, and differences of patient characteristics between the exclusively PO and exclusively SC groups were tested. Differences between the two groups who switched route of administration and the other two groups were not tested, since the reason and date of switch were not known. The crude odds ratio (OR) of the effect of route of administration on MTX intolerance was calculated using logistic regression. Other patient characteristics, such as age, gender, dosage and duration of MTX use and concomitant medication (especially NSAIDs), were subsequently added one by one and the change in OR was observed to detect potential confounding of the effect.

Finally, the prevalence of MTX adverse effects, according to the responses on the MISS questionnaire, was evaluated for differences between the tolerant and intolerant patients and between the exclusively PO and exclusively SC patients. We used the chi-square (χ^2), Mann-Whitney U (M-W U) or Kruskal-Wallis (KW) tests for non-normally distributed variables and the independent sample t-test for normally distributed variables. Statistical analyses were performed with SPSS, version 20 (SPSS Inc., Chicago, USA) and R statistics version 2.15.0 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

In all centres combined, 216 patients were eligible, of who 26 refused participation. Of 190 enrolled patients, four patients were excluded because they were older than 18 years, six because the route of administration was unknown and one because the MISS questionnaire was missing, leaving 179 patients for analysis. Table 1 shows the baseline characteristics, subdivided by MTX intolerance, at the time of completing the MISS questionnaire. MTX intolerance (score ≥ 6 and at least one anticipatory, associative or behavioural symptom) was present in 73 (40.8%) patients. The mean age was 10.6 years and the majority was female (65.9%). Of all patients, 26 used additional DMARDs, next to MTX (etanercept: $n=11$; sulfasalazine: $n=10$; adalimumab: $n=3$; hydroxychloroquine: $n=2$).

Differences between tolerant and intolerant patients in univariate analysis

There was strong evidence that intolerant patients had a longer disease duration, a longer history of MTX treatment and that they took MTX exclusively PO less frequently ($P \leq 0.001$, Table 1). There was some evidence the prevalence of the extended oligoarticular subtype was higher in the intolerant group, whereas the prevalence of the polyarticular subtype was lower (P value around 0.05, Table 1). Furthermore, there was some evidence that intolerant patients had a higher number of active and limited joints. For all other variables tested, there was no evidence of differences between the two groups (Table 1). Folic acid was administered in 56.5% of the tolerant patients and 53.8% of the intolerant patients ($P = 0.79$). As expected, the median MISS score was higher in the intolerant group (median of 11.0 versus 1.0 points, $P < 0.001$).

Differences between patients with different routes of administration in univariate analysis

The oral group was older than the subcutaneous group (mean age 11.6 vs. 9 years, $P = 0.003$, Table 2). Persistent oligoarticular JIA was more common in the subcutaneous group (39.1% vs. 17.9%, $P = 0.006$), whereas polyarticular arthritis was more common in the oral group (40.0% vs. 21.7%, $P = 0.03$). Psoriatic JIA had a prevalence of 10.5% in the oral group and was absent in the subcutaneous group. The median MISS scores per group were the following: PO: 2.0, SC: 4.5, PO > SC: 8.0 and SC > PO: 8.0. There was strong evidence these four groups differed (KW test: $P = 0.002$) and that the median MISS score of the exclusively SC group was higher than the median score of the exclusively PO group (M-W U test: $P = 0.003$). There was no evidence of other differences between the PO and SC groups (Table 2).

Effect of subcutaneous route of administration on MTX intolerance

The crude OR for MTX intolerance of taking MTX exclusively SC was 1.84 (95% CI 0.88-3.83) showing no evidence of increased odds. However, this effect was confounded by age, disease duration, PGA, NSAID use and folic acid use, as indicated by a substantial change in the OR after adjustment for these parameters. After adjustment for age, disease duration and PGA the OR for MTX intolerance of MTX SC changed to 3.37 (95% CI 1.19-10.0, $P = 0.02$, Table 3). Additionally, longer disease duration was associated with MTX intolerance (OR 1.2 [95% CI 1.1-1.4], $P < 0.001$). Inclusion of NSAID and folic acid use in this model yielded an unstable model due to paucity of data regarding additional medication, showing an OR of 10.1 (95% CI 1.79-77.6, $P = 0.01$). Interestingly, MTX dosage and duration of MTX use were no confounders, since inclusion of these variables in the model did not change the OR of route of administration.

Table 2. Baseline characteristics by route of administration

Characteristics	Exclusively PO N=95 (53.1)	Exclusively SC N=46 (25.7)	P value*
Sex, female	66 (69.5)	25 (54.3)	0.08
Age (years), mean \pm SD	11.6 \pm 4.4	9 \pm 4.7	0.003 [†]
JIA subtype			
Oligoarticular, persistent	17 (17.9)	18 (39.1)	0.006
Oligoarticular, extended	14 (14.7)	11 (23.9)	0.18
Polyarticular	38 (40.0)	10 (21.7)	0.03
Systemic	2 (2.1)	2 (4.3)	0.60 [#]
Enthesitis-related	12 (12.6)	3 (6.5)	0.39
Psoriatic arthritis	10 (10.5)	0 (0.0)	0.03
Undifferentiated	2 (2.1)	2 (4.3)	0.60
Disease characteristics			
ANA positive	52 (57.1)	30 (66.7)	0.29
RF positive	5 (7.2)	3 (7.5)	0.99 [#]
HLA-B27 positive	16 (19.5)	5 (12.8)	0.45 [#]
Disease duration (years), mean \pm SD	4 \pm 3.9	3.6 \pm 3.7	0.54 [†]
AJC, median (IQR)	0.0 (0.0-2.0)	0.0 (0.0-0.1)	0.17
JLM, median (IQR)	1.0 (0.0-3.0)	1.0 (0.0-3.0)	0.68
ESR (mm/h), median (IQR)	8.0 (5.0-16.0)	7.0 (5.0-12.5)	0.29
PGA, median (IQR)	2.0 (1.0-3.0)	1.0 (1.0-2.0)	0.11
Parent/patient global assessment, median (IQR)	2.0 (0.0-3.0)	1.0 (0.0-2.0)	0.19
JADAS-10, median (IQR)	5.0 (2.0-10.0)	3.0 (1.0-5.0)	0.12
MTX use			
Duration use (months), median (IQR)	13.0 (6.0-26.0)	17.0 (8.0-26.0)	0.30
Dose (mg/m ² /wk), median (IQR)	11.8 (10.0-13.6)	11.6 (9.7-13.2)	0.43
Additional medication			
Steroids	14 (19.7)	3 (11.5)	0.55 [#]
NSAIDs	39 (54.9)	9 (34.6)	0.08
Folic acid	41 (57.7)	13 (50.0)	0.50
Other DMARDs	19 (26.8)	5 (19.2)	0.45
MTX intolerance: MISS, median (IQR)	2.0 (0.0-7.0)	4.5 (3.0-11.0)	0.003

Abbreviations: see table 1.

Values are the number (%) of patients of non-missing data, except where indicated otherwise. For number of missing values, see Table 1.

* Continuous variables were tested with the Mann-Whitney U test and categorical variables with Pearson's chi-square test, except where indicated otherwise.

[#] Fisher's exact test.

[†] Student's t test.

Prevalence of MTX intolerance, gastrointestinal and behavioural symptoms

Table 4 contains the prevalence of gastrointestinal and behavioural symptoms, by MTX intolerance and route of administration. As expected, the intolerant patients showed a significantly higher prevalence for each item of the questionnaire.

Notably, there was strong evidence that behavioural symptoms (restlessness, crying, irritability and MTX refusal combined) were more frequent among children taking MTX SC ($P = 0.001$). Furthermore, there was weak evidence of a higher prevalence of abdominal pain after administration of MTX in the SC group (43.5% versus 27.4%, $P = 0.056$). The same held true for any symptom of the abdominal pain domain (after MTX, anticipatory and associatively), which was more common in the SC group (50.0% versus 34.7%, $P = 0.082$). There was no evidence of other differences between the PO and SC groups.

DISCUSSION

In this cross-sectional, multicentre study, 40.8% of 179 JIA patients receiving MTX were intolerant to the therapy. There was evidence that receiving MTX exclusively subcutaneously as compared to exclusively orally increased the odds of MTX intolerance (adjusted OR: 3.37 [95% CI 1.19-10.0]). There was strong evidence that the median MISS score was higher in the MTX SC group than in the MTX PO group and that patients taking subcutaneous MTX showed more often behavioural complaints, such as crying, restlessness, irritability and refusal to take MTX. Finally, there was some evidence that they experienced more abdominal pain after MTX administration. Taken together, the odds of MTX intolerance were higher in patients taking MTX SC and they experienced more gastrointestinal and behavioural adverse effects than patients taking MTX PO did.

Table 3. Multivariate analysis of MTX intolerance

Variable	OR (95% CI)	P-value
Route of administration		
PO	Reference	
Switch PO to SC	4.9 (1.4-18.1)	0.01
Switch SC to PO	4.0 (1.1-15.0)	0.03
SC	3.4 (1.2-10.0)	0.02
Age, years	1.0 (0.9-1.1)	0.94
Disease duration, years	1.2 (1.1-1.4)	<0.001
PGA	1.1 (0.9-1.4)	0.36

Abbreviations: CI, confidence interval; OR, odds ratio; PGA, physician's global assessment; PO, oral administration; SC, subcutaneous administration.

Table 4. Prevalence of items on the MISS per intolerance and route of administration

	Tolerant	Intolerant ^a	PO	SC	P-value*
Number of patients	106	73	95	46	
MTX intolerance ^a	0 (0)	73 (100)	28 (29)	20 (43)	0.10
Abdominal pain	28 (26)	50 (68)	33 (35)	23 (50)	0.08
After MTX	23 (22)	43 (59)	26 (27)	20 (43)	0.06
Anticipatory	4 (4)	15 (21)	10 (11)	4 (9)	0.99 [#]
Associative	3 (3)	26 (36)	9 (9)	6 (13)	0.57 [#]
Nausea	34 (32)	69 (95)	48 (51)	30 (65)	0.10
After MTX	29 (27)	61 (84)	43 (45)	24 (52)	0.44
Anticipatory	4 (4)	31 (42)	14 (15)	8 (17)	0.81 [#]
Associative	13 (12)	51 (70)	26 (27)	18 (39)	0.16
Vomiting	4 (4)	33 (45)	17 (18)	9 (20)	0.82 [#]
After MTX	3 (3)	31 (42)	15 (16)	8 (17)	0.81 [#]
Anticipatory	0 (0)	9 (12)	5 (5)	1 (2)	0.66 [#]
Behavioural	39 (37)	71 (97)	45 (47)	35 (76)	0.001
Restlessness	21 (20)	56 (77)	30 (32)	24 (52)	0.02
Crying	12 (11)	34 (47)	16 (17)	18 (39)	0.003
Irritability	14 (13)	57 (78)	29 (31)	25 (54)	0.006
MTX refusal	9 (8)	30 (41)	9 (9)	11 (24)	0.037 [#]

Abbreviations: MISS, methotrexate intolerance severity score; MTX, methotrexate; PO, exclusively orally administered MTX; SC, exclusively subcutaneous.

Values are the number (%) of patients within the groups of (in)tolerance and within the groups of different routes of administration. The groups who switched route of administration were left out of this table, because of potential biases.

^a Intolerance to methotrexate was defined as a score of ≥ 6 on the MISS and at least one associative, anticipatory or behavioural symptom.

* P-value of the null hypothesis of no differences between the PO and the SC group, using Pearson's chi square test, except where indicated otherwise.

[#] Fisher's exact test.

Some other factors were significantly associated with MTX intolerance (Table 1), most notably the duration of MTX use. When analysing these factors together with route of administration in multivariate analysis, these associations were not maintained (data not shown). Patients with persistent oligoarthritis used MTX SC more often (Table 2), probably because these patients tend to be younger and are therefore unable to swallow MTX PO. Another possible explanation is that MTX in these patients is frequently started because of uveitis. Treatment in these cases is performed in cooperation with the ophthalmologist, potentially leading to a different approach.

MTX intolerance is thought to arise as a classical conditioning response, as follows.^{11,15} In MTX therapy, many potential conditioned stimuli are present, such as the yellow colour of the drug. Unconditioned gastrointestinal adverse effects may occur after the intake of MTX. Over the course of weeks, the potential conditioned stimuli can elicit the unconditioned stimuli due to stimulation of the central nervous system, giving rise

to the so-called anticipatory and associative complaints.^{15;25} Why MTX intolerance would occur more often in the MTX SC group is unknown. It could be speculated that the higher prevalence of behavioural complaints in the MTX SC group may be explained by a fear of needles and the negative experience with injections, leading to stress and subsequently to crying, irritability and refusal to take the drug. In turn, these negative experiences could incline patients to value any gastrointestinal side effect more negatively, thus contributing to the high prevalence of the latter in the MTX SC group. These gastrointestinal side effects may also arise due to triggering of the chemoreceptor trigger zone in the medulla oblongata,^{13;14} folic acid depletion^{10;26} or as of yet unknown mechanisms. These mechanisms might be triggered more often in the case of MTX SC, because of its higher bioavailability.¹⁸⁻²¹ Finally, the subcutaneous route of administration was shown to be associated with a higher concentration of long chain MTX polyglutamates (MTX-PGs),²⁷ potentially contributing to the higher prevalence of adverse effects. However, an association between MTX-PGs and the frequency of adverse events could not be demonstrated in a longitudinal study in JIA²⁸ or RA.²⁹

Two accepted and widely used approaches to counter the side effects of MTX are the supplementation of folic or folinic acid 24 hours after MTX administration and the administration of anti-emetics.^{3;10;15;18;26} Some recent studies state there is lack of convincing evidence for these strategies.^{7;30} In our cohort, 55.6% of all patients used folic acid supplementation, which was comparable for the tolerant and intolerant group ($P = 0.79$) and for the SC and PO group ($P = 0.50$). We could not reliably assess the use of anti-emetics.

Previously, we studied a Dutch cohort, in which the prevalence of MTX intolerance was 50.5%. There was strong evidence of an increased frequency in SC patients (67.5% versus 44.5%, $P = 0.001$).¹⁵ Likewise, the occurrence of behavioural complaints was more frequent in the SC group. These results were partly confirmed by the current study. Of note, there were methodological differences between the two studies; the most important being that the Dutch paper assessed patients who switched route of administration together with patients who remained on the same route from the start of MTX. In the current study, patients who switched were excluded from the analyses, because it was suspected that many of these patients switched because of side effects, causing a high prevalence of intolerance in these groups (63.2% and 68.4%, respectively). Furthermore, they might have switched shortly before completing the questionnaire, making the results of the MISS more applicable to the route of administration before the switch.

In a recent British study, a slightly different questionnaire than ours was used to assess the frequency of MTX-induced nausea and vomiting in adolescents and adults.³¹ Nausea and vomiting were reported in 73% and 43% of adolescents respectively. In this study, too, SC administration of MTX was associated with nausea in a multivariate analysis. In addition, in another German study, MTX SC was discontinued more frequently than MTX PO, due to adverse effects. No differences were observed between MTX SC and MTX PO

with respect to efficacy or toxicity.²² Finally, in RA patients MTX SC administration was found to be associated with a higher prevalence of MTX intolerance as well.³²

In our study, we used a self-reporting method to assess the frequency of MTX intolerance. This could potentially lead to overestimation of the prevalence of MTX intolerance. In one centre (Sendenhorst), the questionnaire was sent home, potentially causing mainly MTX intolerant patients to complete and return it, thus overestimating the prevalence of MTX intolerance. However, the prevalence of MTX intolerance at Sendenhorst was in range with the other centres (41.5%). Next, due to missing data, we could not calculate the JADAS or determine the use of additional medication in a sufficient number of patients to include these as covariates in the multivariate analysis. Neither did we know the prevalence of uveitis in our sample. Patients who stopped MTX in the past because of intolerance were not included in the study and questions of the MISS who were left blank, were assigned 0 points, potentially causing the prevalence of MTX intolerance to have been underestimated. Finally, the aim of this study was to investigate the prevalence of MTX intolerance. Therefore, liver toxicity was not taken into account, since it is a well-known side effect of MTX, which is easily dealt with by suspending the drug.

In conclusion, in this cross-sectional study of a German sample of JIA patients, the prevalence of MTX intolerance was high. The odds of MTX intolerance were higher in patients taking MTX SC, they experienced more behavioural adverse effects than patients taking MTX PO did and at least as many gastrointestinal adverse effects as patients on MTX PO, thereby disconfirming the frequently held belief that SC causes fewer side effects,^{8,10,18} and confirming our previous findings.¹⁵ However, given the cross-sectional design of the study, results need to be interpreted with caution. Randomised controlled trials are necessary to obtain definite answers about the differences between SC and PO administration of MTX in terms of safety and efficacy.

ACKNOWLEDGEMENTS

We thank Michael Frosch, Tobias Schwarz and Christiane Zimmer for contributing to the preparation of this manuscript.

FUNDING

This work was supported by the 7th Framework programme of the European Union, SP3-People, support for training and career development for researchers (Marie Curie), Network for Initial Training (ITN), FP7-PEOPLE-2011-ITN, under the Marie Skłodowska-Curie [grant number 289903] to EHPvD; SHARE project, Executive Agency for Health

and Consumers [grant number 2011 1202] to NMW; and the Interdisciplinary Centre of Clinical Research at the University of Münster [grant number IZKF CRA04] to DH. GH receives funding from Abbvie, Pfizer and Roche/Chugai and speaker honoraria from Abbvie, Pfizer and Roche/Chugai, Novartis, UCB. The sponsors had no role in study planning, design, management or data analysis.

REFERENCE LIST

1. Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. *Lancet* 2011; 377(9783):2138-2149.
2. Ruperto N, Murray KJ, Gerloni V, Wulffraat N, de Oliveira SK, Falcini F et al. A randomized trial of parenteral methotrexate comparing an intermediate dose with a higher dose in children with juvenile idiopathic arthritis who failed to respond to standard doses of methotrexate. *Arthritis Rheum* 2004; 50(7):2191-2201.
3. Hashkes PJ, Laxer RM. Medical treatment of juvenile idiopathic arthritis. *JAMA* 2005; 294(13):1671-1684.
4. Foell D, Wulffraat N, Wedderburn LR, Wittkowski H, Frosch M, Gerss J et al. Methotrexate withdrawal at 6 vs 12 months in juvenile idiopathic arthritis in remission: a randomized clinical trial. *JAMA* 2010; 303(13):1266-1273.
5. Prince FH, van Suijlekom-Smit LW. Cost of biologics in the treatment of juvenile idiopathic arthritis: a factor not to be overlooked. *Paediatr Drugs* 2013; 15(4):271-280.
6. Chickermane PR, Khubchandani RP. Evaluation of the benefits of sequential addition of leflunomide in patients with polyarticular course juvenile idiopathic arthritis failing standard dose methotrexate. *Clin Exp Rheumatol* 2015; 33(2):287-292.
7. Nihues T, Lankisch P. Recommendations for the use of methotrexate in juvenile idiopathic arthritis. *Paediatr Drugs* 2006; 8(6):347-356.
8. Alsufyani K, Ortiz-Alvarez O, Cabral DA, Tucker LB, Petty RE, Malleson PN. The role of subcutaneous administration of methotrexate in children with juvenile idiopathic arthritis who have failed oral methotrexate. *J Rheumatol* 2004; 31(1):179-182.
9. Hashkes PJ, Becker ML, Cabral DA, Laxer RM, Paller AS, Rabinovich CE et al. Methotrexate: New Uses for an Old Drug. *J Pediatr* 2014; 164(2):231-236.
10. Murray KJ, Lovell DJ. Advanced therapy for juvenile arthritis. *Best Pract Res Clin Rheumatol* 2002; 16(3):361-378.
11. van der Meer A, Wulffraat NM, Prakken BJ, Gijsbers B, Rademaker CM, Sinnema G. Psychological side effects of MTX treatment in juvenile idiopathic arthritis: a pilot study. *Clin Exp Rheumatol* 2007; 25(3):480-485.
12. Ortiz-Alvarez O, Morishita K, Avery G, Green J, Petty RE, Tucker LB et al. Guidelines for blood test monitoring of methotrexate toxicity in juvenile idiopathic arthritis. *J Rheumatol* 2004; 31(12):2501-2506.
13. Dhir V, Aggarwal A. Methotrexate-related minor adverse effects in rheumatoid arthritis: more than a nuisance. *J Clin Rheumatol* 2012; 18(1):44-46.
14. Veyrat-Follet C, Farinotti R, Palmer JL. Physiology of chemotherapy-induced emesis and antiemetic therapy. Predictive models for evaluation of new compounds. *Drugs* 1997; 53(2):206-234.
15. Bulatovic M, Heijstek MW, Verkaaik M, Van Dijkhuizen EHP, Armbrust W, Hoppenreijns EP et al. High prevalence of methotrexate intolerance in juvenile idiopathic arthritis: development and validation of a methotrexate intolerance severity score. *Arthritis Rheum* 2011; 63(7):2007-2013.
16. Brunner HI, Johnson AL, Barron AC, Passo MH, Griffin TA, Graham TB et al. Gastrointestinal symptoms and their association with health-related quality of life of children with juvenile rheumatoid arthritis: validation of a gastrointestinal symptom questionnaire. *J Clin Rheumatol* 2005; 11(4):194-204.
17. Horneff G. Update on biologics for treatment of juvenile idiopathic arthritis. *Expert Opin Biol Ther* 2013; 13(3):361-376.
18. Ramanan AV, Whitworth P, Baildam EM. Use of methotrexate in juvenile idiopathic arthritis. *Arch Dis Child* 2003; 88(3):197-200.
19. Hamilton RA, Kremer JM. Why intramuscular methotrexate may be more efficacious than oral dosing in patients with rheumatoid arthritis. *Br J Rheumatol* 1997; 36(1):86-90.
20. Teresi ME, Crom WR, Choi KE, Mirro J, Evans WE. Methotrexate bioavailability

- after oral and intramuscular administration in children. *J Pediatr* 1987; 110(5):788-792.
21. Tukova J, Chladek J, Nemcova D, Chladkova J, Dolezalova P. Methotrexate bioavailability after oral and subcutaneous administration in children with juvenile idiopathic arthritis. *Clin Exp Rheumatol* 2009; 27(6):1047-1053.
 22. Klein A, Kaul I, Foeldvari I, Ganser G, Urban A, Horneff G. Efficacy and safety of oral and parenteral methotrexate therapy in children with juvenile idiopathic arthritis: an observational study with patients from the German Methotrexate Registry. *Arthritis Care Res (Hoboken)* 2012; 64(9):1349-1356.
 23. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004; 31(2):390-392.
 24. Consolaro A, Ruperto N, Bazso A, Pistorio A, Magni-Manzoni S, Filocamo G et al. Development and validation of a composite disease activity score for juvenile idiopathic arthritis. *Arthritis Rheum* 2009; 61(5):658-666.
 25. Hornby PJ. Central neurocircuitry associated with emesis. *Am J Med* 2001; 111 Suppl 8A:106S-112S.
 26. Ravelli A, Migliavacca D, Viola S, Ruperto N, Pistorio A, Martini A. Efficacy of folinic acid in reducing methotrexate toxicity in juvenile idiopathic arthritis. *Clin Exp Rheumatol* 1999; 17(5):625-627.
 27. Becker ML, Gaedigk R, Van Haandel L, Thomas B, Lasky A, Hoeltzel M et al. The effect of genotype on methotrexate polyglutamate variability in juvenile idiopathic arthritis and association with drug response. *Arthritis Rheum* 2011; 63(1):276-285.
 28. Bulatovic Calasan M, Den Boer E, de Rotte MC, Vastert SJ, Kamphuis S, De Jonge R et al. Methotrexate polyglutamates in erythrocytes are associated with lower disease activity in juvenile idiopathic arthritis patients. *Ann Rheum Dis* 2013.
 29. de Rotte MC, den BE, de Jong PH, Pluijm SM, Bulatovic CM, Weel AE et al. Methotrexate polyglutamates in erythrocytes are associated with lower disease activity in patients with rheumatoid arthritis. *Ann Rheum Dis* 2013.
 30. Amarilyo G, Rullo OJ, McCurdy DK, Woo JM, Furst DE. Folate usage in MTX-treated juvenile idiopathic arthritis (JIA) patients is inconsistent and highly variable. *Rheumatol Int* 2013; 33(9):2437-2440.
 31. Patil P, Parker RA, Rawcliffe C, Olaleye A, Moore S, Daly N et al. Methotrexate-induced nausea and vomiting in adolescent and young adult patients. *Clin Rheumatol* 2014; 33(3):403-407.
 32. Bulatovic Calasan M, van den Bosch OF, Creemers MC, Custers M, Heurkens AH, van Woerkom JM et al. Prevalence of methotrexate intolerance in rheumatoid arthritis and psoriatic arthritis. *Arthritis Res Ther* 2013; 15(6):R217.

Chapter

10

PREDICTION OF METHOTREXATE INTOLERANCE IN JUVENILE IDIOPATHIC ARTHRITIS: A PROSPECTIVE, OBSERVATIONAL COHORT STUDY

E.H. Pieter van Dijkhuizen,* Maja Bulatović Čalasan,*
Saskia M.F. Pluijm, Maurits C.F.J. de Rotte, Sebastiaan J. Vastert,
Sylvia Kamphuis, Robert de Jonge and Nico M. Wulffraat

* Contributed equally

ABSTRACT

Background

Methotrexate (MTX) is an effective and safe drug in the treatment of juvenile idiopathic arthritis (JIA). Despite its safety, MTX-related gastrointestinal adverse effects before and after MTX administration, termed MTX intolerance, occur frequently, leading to non-compliance and potentially premature MTX termination. The aim of this study was to construct a risk model to predict MTX intolerance.

Methods

In a prospective JIA cohort, clinical variables and single nucleotide polymorphisms were determined at MTX start. The Methotrexate Intolerance Severity Score was employed to measure MTX intolerance in the first year of treatment. MTX intolerance was most prevalent at 6 or 12 months after MTX start, which was defined as the outcome for the prediction model. The model was developed in 152 patients using multivariable logistic regression analysis and subsequently internally validated using bootstrapping.

Results

The prediction model included the following predictors: JIA category, antinuclear antibody, parent/patient assessment of pain, Juvenile Arthritis Disease Activity Score-27, thrombocytes, alanine aminotransferase and creatinine. The model classified 77.5% of patients correctly, and 66.7% of patients after internal validation by bootstrapping. The lowest predicted risk of MTX intolerance was 18.9% and the highest predicted risk was 85.9%. The prediction model was transformed into a risk score (range 0-17). At a cut-off of ≥ 6 , sensitivity was 82.0%, specificity 56.1%, positive predictive value was 58.7% and negative predictive value 80.4%.

Conclusions

This clinical prediction model showed moderate predictive power to detect MTX intolerance. To develop into a clinically usable tool, it should be validated in an independent cohort and updated with new predictors. Such an easy-to-use tool could then assist clinicians in identifying patients at risk to develop MTX intolerance, and in turn to monitor them closely and intervene timely in order to prevent the development of MTX intolerance.

BACKGROUND

Juvenile idiopathic arthritis (JIA) is the most common childhood rheumatic disease.^{1,2} In JIA, methotrexate (MTX) is the cornerstone treatment, due to its efficacy and safety. Serious adverse effects such as hepatotoxicity and bone marrow suppression occur rarely.³ In contrast, MTX-related gastrointestinal adverse effects, such as nausea, abdominal pain and vomiting, occur frequently.⁴⁻¹⁰ Folic acid supplementation is an accepted strategy to prevent and treat these adverse effects.¹¹⁻¹³ Despite folic acid use, many JIA patients experience gastrointestinal adverse effects after MTX intake.⁴⁻¹⁰ JIA patients also experience anticipatory adverse effects, occurring before MTX administration (at the sight of MTX), and associative adverse effects, occurring when thinking of MTX administration (its colour or smell).^{4,5,14} These adverse effects are thought to be a result of classical conditioning to the abovementioned physical symptoms experienced after MTX intake.¹⁴ Importantly, if physical symptoms are absent, conditioned responses cannot develop.¹⁵ Such a combination of symptoms, which we previously termed MTX intolerance,¹⁴ is a significant burden for JIA patients and their parents. Notably, MTX intolerance occurs in up to half of JIA patients on MTX,¹⁴ and can negatively affect their quality of life.⁶ Moreover, over three-quarters of intolerant patients reluctantly used or even refused MTX,¹⁴ which, besides leading to non-compliance, could lead to premature discontinuation of MTX, and even replacement by costly biologicals.^{5,16,17} Such consequences could be avoided, if the development of MTX intolerance is prevented.

To prevent MTX intolerance, it is crucial to predict which patients starting MTX will be at risk to develop it. Thus, clinicians could be able to prevent MTX intolerance in patients at risk by immediate treatment of emerging physical symptoms, which otherwise could give rise to conditioned responses. Treatment of physical symptoms could include lowering the MTX dose,⁴ or starting behavioural therapy⁵ or anti-emetics.¹⁸ Predicting MTX intolerance would enable clinicians to apply such treatment strategies only in those patients who are likely to develop MTX intolerance.

Single nucleotide polymorphisms (SNPs) involved in the MTX metabolic pathways, and clinical predictors have been associated with MTX-related gastrointestinal adverse effects in rheumatoid arthritis (RA)¹⁹⁻²⁸ and JIA, the latter of which were reviewed recently.²⁹ However, to date no model has been constructed to predict MTX intolerance in JIA. The aim of this cohort study was to develop and internally validate such a prediction model, using clinical and genetic predictors.

PATIENTS AND METHODS

Patients and study design

An investigator-initiated observational prospective study on efficacy and adverse effects of MTX in patients starting MTX (ISRCTN13524271) was performed at the University Medical

Centre Utrecht and Erasmus University Medical Centre Rotterdam, The Netherlands, between January 2008 and October 2012. It was approved by the Ethics Committees of the participating centres and the Central Committee on Research involving Human Subjects, and was conducted according to good clinical practice guidelines.

Patients aged 1-18 years, with a confirmed diagnosis of JIA according to International League of Associations for Rheumatology (ILAR) criteria,³⁰ who started MTX, were included. Those who had stopped MTX for at least three months, but re-started MTX due to a relapse, were also included. At the time of MTX start, their clinical data (Table 1) were documented in case report forms and blood for the analysis of SNPs was drawn.

All patients completed the previously developed and validated MTX Intolerance Severity Score (MISS) at 3, 6 and 12 months after MTX start.¹⁴ This questionnaire consists of 12 questions, assessing abdominal pain, nausea and vomiting after or before (anticipatory) MTX intake and when thinking of MTX (associative). Furthermore, it assesses behavioural complaints associated with MTX intake, such as crying, restlessness, irritability and refusal to take the drug. The score ranges from 0 to 36 and those with a score of ≥ 6 , including at least one anticipatory, associative or behavioural symptom, were defined as MTX intolerant.¹⁴

10

Development of MTX intolerance over time and patient selection

To define the outcome for the prediction model, the development of MTX intolerance at 3, 6 and 12 months after MTX start was assessed. For this analysis, of 175 patients starting MTX treatment, 8 patients were excluded due to a diagnosis other than JIA ($n=4$: Lyme disease, colitis, sarcoidosis, 22q11 deletion syndrome) and use of biologicals at MTX start ($n=3$: anakinra; $n=1$: etanercept), resulting in 167 eligible patients (Figure 1). Additionally, 25 patients who completed only one MISS during follow-up were excluded, as their development of MTX intolerance could not be determined. Therefore, the development of MTX intolerance was assessed in 142 patients (Figure 1). In the first year after MTX start, 59 (41.5%) patients were intolerant (score ≥ 6 with at least one anticipatory, associative or behavioural complaint) (Table 2). At 3 months, 22 (15.7%) patients were intolerant. However, intolerance resolved in the majority of these (13 [59.1%]) at 6 months. At 6 months, the number of intolerant patients increased to 33 (24.1%), of whom 24 (72.7%) were newly intolerant. At 12 months, 14 (42.4%) of those intolerant at 6 months stayed intolerant, whereas 8 had less than 6 points on the MISS and 11 did not complete it. The total number of intolerant patients at 12 months was 30 (23.3%), of whom 13 (43.3%) were newly intolerant (Table 2).

Taken together, the majority of patients developing MTX intolerance did so at 6 or 12 months after MTX start. Consequently, the outcome for the prediction model was defined as MTX intolerance at 6 or 12 months after MTX start.

Table 1. Prevalence, univariable ORs (95%-CI) and p-values for potential predictors of MTX intolerance at MTX start

Variables	Cohort, n=152			
	Frequency n (%) ^a	OR (95%-CI)	p-value	
Demographics				
Female	92 (60.5)	1.34 (0.64-2.82)	0.432	
Age at disease onset >8 years	80 (52.6)	0.68 (0.34-1.36)	0.271	
Age at MTX start* >12 years	72 (47.4)	0.54 (0.27-1.07)	0.073	
Disease duration at MTX start >0.5 years	103 (67.8)	0.79 (0.37-1.70)	0.535	
JIA category*^b				
Oligoarticular (persistent/extended)	62 (40.8)	Reference	0.094	
Polyarticular (RF negative/positive)	64 (42.1)	1.91 (0.86-4.24)		
Other (systemic/psoriatic/enthesitis)	26 (17.1)	0.78 (0.27-2.31)		
Disease characteristics				
ANA* ^{b,c} Positive	84 (55.3)	1.98 (0.97-4.07)	0.057	
RF ^c Positive	16 (10.5)	1.52 (0.62-3.72)	0.352	
HLA-B27 ^c Positive	11 (7.2)	0.78 (0.29-2.12)	0.510	
Uveitis Present	21 (13.8)	1.44 (0.55-3.78)	0.455	
Disease activity				
CHAQ disability score ^c	≤0.250	36 (23.7)	Reference	0.395
	0.250-1.875	88 (57.9)	0.61 (0.24-1.55)	
	>1.875	15 (9.9)	0.72 (0.18-2.80)	
Parent/patient assessment of pain* ^{b,c}	≤3 cm	58 (38.2)	Reference	0.086
	3-6 cm	36 (23.7)	2.19 (0.84-5.67)	
	>6 cm	42 (27.6)	0.78 (0.30-2.02)	
Parent/patient global assessment ^c	>2.5 cm	90 (59.2)	0.79 (0.36-1.72)	0.494
	>2	92 (60.5)	2.00 (0.91-4.41)	0.070
Limited joints*	>1	108 (71.1)	2.02 (0.92-4.46)	0.072
PGA ^d	≤2 cm	50 (32.9)	Reference	0.496
	2-5 cm	86 (56.6)	1.35 (0.53-3.47)	
	>5 cm	16 (10.5)	0.87 (0.21-3.60)	
ESR ^c >15 mm/hr	74 (48.7)	1.46 (0.66-3.25)	0.341	
CRP ^c >10 mg/L	49 (32.2)	0.83 (0.40-1.74)	0.544	
JADAS-27* ^{b,c}	≤5	16 (10.5)	Reference	0.048
	5-15	59 (38.8)	0.40 (0.11-1.40)	
	>15	52 (34.2)	0.93 (0.25-3.44)	
Biochemical variables^c				
Haemoglobin >7.5 mmol/L	78 (51.3)	1.18 (0.60-2.32)	0.620	
Leucocytes >7 ×10 ⁹ /L	96 (63.2)	1.21 (0.59-2.47)	0.606	
Thrombocytes* ^b > 350 ×10 ⁹ /L	74 (48.7)	1.61 (0.82-3.16)	0.161	
AST >17 IU/L	96 (63.2)	1.08 (0.50-2.36)	0.635	
ALT* ^b >12 IU/L	101 (66.4)	0.41 (0.19-0.88)	0.019	
Creatinine* ^b >50 µmol/L	56 (36.8)	0.51 (0.24-1.08)	0.069	

Table 1. (continued)

Variables	Cohort, n=152		
	Frequency n (%) ^a	OR (95%-CI)	p-value
Medication			
MTX dose, median (IQR)	mg/m ² /week	9.9 (9.0-11.2)	NA
MTX route	oral	148 (97.4)	NA
MTX restarted		31 (20.4)	1.22 (0.48-3.11)
Folic acid		150 (98.7)	NA
Anti-emetics		5 (3.3)	NA
NSAID		120 (78.9)	0.93 (0.38-2.28)
Single nucleotide polymorphisms^c			
<i>MTHFR</i> rs1801133 C>T	TT	15 (9.9)	0.60 (0.21-1.69)
<i>MTHFR</i> rs1801131 A>C	CC/AC	79 (52.0)	1.65 (0.76-3.62)
<i>MTRR</i> rs1801394 A>G*	GG/AG	117 (77.0)	0.53 (0.24-1.20)
<i>RFC/SLC19A1</i> rs1051266 C>T*	TT	17 (11.2)	1.77 (0.74-4.25)
<i>ITPA</i> rs1127354 C>A	AA/CA	15 (9.9)	0.62 (0.22-1.74)
<i>AMPD1</i> rs17602729 G>A	AA/GA	41 (27.0)	1.46 (0.70-3.05)
<i>ATIC</i> rs2372536 C>G	GG/CG	93 (61.2)	0.84 (0.39-1.83)
<i>ADA22</i> rs73598374 C>T	TT/CT	13 (8.6)	NA
<i>ADORA2A</i> rs5751876 C>T	TT	28 (18.4)	1.54 (0.65-3.64)
<i>MDR-1/ABCB1</i> rs1128503 G>A*	AA	32 (21.1)	1.73 (0.75-3.98)
<i>MDR-1/ABCB1</i> rs1045642 G>A	AA	44 (28.9)	1.40 (0.65-3.01)
<i>MDR-1/ABCB1</i> rs2032582 C>A/T	AA/TT	24 (15.8)	1.51 (0.63-3.64)
<i>MRP-1/ABCC1</i> rs35592 T>C	CC/TC	52 (34.2)	0.79 (0.39-1.57)
<i>MRP-1/ABCC1</i> rs3784862 A>G	GG/AG	73 (48.0)	0.97 (0.50-1.91)
<i>MRP-2/ABCC2</i> rs4148396 C>T	TT	18 (11.8)	1.57 (0.60-4.08)
<i>MRP-2/ABCC2</i> rs717620 C>T	TT/CT	44 (28.9)	0.82 (0.37-1.82)
<i>MRP-3/ABCC3</i> rs4793665 T>C	CC/TC	92 (60.5)	0.73 (0.36-1.49)
<i>MRP-3/ABCC3</i> rs3785911 A>C*	CC/AC	78 (51.3)	1.67 (0.84-3.32)
<i>MRP-4/ABCC4</i> rs868853 T>C	CC/TC	22 (14.5)	0.88 (0.35-2.18)
<i>MRP-4/ABCC4</i> rs2274407 C>A	AA/CA	20 (13.2)	1.33 (0.48-3.73)
<i>MRP-5/ABCC5</i> rs2139560 G>A	AA/GA	92 (60.5)	1.31 (0.64-2.68)
<i>BCRP/ABCG2</i> rs13120400 T>C	CC/TC	63 (41.4)	0.77 (0.38-1.59)
<i>BCRP/ABCG2</i> rs2231142 G>T	TT/GT	30 (19.7)	0.96 (0.42-2.20)
<i>FPGS</i> rs4451422 A>C	CC/AC	102 (67.1)	1.37 (0.63-2.94)
<i>GGH</i> rs10106587 A>C	CC/AC	73 (48.0)	1.20 (0.59-2.46)
<i>GGH</i> rs3758149 G>A	AA/GA	77 (50.7)	1.20 (0.57-2.55)
<i>PCFT/SLC46A1</i> rs2239907 C>T	TT/CT	104 (68.4)	1.49 (0.69-3.23)

Abbreviations: ALT, alanine aminotransferase; ANA, antinuclear antibody; AST, asparagine aminotransferase; CHAQ, childhood health assessment questionnaire; CI, confidence interval; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; HLA, human leucocyte antigen; IQR, interquartile range; IU, international units; JADAS, juvenile arthritis disease activity score; JIA, juvenile idiopathic arthritis; MICE, multivariate imputation by chained equations; MTX, methotrexate; NSAID, non-steroidal anti-inflammatory drug; OR, odds ratio; PGA, physician global assessment; RF, rheumatoid factor.

^aVariables associated with the outcome at p<0.20 in the univariable logistic regression analysis. Variables with

Table 1. (continued)

observed frequencies of <5 in the cross-tabulation with the outcome were excluded from the univariable logistic analysis: MTX route, use of folic acid, use of anti-emetics and ADA22 rs73598374.

^a Frequencies are based on observed data, not imputed data.

^b JIA category, ANA, parent/patient assessment of pain, JADAS-27, thrombocytes, ALT and creatinine were included in the multivariable logistic regression analysis.

^c MICE was used to impute missing values in the following variables (percentage of missing values): HLA-B27 (60.5), RF (19.1), JADAS-27 (16.4), CRP (15.8), parent/patient global assessment (11.8), *RFC/SLC19A1* rs1051266 (11.8), creatinine (11.2), parent/patient assessment of pain (10.5), CHAQ disability score (8.6), *MDR-1/ABCB1* rs2032582 (8.6), ALT (7.9), AST (7.2), ESR (5.3), *GGH* rs3758149 (4.6), *MRP-2/ABCC2* rs717620 (3.9), *MRP-4/ABCC4* rs868853 (3.9), *MRP-5/ABCC5* rs2139560 (3.9), *GGH* rs10106587 (3.9), *MTHFR* rs1801131 (3.3), *ATIC* rs2372536 (3.3), *ADORA2A* rs5751876 (3.3), *MRP-1/ABCC1* rs3784862 (3.3), *MRP-2/ABCC2* rs4148396 (3.3), *MRP-3/ABCC3* rs4793665 (3.3), *BCRP/ABCG2* rs13120400 (3.3), *PCFT/SLC46A1* rs2239907 (3.3), *MTHFR* rs1801133 (2.6), *MTRR* rs1801394 (2.6), *ITPA* rs1127354 (2.6), *AMPD1* rs17602729 (2.6), ADA22 rs73598374 (2.6), *MDR-1/ABCB1* rs1128503 and rs1045642 (2.6), *MRP-1/ABCC1* rs35592 (2.6), *MRP-3/ABCC3* rs3785911 (2.6), *MRP-4/ABCC4* rs2274407 (2.6), *BCRP/ABCG2* rs2231142 (2.6), *FPGS* rs4451422 (2.6), thrombocytes (2.0), ANA (2.0), hemoglobin (1.3), leucocytes (1.3).

^d PGA was determined retrospectively by an experienced physician (SJV) in 20 visits (13.2%).

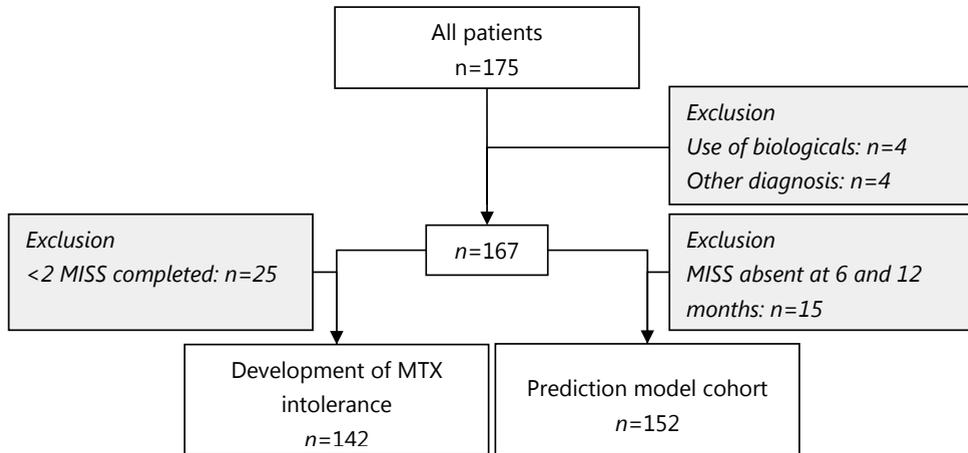


Figure 1. Flowchart. **Abbreviations:** MISS, Methotrexate Intolerance Severity Score; MTX, methotrexate.

For the construction of the prediction model, patients with a completed MISS at 6 or 12 months were re-selected from the eligible cohort of 167 patients, resulting in 152 included patients (Figure 1).

Potential clinical and genetic predictors

Potential clinical predictors (demographics, JIA category, disease characteristics, disease activity and biochemical measurements) were identified at baseline (Table 1). Potential genetic predictors were SNPs involved in the MTX metabolic pathways, with a high polymorphic allele frequency and documented functional effects.³¹ SNPs were determined

Table 2. MTX intolerance development

Time point	N	Intolerance, n(%) ^a
3 months	140 ^b	22 (15.7)
6 months	137 ^b	33 (24.1)
12 months	129 ^b	30 (23.3)
First treatment year	142	59 (41.5)
6 or 12 months ^d	152 ^c	51 (33.6)

Abbreviations: MTX, methotrexate; n, number of patients

^a Frequencies are based on observed data; ^b Patients still on MTX; ^c Cohort for prediction model construction; ^d Outcome was imputed in 21.7% of cases

in the following genes: methylenetetrahydrofolate reductase (*MTHFR*), reduced folate carrier (*RFC*), methionine synthase reductase (*MTRR*), inosine triphosphatase (*ITPA*), adenosine monophosphate deaminase (*AMPD*), aminoimidazole-4-carboxamide ribonucleotide transformylase (*ATIC*), adenosine-deaminase (*ADA*), adenosine A2A receptor (*ADORA2A*), multidrug resistance (*MDR*) 1, multidrug resistance protein (*MRP*) 1-5, breast cancer resistance protein (*BCRP*), foylpolylglutamate synthase (*FPGS*), gamma glutamyl hydrolase (*GGH*) and proton-coupled folate transporter (*PCFT*) (Table 1).

10

Statistical analysis

Prediction model construction

The prediction model was constructed in several steps. First, missing values were imputed using multivariate imputation by chained equations (MICE).³² This was done to ensure that all collected data could be used for the development of the model. Second, to facilitate implementation of the model in daily clinical practice, continuous variables were dichotomised or categorised, according to patterns in the data or the risk gradients across percentiles, and the cut-off points with the lowest p-value on the log-likelihood ratio test (i.e. those yielding the optimal association) were chosen.³³ Third, all variables were entered in a univariable logistic regression analysis. The results are presented as regression coefficients (β) and odds ratios (OR) with 95% confidence intervals (95% CI). The regression coefficients are an indication of the direction and the magnitude of the effect of the individual predictors, whereas the ORs with 95% CI indicate the significance of the association.

Variables with a p-value <0.20 on the log-likelihood ratio test in the univariable analysis were eligible for inclusion in the multivariable logistic regression analysis. The maximum number of included variables equalled the square root of the number of cases (MTX intolerant patients) in the cohort. If more variables were eligible than the allowed maximum, or if variables correlated (Spearman's $|\rho| > 0.40$), those with the lowest p-value on the log-likelihood ratio test were included in the multivariable analysis. In

addition, presence of effect modification by the predictors in the model was assessed. Effect modification is the situation in which the effect of one predictor on the outcome is modified by the value of another factor. For example, the effect of a predictor may differ between boys and girls. Statistically, this is tested by adding interaction terms to the model, allowing the regression coefficients to take different values for different categories of patients.

Predictive power of the model was assessed with the C-statistic, which reflects the percentage of patients classified correctly. To determine whether the model fit the data well, the Hosmer-Lemeshow test was employed. Multicollinearity was tested with variance inflation factors (VIF).

Prediction model validation and risk score computation

All prediction models need to be validated. Since no independent cohort was available, the model was internally validated using an established statistical technique, called bootstrap³⁴⁻³⁶. In short, 200 bootstrap cohorts (of equal size as the original dataset, $n=152$) were randomly drawn, with replacement, from the cases in the original dataset. Next, to each bootstrap cohort, bootstrap multivariable models were fitted (200 in total) using exactly the same methods as described above for the original model, and the corresponding C-statistics (C_{boot}) were determined. Then, the probability of MTX intolerance of the patients in the original dataset was calculated using each of these multivariable models, resulting in another set of C-statistics ($C_{boot-original}$) reflecting the percentage of patients predicted correctly according to each of these models. The difference between $C_{boot-original}$ values and C_{boot} values is an estimate of the so-called optimism value (i.e. how much the original model fitted to the original dataset was optimistic compared to the “real” performance of the model in the population). Therefore, in order to obtain the final adjusted C-statistic, indicating the “real” performance of the model in the population,³⁶ two additional steps need to be performed: a) Subtraction of $C_{boot-original}$ values from C_{boot} values and averaging them in order to obtain the optimism value; b) Subtraction of this optimism value from $C_{original}$ (the C-statistic of the original model, developed in the original dataset), thus obtaining the final adjusted C-statistic. Furthermore, to correct for overfitting, the regression coefficients were reduced with a shrinkage factor, calculated from the bootstrap re-sampling.

All the above-mentioned procedures were performed twice. Firstly, only the routinely available clinical variables were considered as potential predictors. Secondly, SNPs were also considered as potential predictors in order to determine whether they contributed to the prediction of MTX intolerance.

To compute a risk score of becoming MTX intolerant, the shrunken regression coefficients were multiplied and rounded off to obtain simple scores that sum up to a total risk score. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and accuracy of various cut-off points were calculated.

Statistical analyses were carried out with R statistics version 2.15.0 (R Foundation for Statistical Computing, Vienna, Austria), using the packages Hmisc (by Frank E Harrell Jr with contributions from many other users, version 3.9-3, 2012) and mice.³²

RESULTS

Baseline characteristics of the prediction model cohort

The prediction model was constructed in 152 patients. According to the outcome as defined above, 51 (33.6%) patients were MTX intolerant (Table 2). Intolerant and tolerant patients did not differ regarding the proportion of MTX re-starters, MTX dose, route of administration, concomitant medication use or disease activity (Juvenile Arthritis Disease Activity score [JADAS-27]) at 6 and 12 months after MTX start (data not shown).

Nineteen (12.5%) patients discontinued MTX treatment during the follow-up, because of MTX intolerance (n=8), disease remission (n=3), insufficient effect (n=2), MTX toxicity (increased liver enzymes: n=1) or other reasons (n=5). Patients also switched the route of administration due to gastrointestinal complaints (either from oral to subcutaneous or vice versa): 8 patients after 3 months, 6 patients after 6 months and 1 patient after 12 months.

Baseline characteristics are depicted in Table 1. Thirty-one patients (20.4%) had re-started MTX treatment due to a relapse after at least three months discontinuation. The majority of patients had either oligoarticular or polyarticular JIA (82.9%), with high disease activity (median JADAS-27 of 12.7 [interquartile range 7.6-18.2]). Median MTX dose was 9.9 mg/m²/week, administered mostly as oral MTX (97.4%) with concomitant use of folic acid (98.7%).

Clinical prediction model

First, a model was constructed, containing clinical variables only, excluding the SNPs. Ten clinical variables were associated in the univariable analysis with MTX intolerance (p<0.20; Table 1). The maximum number of variables allowed in the multivariable analysis was seven. Those with the lowest p-value were selected for the clinical prediction model, namely JIA category, JADAS-27, parent/patient assessment of pain, antinuclear antibody (ANA), alanine aminotransferase (ALT), thrombocytes and creatinine, and an interaction term between creatinine and JIA category was added. The C-statistic of the clinical prediction model was 77.5% (Table 3). The model fit the data well, as shown by a non-significant Hosmer-Lemeshow test (p=0.705). There was no multicollinearity (data not shown).

Table 3. Prediction model and scores for MTX intolerance

Predictors	OR (95%-CI)	p-value	β^a	Score ^b	
JIA category					
Oligoarticular (persistent/extended)	Reference			0	
Polyarticular (RF negative/positive)	4.99 (1.36-18.34)	0.016	0.914	5	
Other (systemic/psoriatic/enthesitis)	0.93 (0.16-5.49)	0.935	-0.042	0	
ANA	Positive	1.98 (0.83-4.68)	0.122	0.387	2
Parent/patient assessment of pain	≤3 cm	Reference		0	
	3-6 cm	2.06 (0.72-5.89)	0.175	0.412	2
	>6 cm	0.60 (0.17-2.07)	0.421	-0.288	-1
JADAS-27	≤5	Reference		0	
	5-15	0.35 (0.08-1.56)	0.168	-0.599	-3
	>15	0.77 (0.14-4.32)	0.766	-0.150	-1
Thrombocytes	>350 × 10 ⁹ /L	1.27 (0.49-3.27)	0.621	0.136	1
ALT	>12 IU/L	0.39 (0.16-0.96)	0.040	-0.534	-3
Creatinine	>50 μmol/L	1.37 (0.33-5.67)	0.665	0.179	1
Interaction term creatinine*JIA category					
	>50 μmol/L & polyarticular arthritis	0.17 (0.02-1.35)	0.093	-1.022	-5
	>50 μmol/L & other JIA category	0.82 (0.07-9.74)	0.878	-0.110	-1
Constant			-0.039	7	
C-statistic	77.5%				
C-statistic (optimism-corrected by bootstrap)	66.7%				
Hosmer-Lemeshow test (p-value)	0.705				

Abbreviations: ALT, alanine aminotransferase; ANA, anti-nuclear antibody; CI, confidence interval; JADAS, juvenile arthritis disease activity score; JIA, juvenile idiopathic arthritis; MTX, methotrexate; OR, odds ratio; RF, rheumatoid factor.

^a These are shrunk coefficients (by factor 0.5688) to correct for overfitting.

^b Shrunk coefficients were multiplied by 5 and rounded off to the nearest integer. The constant was adjusted to obtain the minimum score of 0.

Clinical-genetic prediction model

Next, SNPs were considered as potential predictors in order to determine their contribution to MTX intolerance prediction. Four SNPs in the *MTRR*, *RFC*, *MDR-1* and *MRP-3* genes had univariable p-values of <0.20, however these p-values (range: 0.123-0.194) were generally higher than those of the clinical model variables (range: 0.048-0.161) (Table 1). Hence, since seven variables with the smallest p-values were selected for multivariable analysis, only the *MTRR* rs1801394 SNP, next to six clinical variables (those from the abovementioned clinical model, excluding thrombocytes), were included in the model. The model's C-statistic was 77.7%.

Prediction model validation

Both the clinical and the clinical-genetic prediction model were internally validated using bootstrapping. Upon internal validation, the corrected C-statistic of the clinical model was 66.7%, whereas the corrected C-statistic of the clinical-genetic model was 64.6%.

Since the clinical-genetic model did not perform better than the model with clinical variables, the latter was given preference as clinical variables are readily available at MTX start, making it easier to apply the model in clinical practice.

Risk score

To enable health care professionals to use the model easily, the shrunken regression coefficients of the clinical model's predictors, transformed into simple scores, were used to compute an individual risk score for being MTX intolerant. This score ranged from 0 to 17 points, with a higher score reflecting a higher probability of MTX intolerance (Table 3). The lowest predicted risk of being MTX intolerant was 18.9%, if the following predictors were present: oligoarticular JIA, negative ANA, parent/patient assessment of pain >6 cm, JADAS-27 of 5-15 points, thrombocytes $\leq 350 \times 10^9/L$, ALT >12 IU/L and creatinine $\leq 50 \mu\text{mol/L}$. The combination of these predictors resulted in a score of 0 [7 (the constant) + 0 + 0 + (-1) + (-3) + 0 + (-3) + 0] (Table 3). On the other hand, the highest predicted risk of being MTX intolerant was 85.9%, if the following predictors were present: polyarticular JIA, positive ANA, parent/patient assessment of pain of 3-6 cm, JADAS-27 ≤ 5 points, thrombocytes $> 350 \times 10^9/L$, ALT <12 IU/L and creatinine $\leq 50 \mu\text{mol/L}$. The combination of these predictors resulted in a score of 17 [7 + 5 + 2 + 2 + 0 + 1 + 0 + 0].

Within the 0-17 range, the diagnostic accuracy of different cut-off scores for predicting the risk of being MTX intolerant was evaluated by computing the corresponding sensitivity, specificity, PPV, NPV, and accuracy (Table 4). Our goal was to correctly identify as many future MTX intolerant patients as possible (high sensitivity), while attempting to avoid misidentification of tolerant patients as intolerant patients (moderate specificity). This was reached at the cut-off score ≥ 6 , where 82% of intolerant patients and 56.1% of tolerant patients were identified correctly.

Table 4. Diagnostic parameters of the risk score for various cut-off scores

Cut-off	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
≥ 4	93.4	29.9	50.3	85.7	57.3
≥ 5	87.8	46.1	55.3	83.3	64.1
≥ 6	82.0	56.1	58.7	80.4	67.3
≥ 7	69.2	69.9	63.6	74.9	69.6
≥ 8	58.7	80.3	69.4	71.9	71.0
≥ 9	46.0	86.8	72.6	67.9	69.2

Abbreviations: NPV, negative predictive value; PPV, positive predictive value.

DISCUSSION

We developed and internally validated a prediction model for MTX intolerance at 6 or 12 months after MTX start in a large JIA cohort, consisting of routine clinical variables: JIA category, JADAS-27, parent/patient assessment of pain, ANA, ALT, thrombocytes, creatinine and an interaction term between creatinine and JIA category. The model classified 77.5% of patients correctly, and 66.7% after internal validation. It should be validated in an independent cohort and updated with other predictors.

In our model, patients who had more pain (>6 cm), higher baseline disease activity assessed with JADAS-27 and higher ALT, had a lower risk to become MTX intolerant. On the other hand, patients with positive ANA, who had less pain (3-6 cm), higher thrombocyte levels and higher creatinine, had an increased risk of MTX intolerance. Creatinine level and age were correlated, so creatinine can be regarded as a surrogate marker for age (median age was 7.5 years [patients with creatinine ≤ 50 $\mu\text{mol/L}$] versus 13.7 years [creatinine >50 $\mu\text{mol/L}$]). The relationship between JIA category, creatinine (age) and MTX intolerance was complex: In younger patients, polyarticular JIA was a strong predictor for intolerance (score 5, Table 3), whereas in older patients this effect disappeared (score 5 for polyarticular JIA and -5 for the interaction term between older patients (higher creatinine) and polyarticular JIA).

To predict which patients are prone to develop MTX intolerance, our risk score could be readily used by clinicians, since it is based on clinical variables, which are routinely determined and available for all JIA patients before MTX start. At the cut-off score of ≥ 6 , as many as 82% of intolerant patients were classified correctly (high sensitivity), while maintaining correct classification of 56.1% of tolerant patients (modest specificity). Table 4 provides the sensitivity and specificity of other potential cut off points.

Identification of patients at risk increases patients' and clinicians' awareness of MTX intolerance. In patients at risk, clinicians should frequently (i.e. every 4 weeks) monitor MTX-related gastrointestinal adverse effects, using the MISS, from the very start of MTX treatment. This would enable clinicians to treat the emerging physical symptoms early, for example by lowering MTX dose,⁴ adding anti-emetics¹⁸ or applying behavioural therapy,⁵ thus preventing the development of a classical conditioning response¹⁵ and hence MTX intolerance. The effect of these timely interventions on the development of MTX intolerance should be determined in a clinical trial.

The outcome of our prediction model was defined as MTX intolerance at 6 or 12 months after MTX start, since the majority of patients developing MTX intolerance did so at these time-points. The later onset of MTX intolerance is consistent with the notion that the development of MTX intolerance is governed by a classical conditioning response, which worsens over time.^{5,14} Moreover, in our previous cross-sectional study in patients with longer MTX use (interquartile range: 0.6-3.6 years), we demonstrated higher prevalence of MTX intolerance (50.5-67.5%) compared to the prevalence of 34.1% in the present longitudinal study during the first year of MTX treatment.¹⁴ This also supports

the notion that MTX intolerance takes time to develop and that longer MTX use may increase the risk of MTX intolerance. To determine whether the risk of MTX intolerance indeed increases with longer MTX use, development of MTX intolerance should be monitored beyond one year of MTX use. Nevertheless, MTX intolerance ensued in 15.8% of patients already after 3 months of MTX use. Interestingly, patients who had restarted MTX had a higher risk of becoming intolerant after 3 months than those newly starting MTX (36% versus 12.7%, $p=0.015$).

To our knowledge, no previous studies have developed a similar model and a corresponding risk score to predict the occurrence of MTX-induced gastrointestinal adverse effects in JIA. In a recently published paper, predictors for MTX adverse events in JIA patients, including the predictors in the current model, were reviewed. Only a few candidate predictors were elucidated, and validation of these lacked.²⁹

Our study did not identify genotypes as predictors for intolerance. In contrast, in RA, two studies identified combinations of risk genotypes to predict adverse effects in general and gastrointestinal adverse effects in particular.^{20,26} In our study, only 4 of 27 SNPs were moderately associated with MTX intolerance and only one SNP could be included in the clinical-genetic model, which had comparable predictive power as the clinical model. Previously, in RA and JIA, significant associations ($p<0.05$) were reported between SNPs in the *MTHFR*, *ATIC*, *ADORA*, *MRP2/ABCC2* and *GGH* genes and gastrointestinal adverse effects.^{19-22,24-28;37,38} SNPs in these genes were not associated with MTX intolerance in our study, which could be due to disparities in patient groups (RA versus JIA), cohorts (cross-sectional versus longitudinal), and the definition of MTX-induced gastrointestinal complaints (after MTX versus before *and* after MTX use). These results taken together with our current study show that it is still difficult to predict reliably the risk of developing MTX adverse events in general and MTX intolerance in particular.

The strengths of our study were that MTX intolerance was assessed using a validated questionnaire. In addition, the model was constructed and internally validated in a large prospective JIA cohort. Internal validation using bootstrapping is an established method to estimate the performance of a prediction model in the population, comparable to external validation in an independent cohort.³⁴⁻³⁶

In conclusion, we developed and internally validated a clinical prediction model for MTX intolerance in a large JIA cohort. It is an easy-to-use tool to identify patients at risk of developing MTX intolerance, and in turn to monitor them closely and intervene timely, in order to prevent MTX intolerance and its negative impact on patients' daily lives, compliance and continuation of an effective treatment. In its current composition, the model performs moderately well and should be validated in an independent cohort and updated with new predictor variables before it can be broadly used in clinical practice.

ACKNOWLEDGEMENTS

We wish to acknowledge: M.J.C. Eijkemans for statistical support; A. Blaauw, M.J.W. van Opdorp and A. van Dijk for valuable assistance during patient inclusion, case report form completion and investigator site file maintenance; and B. van Zelst and P. Griffioen for SNP analysis.

FUNDING

EHPvD was supported by funding from the 7th Framework programme of the EU, SP3-People, support for training and career development for researchers (Marie Curie), Network for Initial Training (ITN), FP7-PEOPLE-2011-ITN, under the Marie Skłodowska-Curie grant agreement No 289903. The project was furthermore supported by the SHARE project, EAHC grant number 2011 1202 [grant to NMW] and the Dutch Arthritis Association [NR 07-01-402 to NMW, NR 06-2-402 to RdJ]. The sponsors had no role in the study planning, design, management or data analysis.

REFERENCE LIST

1. Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. *Lancet* 2011; 377(9783):2138-2149.
2. Ravelli A, Martini A. Juvenile idiopathic arthritis. *Lancet* 2007; 369(9563):767-778.
3. Ortiz-Alvarez O, Morishita K, Avery G, Green J, Petty RE, Tucker LB et al. Guidelines for blood test monitoring of methotrexate toxicity in juvenile idiopathic arthritis. *J Rheumatol* 2004; 31(12):2501-2506.
4. Murray KJ, Lovell DJ. Advanced therapy for juvenile arthritis. *Best Pract Res Clin Rheumatol* 2002; 16(3):361-378.
5. van der Meer A, Wulffraat NM, Prakken BJ, Gijsbers B, Rademaker CM, Sinnema G. Psychological side effects of MTX treatment in juvenile idiopathic arthritis: a pilot study. *Clin Exp Rheumatol* 2007; 25(3):480-485.
6. Brunner HI, Johnson AL, Barron AC, Passo MH, Griffin TA, Graham TB et al. Gastrointestinal symptoms and their association with health-related quality of life of children with juvenile rheumatoid arthritis: validation of a gastrointestinal symptom questionnaire. *J Clin Rheumatol* 2005; 11(4):194-204.
7. Hashkes PJ, Laxer RM. Medical treatment of juvenile idiopathic arthritis. *JAMA* 2005; 294(13):1671-1684.
8. Ruperto N, Murray KJ, Gerloni V, Wulffraat N, deOliveiraSK, FalciniFetal. A randomized trial of parenteral methotrexate comparing an intermediate dose with a higher dose in children with juvenile idiopathic arthritis who failed to respond to standard doses of methotrexate. *Arthritis Rheum* 2004; 50(7):2191-2201.
9. Ramanan AV, Whitworth P, Baildam EM. Use of methotrexate in juvenile idiopathic arthritis. *Arch Dis Child* 2003; 88(3):197-200.
10. Alsufyani K, Ortiz-Alvarez O, Cabral DA, Tucker LB, Petty RE, Malleson PN. The role of subcutaneous administration of methotrexate in children with juvenile idiopathic arthritis who have failed oral methotrexate. *J Rheumatol* 2004; 31(1):179-182.
11. Ravelli A, Migliavacca D, Viola S, Ruperto N, Pistorio A, Martini A. Efficacy of folinic acid in reducing methotrexate toxicity in juvenile idiopathic arthritis. *Clin Exp Rheumatol* 1999; 17(5):625-627.
12. Alarcon GS, Morgan SL. Folinic acid to prevent side effects of methotrexate in juvenile rheumatoid arthritis. *J Rheumatol* 1996; 23(12):2184-2185.
13. Hunt PG, Rose CD, Ilvain-Simpson G, Tejani S. The effects of daily intake of folic acid on the efficacy of methotrexate therapy in children with juvenile rheumatoid arthritis. A controlled study. *J Rheumatol* 1997; 24(11):2230-2232.
14. Bulatovic M, Heijstek MW, Verkaaik M, van Dijkhuizen EHP, Armbrust W, Hoppenreijns EP et al. High prevalence of methotrexate intolerance in juvenile idiopathic arthritis: development and validation of a methotrexate intolerance severity score. *Arthritis Rheum* 2011; 63(7):2007-2013.
15. Figueroa-Moseley C, Jean-Pierre P, Roscoe JA, Ryan JL, Kohli S, Palesh OG et al. Behavioral interventions in treating anticipatory nausea and vomiting. *J Natl Compr Canc Netw* 2007; 5(1):44-50.
16. Beukelman T, Patkar NM, Saag KG, Tolleson-Rinehart S, Cron RQ, DeWitt EM et al. 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: initiation and safety monitoring of therapeutic agents for the treatment of arthritis and systemic features. *Arthritis Care Res (Hoboken)* 2011; 63(4):465-482.
17. Horneff G. Update on biologicals for treatment of juvenile idiopathic arthritis. *Expert Opin Biol Ther* 2013.
18. Ravelli A, Martini A. Methotrexate in juvenile idiopathic arthritis: answers and questions. *J Rheumatol* 2000; 27(8):1830-1833.
19. Davila L, Ranganathan P. Pharmacogenetics: implications for therapy in rheumatic diseases. *Nat Rev Rheumatol* 2011; 7(9):537-550.

20. Dervieux T, Greenstein N, Kremer J. Pharmacogenomic and metabolic biomarkers in the folate pathway and their association with methotrexate effects during dosage escalation in rheumatoid arthritis. *Arthritis Rheum* 2006; 54(10):3095-3103.
21. Fisher MC, Cronstein BN. Metaanalysis of methylenetetrahydrofolate reductase (MTHFR) polymorphisms affecting methotrexate toxicity. *J Rheumatol* 2009; 36(3):539-545.
22. Hider SL, Thomson W, Mack LF, Armstrong DJ, Shadforth M, Bruce IN. Polymorphisms within the adenosine receptor 2a gene are associated with adverse events in RA patients treated with MTX. *Rheumatology (Oxford)* 2008; 47(8):1156-1159.
23. Malik F, Ranganathan P. Methotrexate pharmacogenetics in rheumatoid arthritis: a status report. *Pharmacogenomics* 2013; 14(3):305-314.
24. Ranganathan P, Culverhouse R, Marsh S, Mody A, Scott-Horton TJ, Brasington R et al. Methotrexate (MTX) pathway gene polymorphisms and their effects on MTX toxicity in Caucasian and African American patients with rheumatoid arthritis. *J Rheumatol* 2008; 35(4):572-579.
25. Stamp LK, Chapman PT, O'Donnell JL, Zhang M, James J, Frampton C et al. Polymorphisms within the folate pathway predict folate concentrations but are not associated with disease activity in rheumatoid arthritis patients on methotrexate. *Pharmacogenet Genomics* 2010; 20(6):367-376.
26. Weisman MH, Furst DE, Park GS, Kremer JM, Smith KM, Wallace DJ et al. Risk genotypes in folate-dependent enzymes and their association with methotrexate-related side effects in rheumatoid arthritis. *Arthritis Rheum* 2006; 54(2):607-612.
27. Wessels JA, Kooloos WM, De JR, De Vries-Bouwstra JK, Allaart CF, Linssen A et al. Relationship between genetic variants in the adenosine pathway and outcome of methotrexate treatment in patients with recent-onset rheumatoid arthritis. *Arthritis Rheum* 2006; 54(9):2830-2839.
28. Wessels JA, De Vries-Bouwstra JK, Heijmans BT, Slagboom PE, Goekoop-Ruiterman YP, Allaart CF et al. Efficacy and toxicity of methotrexate in early rheumatoid arthritis are associated with single-nucleotide polymorphisms in genes coding for folate pathway enzymes. *Arthritis Rheum* 2006; 54(4):1087-1095.
29. van Dijkhuizen EP, Wulffraat NM. Prediction of methotrexate efficacy and adverse events in patients with juvenile idiopathic arthritis: a systematic literature review. *Pediatr Rheumatol Online J* 2014; 12:51.
30. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004; 31(2):390-392.
31. Bulatovic M, Heijstek MW, van Dijkhuizen EH, Wulffraat NM, Pluijm SM, De JR. Prediction of clinical non-response to methotrexate treatment in juvenile idiopathic arthritis. *Ann Rheum Dis* 2012; 71(9):1484-1489.
32. Van Buuren S, Groothuis-Oudshoorn K. MICE: Multivariate Imputation by Chained Equations in R. *J Stat Soft* 2011; 45(3):1-67.
33. Hosmer DW, Lemeshow S. Applied Logistic Regression. 1989. New York: John Wiley & Sons.
34. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; 15(4):361-387.
35. Steyerberg EW, Harrell Jr FE, Borsboom GJJM, Eijkemans MJC, Vergouwe Y, Habbema JD. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology* 2001; 54(8):774-781.
36. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York: Springer; 2009.
37. Schmeling H, Biber D, Heins S, Horneff G. Influence of methylenetetrahydrofolate reductase polymorphisms on efficacy and toxicity of methotrexate in patients

with juvenile idiopathic arthritis. *J Rheumatol* 2005; 32(9):1832-1836.

38. Tukova J, Chladek J, Hroch M, Nemcova D, Hoza J, Dolezalova P. 677TT genotype is associated with elevated risk of

methotrexate (MTX) toxicity in juvenile idiopathic arthritis: treatment outcome, erythrocyte concentrations of MTX and folates, and MTHFR polymorphisms. *J Rheumatol* 2010; 37(10):2180-2186.

Chapter

GENERAL DISCUSSION

11

JUVENILE IDIOPATHIC ARTHRITIS

In the area of auto-immune diseases, many pathologies are defined and grouped together based on prevailing symptoms.^{1,2} For example, juvenile idiopathic arthritis (JIA), subject of **Part I** of this thesis, is defined as any arthritis of unknown cause, lasting more than six weeks occurring before the age of 16.³⁻⁵ Thus, JIA is an umbrella-term covering a broad spectrum of different diseases, all of which manifest chronic arthritis.^{1,2} JIA has been classified into subgroups,³ but the current classification is far from perfect.^{1;2;6;7} Even within categories, disease severity differs widely. For example, for the oligoarticular category disease activity can range from the inflammation of a single joint to the involvement of multiple joints mimicking polyarthritis, so-called extended oligoarthritis.³ Consequently, the prognosis of children with JIA is variable, even within categories,^{8;9} and even though many children with JIA achieve inactive disease or disease remission,¹⁰⁻¹³ still a substantial number continue to present active disease and develop joint damage, which might eventually result in disability.¹⁴⁻²²

In therapeutic strategies for adult rheumatology, the concept of window-of-opportunity has been introduced.²³⁻²⁶ This denotes the idea of a short time frame at the onset of symptoms in which the disease is more susceptible to treatment. Hence, better outcomes are achieved with less effort if therapy is started within this time frame. In paediatric rheumatology, this concept has not been proven as rigorously as in adult rheumatology,²⁷⁻³⁰ but it is nonetheless hypothesised to exist. At any rate, early and effective treatment of JIA is beneficial to any child, as it reduces the period of time spent with active disease and resulting pain, disabilities, distress and risk of developing joint damage.

Nowadays, effective drugs are available for the treatment of JIA, which are usually administered in a step up approach.³¹ Treatment is typically started with intra-articular steroid injections. In case of insufficient response or precocious relapse of disease, systemic treatment with methotrexate (MTX) is started. MTX is a cornerstone in the treatment of JIA and shows efficacy in about 70% of patients.^{11;12} In case MTX does not suffice, biological agents, such as tumour necrosis factor (TNF)- α inhibitors,³²⁻³⁶ interleukin (IL)-1 inhibitors³⁰ or IL-6 inhibitors^{32;37;38} can be added to the treatment. Since typically a few months will pass between each step, in which the effect of current treatment is assessed, it can easily take up to one year, or even longer, before a patient who needs biological therapy, receives adequate treatment. Given the abovementioned goal of initiating effective therapy early in the course of disease, this is clearly undesirable.

On the other hand, it is not feasible to administer biological agents to just any child with JIA. This would lead to an unduly increase in the costs of health care and an unnecessary exposure of many children to potential side effects, such as an increased risk of infections,^{36;37;39-47} caused by treatment they do not need. Therefore, we are in need of models that predict early in the course of disease, which children will need biological therapy.

Much research has already been performed in this area. These studies were reviewed in **Chapter 2**. From this review, it appeared, however, that most studies were retrospective and had only been concerned with easily accessible, clinical information. These variables included the JIA category, the number and distribution of active joints, the physician's global assessment of disease activity and the parent's or patient's global assessment of well-being, as well as routine laboratory parameters, such as a full blood count and the erythrocyte sedimentation rate. The emergent picture from **Chapter 2** was that these variables could not accurately predict the prognosis of JIA, since, even though some of these variables were found to be predictive in some studies, others could not confirm these findings, or validation had not been performed yet. This conclusion held for predictors of all outcomes assessed in **Chapter 2**, i.e., disease activity, joint damage, functional ability and quality of life. An exception to this general rule was the finding of a worse prognosis for the polyarticular categories of JIA in multiple studies, but as noted above, prognosis differs even within the categories.

Since the publication of the review in **Chapter 2**, another study was performed in which an attempt was made to create a model for the prediction of the prognosis of JIA, using clinical variables only.⁴⁸ This model achieved a concordance-statistic (c-statistic) of 0.87 in 75% of 609 subjects and 0.85 in the remaining 25% of subjects, used as test data. Therefore, it seemed that JIA prognosis could be predicted using clinical variables only, after all. Indeed, some of the variables in the model were also predictive in some studies included in the review in **Chapter 2**. However, it is too early to draw the conclusion that the prognosis of JIA can be predicted using clinical variables only, due to some ambiguity in the methodology. First of all, the outcome of the prediction model was developed ad-hoc and had not been validated, making it unclear exactly what was being predicted by the model. Notably, this outcome allowed some residual disease activity up to 18 months after diagnosis even in the category designated as good prognosis. Yet, such disease activity is hardly reconcilable with the aim to induce early disease remission. Secondly, the adopted procedure of backward stepwise predictor selection in training data is known to produce optimism, since all of these steps are data driven and are not necessarily replicable in test data.⁴⁹ Yet, the optimism, i.e., the difference between the c-statistic in training data and test data, is surprisingly small in this study. It is unclear why this is so, but might be an indication that the c-statistic in test data is still too optimistic. We tried to apply the model to the data we collected for our study in **Chapter 3**. Since the ad-hoc developed outcome was not available to us, we used the number of visits in inactive disease according to Wallace criteria instead,⁵⁰⁻⁵³ dichotomised in such a way that the frequency of favourable and unfavourable outcome in our cohort was similar to the one reported in the study.⁴⁸ The motivation for this was that a robust model would predict the prognosis of JIA accurately, independently of the outcome used. Wallace criteria are the most clear cut criteria to distinguish active and inactive disease available nowadays. The area under the curve (AUC) of the receiver operating characteristics (ROC)

curve was 0.54, indicating poor performance. This performance increased to an AUC of 0.68 when dichotomising patients into a group without any follow up visit in inactive disease *versus* those who had at least one visit in inactive disease. Note, however, that in this case many patients with an unfavourable prognosis are misclassified in the group of patients with a favourable prognosis (e.g., patients with only one visit in inactive disease). Moreover, the improved performance of the model was solely due to the fact that not many patients ($n = 11$) were classified into the group with unfavourable prognosis and that the model likewise classified few patients as having an unfavourable prognosis ($n = 8$). Yet, only 1 out of 11 patients with a truly unfavourable prognosis was classified correctly. Thus, the predictive performance of the model remained poor.

These results made clear that further studies were needed. More in particular, in comparison with the studies reviewed in **Chapter 2**, prospective studies of JIA patients enrolled at diagnosis, assessing not only clinical parameters, but also other categories of potential predictors, were still lacking. Ideally, these studies should assess the outcome using validated criteria, such as Wallace criteria, taken longitudinally. The latter requirement is important, since JIA has a fluctuating disease course,^{51;54;55} implying that an analysis at a single time point can lead to misclassification of patients.

We performed such a prospective study (**Chapter 3**), in which we enrolled 169 treatment-naïve JIA patients in a short time frame after the start of the first symptoms (at most six months). Besides clinical parameters, we collected a faecal sample at baseline to characterise the gut microbiota composition in all patients. Additionally, a large panel of inflammation-related compounds was determined in blood plasma. Patients were followed every six months for two years and disease activity was assessed using the internationally validated Wallace criteria.⁵⁰⁻⁵³ The outcome was modelled longitudinally, making use of statistical techniques that allow repeated measurements of a single patient at variable moments in time, such as generalised estimating equations.⁵⁶⁻⁵⁸ Moreover, we performed survival analysis, modelling the time to first inactive disease using a Cox proportional hazards model,⁵⁹ as well as the time to inactive disease, adopting a recurrent event approach.⁶⁰

When taking all patients in the cohort together ($N = 152$, after exclusion of patients lost to follow up and those enrolled after the start of MTX), no accurate prediction model could be developed (AUC <0.65 in test data). However, when analysing subgroups of patients separately, i.e., oligoarticular patients, polyarticular rheumatoid factor (RF) negative patients and antinuclear antibody (ANA) positive patients, better performing models were obtained (AUC in test data 0.69-0.72). Significant predictors in these models were the juvenile arthritis disease activity score (JADAS)-71 and the relative abundance of the operational taxonomic unit (OTU) *Mogibacteriaceae* for oligoarticular patients; the duration of morning stiffness, haemoglobin level and the chemokine C-X-C motif ligand 9 (CXCL-9) level for polyarticular RF negative patients; and the duration of morning stiffness and haemoglobin level for ANA positive patients. Granted, the predictive performance of the models was still moderate, implying that further refinement is needed.

Nonetheless, these findings highlighted a few important aspects. First of all, with respect to the predictors, the association of morning stiffness with JIA outcome (which was independent of the statistical algorithm chosen and one of the few associations also in the entire cohort of patients) underlined the value of patient-reported measures. This finding was re-emphasised since the presence of morning stiffness was also an important determinant of patient dissatisfaction with their condition [Del Giudice *et al.*, in preparation]. Patient satisfaction should be an important – though not the sole – goal of treatment in JIA. Moreover, the patient's self-assessment was also found to be a crucial component in a decision-rule for the escalation of therapy to biologicals [Swart *et al.*, submitted]. All these results combined clearly advocate the role of patients as partners of the physician in the treatment of JIA: the patient's evaluation of disease activity is predictive of the outcome, is related to their satisfaction and leads to better decision rules regarding the escalation of therapy.

The finding of CXCL-9 as a predictor in polyarticular RF negative patients was also very interesting, as this chemokine is induced by interferon (IFN)- γ and is itself an inducer and chemoattractant of T helper (Th)-1 and Th-17 cells.⁶¹⁻⁶⁴ Since an imbalance between Th-1 and Th-17 cells on the one hand and regulatory T cells on the other hand plays a crucial role in the pathogenesis of JIA,^{5,65} this chemokine might also be involved. This hypothesis merits evaluation and validation in future studies. Next, the OTU *Mogibacteriaceae* was a predictor in oligoarticular patients. Nothing is known about this OTU in the context of rheumatologic disorders, but in our own comparison (**Chapter 4**) no differences in relative abundance were found between JIA patients and healthy children. Here, too, our research opened ways towards new studies to assess if gut microbiota in general and specifically *Mogibacteriaceae* are involved in the pathogenesis of JIA, or if our finding was due to a statistical correlation without mechanistic implications (see also below in the discussion of **Chapter 4**). Finally, not surprisingly, increased disease activity at baseline, as indicated by a higher JADAS-71 score in oligoarticular patients and a lower haemoglobin level (inversely correlated with the erythrocyte sedimentation rate) in polyarticular RF negative patients and ANA positive patients predicted a worse prognosis, implying that patients with more active disease should be eligible to get more extensive treatment earlier, something which is already routinely done.³¹

A second important learning point from the results in **Chapter 3** stems from the fact that the predictive accuracy of the models improved when they were fitted to more homogeneous subgroups of patients. This suggests that the heterogeneity of JIA impeded the accurate prediction of the prognosis. Indeed, the prognosis of JIA is probably dependent on the various diseases underlying it, which to date are unknown. To gain insight into these pathologies, biomarkers should be sought that delineate homogeneous subgroups of JIA patients, similar to IL-18⁶⁶⁻⁶⁸ and the S100 proteins⁶⁹⁻⁷¹ for the systemic category of JIA. Indeed, already over a decade ago, some suggested to revise the JIA classification to this end.^{1,2,6,7} One suggestion was to categorise ANA positive patients separately, as

these would constitute a distinct category of patients.^{1,2,6,7} It will be interesting to see if CXCL-9 can fulfil such a role as biomarker of disease. The identified cut-off point of 30 pg/ml in our cohort may serve as a starting point to delineate a subgroup of JIA patients and observe if these have similar disease characteristics and outcomes.

Further insight can be gained by clarifying the aetiology of JIA. This is what **Chapter 4** contributes to. As already mentioned in the general introduction to this thesis, JIA, like other autoimmune diseases, is thought to arise from environmental factors in genetically predisposed individuals.^{5,65} One of these environmental factors could be the gut microbiota composition.⁷²⁻⁷⁴ Several observations support this hypothesis. For example, the number of Th-17 and Forkhead box P3 (FoxP3) positive regulatory T cells was reduced in mice lacking microbiota.⁷²⁻⁷⁴ Administration of stools from healthy mice to these germ-free mice led to a reconstitution of Th17 numbers.⁷⁵ These cell populations are of critical importance in immune homeostasis and play a key role in the pathogenesis of autoimmune diseases such as JIA.^{5,65,73} Additional observations described an increased risk to develop autoimmune diseases following the use of antibiotics in childhood. This risk was dependent on the timing, number and dosage of the antibiotic prescriptions.^{76,77} Furthermore, infants who were breast fed for less than four months had a higher risk to develop JIA⁷⁸ and possibly also multiple sclerosis.⁷⁹ Finally, birth by caesarean section and early introduction of gluten increased the risk of autoimmune (type 1) diabetes.^{80,81} Since antibiotics, diet, breastfeeding and caesarean section all induce changes in the gut microbiota composition, it is tempting to hypothesise that all of these associations are mediated by alterations in the microbiota. It is, however, currently unclear exactly how the microbiota would exert this role. Hypothesised mechanisms include molecular mimicry, bystander activation and antigenic modification.⁷²⁻⁷⁴

Motivated by these observations and hypotheses, we undertook to study differences in the composition of gut microbiota between JIA patients and healthy children (**Chapter 4**). We demonstrated that the composition of the gut microbiota was different between treatment-naïve, Italian JIA patients at baseline and Italian healthy controls of similar age as the patients. No differences were found between patients at baseline, in inactive disease and during persistent activity, suggesting that the distinct profile was related to JIA, rather than disease activity. Indeed, correlations with disease activity scores were weak also at baseline. Differences in the gut microbiota composition between Dutch patients and age- and gender-matched healthy children were less evident, i.e., no evidence of differences remained after false discovery rate correction. Most variability in the microbiota profile was related to the age of the subjects both in Italian and Dutch patients. This could be part of the explanation of the differences between the Italian and Dutch populations in our study, since Dutch patients were on average older than Italian patients.

Everything taken together, there were differences in gut microbiota composition between JIA patients and healthy children. It is impossible at this moment to attribute

a causal role to gut microbiota in the development of autoimmune diseases. Part of the reason is that none of the OTUs was differently abundant in Italian patients and Dutch patients alike. A probable explanation for this observation is that differences in diet and potentially other environmental factors induce differences in the gut microbiota composition between these geographically diverse regions. Nevertheless, the observation that patients with similar disease do not share similar gut microbiota compositions, raises some doubts whether microbiota play a role in the development of JIA. It may however be that different microbial profiles play a role in Dutch and Italian populations, perhaps leading to the somewhat different presentation of the disease in both populations.

A more fundamental impediment in the attribution of causality is that potential causal pathways remain concealed in our associative study, as alluded to in the general introduction to this thesis. Generally speaking, four possible explanations of the association between gut microbiota and JIA exist: First, alterations in the microbiota profile might predispose an individual to develop JIA. However, alternatively, the causal pathway might be reversed, the immune inactivation and loss of tolerance in JIA leading to alterations in the gut microbiota profile. The third possibility is that both JIA and the alterations in the gut microbiota composition are the consequence of another, so far unknown, factor. In this case, determination of the gut microbiota composition may still be useful as a diagnostic test, but will not lead towards a potential therapeutic target. Finally, alterations of the gut microbiota and the occurrence of JIA might be entirely unrelated. The general way forward to solve this conundrum seems to be to prove or disprove causality in a randomised placebo-controlled trial, in which faecal transplants of healthy subjects are administered to JIA patients. Similar studies have been performed in inflammatory bowel disorders (IBD), showing promising but varying preliminary results.⁸²⁻⁸⁵

JUVENILE DERMATOMYOSITIS

Part II of the thesis concerned juvenile dermatomyositis (JDM). JDM is a chronic systemic inflammatory disorder, characterised by skeletal muscle inflammation and rash.⁸⁶⁻⁸⁸ Like JIA, it has a variable prognosis with about 24-40% of patients experiencing a monocyclic course, which resolves with appropriate therapy within the course of 2 years. On the other hand, 50-60% of patients suffer from a chronic course. The overall mortality is 2-3%.⁸⁸ In comparison with JIA, even less studies have been performed to find prognostic markers for the evolution of JDM. On top of this, outcome parameters for the measurement of disease activity have not yet been unequivocally defined and validated.⁸⁹⁻⁹²

Therefore, in our study of clinical signs and symptoms associated with disease activity, we chose a pragmatic approach and used the four parameters used in the paediatric rheumatology international trials organisation (PRINTO) criteria for inactive disease.⁹¹ However, rather than dichotomising these parameters and applying the criteria, we

chose to model them as continuous markers, thus benefiting from a gain of information contained in the parameters. Moreover, we evaluated disease activity repeatedly over time. Thus, patients contributed multiple visits, in the course of various years after diagnosis.

This approach led to many statistical challenges, described in **Chapter 5**. First, since many visits occurred in disease remission, the continuous outcome parameters showed an excess of the best possible value, together with a long tail of abnormal values for visits not in remission. Thus, the outcome parameters were not normally distributed. This was resolved using a hurdle-like approach, in which the probability of achieving a non-pathological value was modelled separately from the distribution of pathological values.⁹³ Secondly, multiple visits from the same patients were not independent from each other. Therefore, subject-specific random intercepts and slopes for the time elapsed since diagnosis were introduced in the model.⁹⁴ Thirdly, the four outcome parameters were correlated. This correlation was modelled by specifying a multivariate normal distribution for the subject-specific random intercepts of the four outcome parameters. The covariance matrix for this distribution was estimated from the data. This approach was based on previously published work.⁹⁵ The model was fitted in a Bayesian framework, as this allowed greater flexibility in model definition and shifted attention towards the possible values a parameter can take, given prior assumptions and the data, rather than hypothesis testing.

The efforts paid off, as the resulting model fitted the observed values well. Various clinical signs and symptoms were associated with the four outcome parameters. **Chapter 6** discusses potential explanations for these associations and draws clinical conclusions. For example, the model drew attention to myalgia and dysphonia as signs of markedly elevated disease activity. Joint swelling and contractures were associated with impairments in the performance of the childhood myositis assessment scale (CMAS) and manual muscle testing of 8 muscle groups (MMT8). Cutaneous symptoms, such as nail fold changes and periorbital rash, were associated with strictly muscular outcome measures, such as creatine kinase (CK) and the CMAS.^{89,90} Thus, cutaneous and muscular disease were correlated to a certain degree. This finding lent support to the consensus-based expert opinion that nail fold capillaroscopy should be performed regularly and that ongoing skin disease reflects ongoing systemic disease and should be treated adequately.⁹⁶ Other correlations were found between the CMAS and the MMT8, whereas the physician's global assessment of disease activity (PGA) was the parameter capturing mostly unique information. This was not surprising, given that the PGA was the only outcome parameter taking account of cutaneous disease directly.

Our results cannot be used to predict the prognosis of JDM patients, since just one parameter at diagnosis (arthritis) was associated with disease activity during follow up. However, the model could form a starting point for such prediction models. For example, since myalgia, dysphonia and haematuria were associated with considerably higher disease activity, it will be interesting to see in future studies if these parameters could

serve as clinical predictive markers. Alternatively, one could design a study to start more aggressive therapy in patients presenting with these signs and symptoms and observe if their outcome improves.

Finally, our results contribute to the discussion of the outcome parameter set to be used in JDM. Given that the CMAS and MMT8 were correlated and that the full performance of the CMAS requires quite some time, it can be recommended to merge these two measurements into one instrument capturing relevant information of both. Moreover, the finding that PGA captured unique information supports the proposed adaptation of the PRINTO criteria of inactive disease, i.e., by making a normal PGA compulsory in the definition of inactive disease.⁹² An open question that cannot be answered using our data is whether the outcome parameter set should include a specific cutaneous instrument, for example the skin disease activity score (DAS) or physician's assessment of skin disease activity on a 10-cm visual analog scale (VAS).⁹⁷

METHOTREXATE

The approach taken in **Part I** and, to a lesser extent, **Part II** of this thesis is one of predicting disease evolution prior to the start of treatment. Alternatively, one could take the start of systemic therapy for granted and predict the probability of response and the occurrence of adverse effects, instead. This is the approach taken in **Part III** where (mainly) adverse effects and efficacy of methotrexate (MTX) are studied.

Existing literature is reviewed in **Chapter 7**. Many studies ($n = 15$) assessed a range of candidate predictors for MTX efficacy, including clinical and laboratory variables, but also single nucleotide polymorphisms (SNPs) in genes involved in MTX metabolism. Some interesting candidate predictors were found, such as S100A8/A9 levels⁹⁸ (also a biomarker for systemic JIA^{69,70} and currently being studied as a biomarker for the cessation of therapy), long-chain MTX polyglutamates,⁹⁹ osteopontin level¹⁰⁰ and some SNPs in transporter genes involved in influx and efflux of MTX in and out of cells.¹⁰¹⁻¹⁰⁴ However, validation of these predictors in independent studies was lacking.

Since the publication of the literature review in **Chapter 7** another study was published in which gene expression in peripheral blood mononuclear cells (PBMCs) was analysed, in an attempt to find different patterns between responders and non-responders to MTX.¹⁰⁵ Using unsupervised clustering together with healthy controls, three clusters were identified, one enriched for healthy controls, one for responders and one for non-responders. Differentially expressed genes were involved in cytokine signalling pathways, including interleukin (IL)-1, IL-6 and IL-17, as well as the janus kinase (JAK)-signal transducer and activator of transcription (STAT) pathway, peroxisome proliferator-activated receptor (PPAR) signalling pathway and cluster of differentiation (CD) 40 pathway.

Much less research was dedicated to find predictors for the occurrence of MTX adverse effects (**Chapter 7**). Potential predictors included the alanine aminotransferase (ALT) level and SNPs in genes involved in the breakdown of MTX polyglutamates and folate metabolism. Here, too, validation in independent cohorts was lacking.

The low number of studies assessing predictors for MTX adverse events might be due to the general belief that MTX is a safe drug. Mainly, liver toxicity and gastrointestinal side effects have been described, of which the former is transient and resolves after suspension of MTX.¹⁰⁶⁻¹¹² However, gastrointestinal side effects occur frequently and can have a profound impact on the quality of life.¹¹¹ These symptoms can also occur before MTX intake (anticipatory complaints) or when being reminded of MTX (associative complaints). When these complaints occur together, they are termed MTX intolerance. These symptoms are frequently accompanied by behavioural disturbances, such as refusal to take the drug, or irritability.¹⁰⁹ MTX intolerance is thought to arise due to a conditioning reaction after the onset of gastrointestinal side effects.^{109;113} We developed the methotrexate intolerance severity score (MISS) and showed that MTX intolerance is very frequent indeed (**Chapter 8**). Up to one half of patients in a cross-sectional analysis of a cohort of JIA patients on MTX satisfied the definition of MTX intolerance (a score of at least six points on the MISS, with at least one point due to anticipatory, associative or behavioural complaints). Notably, patients using MTX subcutaneously (SC) reported MTX intolerance more often than patients on oral MTX (67.5% *versus* 44.5%, $p = 0.001$). This was surprising, as it was previously thought that MTX SC, by evading the gastrointestinal tract, would cause less gastrointestinal side effects.^{106;112} A potential explanation of this observation is that MTX intolerance is induced by the negative and sometimes painful experience of having to inject the drug, causing any gastrointestinal side effect to be valued more negatively.

These results were corroborated by our study of a German cohort of JIA patients (**Chapter 9**), in which slightly over 40% of patients were intolerant to MTX. In this cohort, when analysing the effect of route of administration on the occurrence of MTX intolerance, patients who had switched route of administration were excluded, as it was believed that the reason for switch might have been the occurrence of MTX intolerance in the past, leading to a case of confounding by indication. Indeed, the prevalence of MTX intolerance was very high in the groups who switched (>60%), indicating that switching the route of administration does not abate the complaints of MTX intolerance. After exclusion of these groups and correction by confounding variables, the odds of being MTX intolerant were higher for patients who took MTX exclusively subcutaneously (adjusted odds ratio 3.37, 95% confidence interval 1.19-10.0).

A high prevalence of MTX intolerance has subsequently also been found in other studies. In a longitudinal cohort study of Czech JIA patients, 25.5% and 30.6% presented MTX intolerance according to the MISS at 6 and 12 months after the start of MTX, respectively, whereas 45.5% of patients were intolerant to MTX at any time during one

year of follow up.¹¹⁴ No differences were seen between oral and SC MTX, but the group of patients taking oral MTX was very small ($n = 7$). High frequencies of MTX intolerance were also reported for French IBD patients on MTX (31%)¹¹⁵ and Pakistani adult rheumatoid arthritis (RA) patients on oral MTX (33.3%).¹¹⁶ Conversely, the prevalence in Dutch RA patients was much lower, but it was still more frequent in patients taking MTX SC than in patients on oral MTX (20.6% versus 6.2%, $p < 0.001$).¹¹⁷ Additionally, using a somewhat different questionnaire than the MISS, in a British study 73% of 49 adolescents with inflammatory arthritis experienced nausea, significantly more often with MTX SC than oral.¹¹⁸

Motivated by the frequency of MTX intolerance and its severity, we tried to predict its occurrence in advance, with the ultimate goal to come up with preventive strategies (**Chapter 10**). However, this proved very difficult, using clinical data and SNPs in genes involved in MTX metabolism. The c -statistic, adjusted for optimism by bootstrap, was 0.67 at most, indicating poor performance. This failure could be due to the psychological nature of MTX intolerance, which might not be predictable with these types of data.

To sum up, how should one go about the escalation of therapy in JIA? Currently, no accurate predictions of the prognosis of JIA are possible (**Part I**). Therefore, the decision to escalate to systemic therapy should be based on ad-hoc evaluation of disease activity (treat-to-target).¹¹⁹ However, given the encouraging results obtained thus far for the prediction of MTX efficacy (**Chapter 7**),¹⁰⁵ it is reasonable to expect that in the near future it will be possible to know in advance the likelihood of response to MTX for an individual patient. If it is low, a biological agent can be started; if it is high, MTX can be chosen. That leaves the question of the appropriate route of administration. The biological availability of MTX SC is higher than that of oral MTX,^{106;120-122} however no reports of increased clinical efficacy exist yet (although preliminary results of our comparison of two different studies may suggest that MTX SC is more efficacious indeed; unpublished data). Contrariwise, in two cohort studies, the efficacy of MTX SC was similar to oral MTX.^{105;123} Moreover, it has now been shown in multiple studies that MTX SC is more frequently associated with MTX intolerance (**Chapters 8 and 9**).^{117;118} Taking everything together, a valid strategy is to start MTX orally and switch to SC for individual patients, in case of inefficacy or side effects, since MTX response has been observed in patients who did not respond to oral MTX previously.¹¹² With respect to MTX intolerance, unfortunately, its occurrence cannot be predicted (**Chapter 10**). Neither can it be prevented with anti-emetics,¹¹⁵ nor treated by countermeasures, including anti-emetics¹²⁴ or cognitive behavioural therapy (unpublished data). Although no longitudinal studies regarding the persistence of MTX intolerance have been performed, the observation in **Chapter 10** that MTX restarters were more prone to develop MTX intolerance compared to MTX starters, suggests that MTX intolerance is a persistent problem indeed. Therefore, currently, should MTX intolerance ensue, the only remaining option is to switch to alternative treatment, such as biological agents.

FROM STATISTICAL MODELS TO ALGORITHM-BASED MEDICINE

The statistical analyses presented in this thesis were carried out using a wide array of statistical models, ranging from simple hypothesis testing using Mann Whitney U tests and χ^2 -tests, to logistic regression models and longitudinal models such as generalised estimating equations and survival analysis. Additionally, the JDM cohort was analysed using Bayesian statistics. These models are relatively straightforward to interpret as all of them provide point estimates of parameters such as odds ratios, hazard ratios or regression coefficients, together with a measure of uncertainty. Using such estimates, one can deduce if a certain variable increases or decreases the probability of the outcome occurring. These deductions can then be verified against existing knowledge (see, e.g., the discussion of **Chapter 6**).

Yet, already with respect to these models one can perceive some difficulties. Many clinicians will be familiar with logistic regression and survival analysis, but will not have affinity with generalised estimating equations or Bayesian statistics. This may make the results of such models more difficult to grasp and, consequently, may hamper their clinical application. Here, too, one can maybe speak of a valley of death, a term much used with respect to laboratory studies, denoting the difficulty in translating results obtained in the laboratory into clinically tangible products.¹²⁵

It is foreseeable that these problems will grow only worse. With the advent of more powerful computers and the availability of machine learning algorithms in easy-to-use software, such as R statistics, ever more clinical trials and cohort studies will be analysed using these techniques. However, many of these machine learning techniques use algorithms that are not easy to grasp. Hence, the result of the model, for example a probability to achieve the outcome, is not produced in an intuitive way, making it difficult to check if they are in rapport with prior knowledge. For example, the random forest algorithm (used in **Chapter 4**), developed almost two decades ago, roughly works as follows:¹²⁶ Data are bootstrapped to the original size. A predefined number of variables is selected at random from among all candidate predictors and these variables are searched for the best split. This procedure is repeated until a fully grown decision tree has been produced. The out-of-bag error rate for this tree is determined by predicting samples left out of the bootstrap data and comparing the predicted classification with the observed outcome. A predetermined number of trees is grown using this procedure. The probability of achieving the outcome is determined by the proportion of trees that classify a sample as having achieved the outcome. The results returned by the model contain a classification of all samples, together with an out-of-bag error rate, as well as a measure of variable importance, indicating what variables were most important for the correct classification of samples.

From this discussion it is clear that regression coefficients are lacking in the results of the random forest. Neither would one necessarily be able to verify the predicted

probabilities by inspecting all trees and classifying cases in the data accordingly. The consequence of all this is that the way the model arrived at the predicted probabilities remains somewhat obscure. Similar comments apply to other machine learning algorithms, which may be even less intuitive to understand. Using these models, only the computer can do the classification task. To sum up, previously, logistic regression models were understandable, regression coefficients could be transformed in an easily interpretable score of additions and subtractions and a cut-off point could be calculated together with its sensitivity and specificity to classify cases correctly (see, e.g., **Chapter 10**). With machine learning, the computer has to be trusted to produce the correct results.

Granted, a simple additive risk score remains preferable over complex and difficult-to-interpret statistical models. But in case the former are insufficient for accurate predictions, the latter may have to be implemented. This requires a cultural change on behalf of the scientist and the physician. The scientist needs to make sure the used techniques are understandable at least at some basic level and, more importantly, has to demonstrate unequivocally that the produced model leads to better outcomes. This implies that demonstration of good predictive performance of a model is not sufficient. Rather, a randomised trial needs to be performed to show that patients treated according to the algorithm are better off compared to the standard of care. Hence, patients need to be randomised to receive either standard of care or treatment according to the algorithm, which could involve early escalation of treatment in case of a high predicted probability of persistently active disease, or a switch to biological agents in case of a high predicted probability of non-response to MTX. The primary outcome of such a trial could be inactive disease; however, secondary outcomes should involve assessments of side effects, costs, long-term sequelae such as joint erosions, but also patient satisfaction. Acceptance rates of the algorithm will be greatly increased if the algorithm is demonstrated to be superior to standard of care in these respects. To be sure, showing that the algorithm performs well in terms of prediction accuracy in test data remains the first step. But nobody should think that to be the final step in the adoption of the algorithm.

In so doing, the care of patients with JIA (and potentially other autoimmune diseases too) will change towards an approach that can be termed algorithm-based medicine (ABM). This means that treatment decisions will be based on algorithms that predict the probability of achievement of inactive disease, response to therapy or development of adverse effects. ABM requires a change of mind set on behalf of the physician. No longer will his or her preference determine the choice of treatment and no longer will his or her experience guide therapeutic decisions. Rather, this choice will be based on a predetermined algorithm.

The potential rewards of such a paradigm shift are big. ABM has been practiced in the field of haemato-oncology for a few decades now, leading to an unprecedented increase in the survival rates of children with acute lymphoblastic leukaemia, due to algorithms that are continually being refined. It is about time children with chronic autoimmune diseases start to benefit from this approach, also.

WRONG BUT USEFUL

This thesis started with a citation of George Box: “Essentially, all models are wrong, but some are useful.”¹²⁷ By way of conclusion, how do the presented models fare in this respect?

Regarding the model for MTX intolerance, it showed poor performance after optimism correction. Therefore, it is not very likely to be used clinically. However, it served to show that prediction of MTX intolerance is unlikely to be achieved (at least, using the data available in **Chapter 10**). This means the focus should be on early detection of MTX intolerance, to permit taking appropriate action, once it ensues. On the other hand, the frequency of occurrence of MTX intolerance and the fact that it occurs more often in patients taking MTX subcutaneously has now been shown in many studies. This finding aids physicians in deciding the route of administration of MTX.

The models for the prognosis of JIA show that prediction is complicated. They may serve as a basis for future refinement. Yet, it seems that prediction of the prognosis will remain difficult due to the heterogeneity of JIA. Much attention should therefore be paid to a better characterisation of JIA and its underlying diseases. The models may contribute something in that respect, also: The predictors found in the three subgroups of patients (oligoarticular, polyarticular RF negative and ANA positive) are promising and merit further attention to elucidate their potential pathophysiologic role. This is especially true for CXCL-9, which has been related to autoimmune diseases previously and might be a biomarker of disease, similar to other biomarkers for systemic JIA. The evaluation of morning stiffness is important, as it is a good example of the value of patient reported outcomes.

Multiple studies have shown differences in gut microbiota composition between healthy children and JIA patients, making such an association very likely to exist. However, our study has shown that the relationship is complex and differs among age groups and geographical regions, but not among various disease activity states. Future research should disentangle these relationships and provide an answer to the all-important question if gut microbiota are involved in the pathogenesis of autoimmune diseases.

Finally, with respect to JDM, the associations elucidated by the model correspond to previously found patterns and can be explained with clinical observations, making it very likely that the model captured some true relationships. These clinical signs and symptoms associated with disease activity could form the basis for therapeutic decisions and a better delineation of JDM patients into more homogeneous subgroups. Moreover, the model provides input for an informed discussion about the usefulness of various outcome measures in JDM.

REFERENCE LIST

1. Martini A. Are the number of joints involved or the presence of psoriasis still useful tools to identify homogeneous disease entities in juvenile idiopathic arthritis? *J Rheumatol* 2003; 30(9):1900-1903.
2. Martini A. It is time to rethink juvenile idiopathic arthritis classification and nomenclature. *Ann Rheum Dis* 2012; 71(9):1437-1439.
3. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004; 31(2):390-392.
4. Ravelli A, Martini A. Juvenile idiopathic arthritis. *Lancet* 2007; 369(9563):767-778.
5. Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. *Lancet* 2011; 377(9783):2138-2149.
6. Ravelli A, Varnier GC, Oliveira S, Castell E, Arguedas O, Magnani A et al. Antinuclear antibody-positive patients should be grouped as a separate category in the classification of juvenile idiopathic arthritis. *Arthritis Rheum* 2011; 63(1):267-275.
7. van den Broek T, Hoppenreijns E, Meering J, Scholman R, Otten HG, Swart JF et al. Cytokine profiling at disease onset: support for classification of young antinuclear antibody-positive patients as a separate category of juvenile idiopathic arthritis. *Ann Rheum Dis* 2015; 74(2):470-472.
8. Adib N, Silman A, Thomson W. Outcome following onset of juvenile idiopathic inflammatory arthritis: I. Frequency of different outcomes. *Rheumatology* 2005; 44(8):995-1001.
9. Adib N, Silman A, Thomson W. Outcome following onset of juvenile idiopathic inflammatory arthritis: II. Predictors of outcome in juvenile arthritis. *Rheumatology* 2005; 44(8):1002-1007.
10. Guzman J, Oen K, Tucker LB, Huber AM, Shiff N, Boire G et al. The outcomes of juvenile idiopathic arthritis in children managed with contemporary treatments: results from the ReACCh-Out cohort. *Ann Rheum Dis* 2015; 74(10):1854-1860.
11. Foell D, Wulffraat N, Wedderburn LR, Wittkowski H, Frosch M, Gerss J et al. Methotrexate withdrawal at 6 vs 12 months in juvenile idiopathic arthritis in remission: a randomized clinical trial. *JAMA* 2010; 303(13):1266-1273.
12. Ruperto N, Murray KJ, Gerloni V, Wulffraat N, deOliveira SK, Falcini F et al. A randomized trial of parenteral methotrexate comparing an intermediate dose with a higher dose in children with juvenile idiopathic arthritis who failed to respond to standard doses of methotrexate. *Arthritis Rheum* 2004; 50(7):2191-2201.
13. Solari N, Palmisani E, Consolaro A, Pistorio A, Viola S, Buoncompagni A et al. Factors associated with achievement of inactive disease in children with juvenile idiopathic arthritis treated with etanercept. *J Rheumatol* 2013; 40(2):192-200.
14. Palmisani E, Solari N, Magni-Manzoni S, Pistorio A, Labo E, Panigada S et al. Correlation between juvenile idiopathic arthritis activity and damage measures in early, advanced, and longstanding disease. *Arthritis Rheum* 2006; 55(6):843-849.
15. Nielsen S, Ruperto N, Gerloni V, Simonini G, Cortis E, Lepore L et al. Preliminary evidence that etanercept may reduce radiographic progression in juvenile idiopathic arthritis. *Clin Exp Rheumatol* 2008; 26(4):688-692.
16. Taxter AJ, Wileyto EP, Behrens EM, Weiss PF. Patient-reported Outcomes across Categories of Juvenile Idiopathic Arthritis. *J Rheumatol* 2015; 42(10):1914-1921.
17. Magni-Manzoni S, Rossi F, Pistorio A, Temporini F, Viola S, Beluffi G et al. Prognostic Factors for Radiographic Progression, Radiographic Damage, and Disability in Juvenile Idiopathic Arthritis. *Arthritis Rheum* 2003; 48(12):3509-3517.
18. Oen K, Reed M, Malleson PN, Cabral DA, Petty RE, Rosenberg AM et al. Radiologic outcome and its relationship to functional

- disability in juvenile rheumatoid arthritis. *J Rheumatol* 2003; 30(4):832-840.
19. Shoop-Worrall SJW, Kearsley-Fleet L, Thomson W, Verstappen SMM, Hyrich KL. How common is remission in juvenile idiopathic arthritis: A systematic review. *Semin Arthritis Rheum* 2017.
 20. Berntson L, Nordal E, Aalto K, Peltoniemi S, Herlin T, Zak M et al. HLA-B27 predicts a more chronic disease course in an 8-year followup cohort of patients with juvenile idiopathic arthritis. *J Rheumatol* 2013; 40(5):725-731.
 21. Selvaag AM, Flato B, Dale K, Lien G, Vinje O, Smerdel-Ramoya A et al. Radiographic and clinical outcome in early juvenile rheumatoid arthritis and juvenile spondyloarthritis: A 3-year prospective study. *J Rheumatol* 2006; 33(7):1382-1391.
 22. Tollisen A, Selvaag AM, Aulie HA, Lilleby V, Aasland A, Lerdal A et al. Physical Functioning, Pain and Health-Related Quality of Life in Adults with Juvenile Idiopathic Arthritis: A Longitudinal 30-Year Follow-Up Study. *Arthritis Care Res (Hoboken)* 2017.
 23. Boers M. Understanding the window of opportunity concept in early rheumatoid arthritis. *Arthritis Rheum* 2003; 48(7):1771-1774.
 24. Mottonen T, Hannonen P, Korpela M, Nissila M, Kautiainen H, Ilonen J et al. Delay to institution of therapy and induction of remission using single-drug or combination-disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. *Arthritis Rheum* 2002; 46(4):894-898.
 25. O'Dell JR. Treating rheumatoid arthritis early: a window of opportunity? *Arthritis Rheum* 2002; 46(2):283-285.
 26. Quinn MA, Emery P. Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. *Clin Exp Rheumatol* 2003; 21(5 Suppl 31):S154-S157.
 27. Albers HM, Wessels JA, van der Straaten RJ, Brinkman DM, Suijlekom-Smit LW, Kamphuis SS et al. Time to treatment as an important factor for the response to methotrexate in juvenile idiopathic arthritis. *Arthritis Rheum* 2009; 61(1):46-51.
 28. Wallace CA, Giannini EH, Spalding SJ, Hashkes PJ, O'Neil KM, Zeff AS et al. Trial of early aggressive therapy in polyarticular juvenile idiopathic arthritis. *Arthritis Rheum* 2012; 64(6):2012-2021.
 29. Tynjala P, Vahasalo P, Tarkiainen M, Kroger L, Aalto K, Malin M et al. Aggressive combination drug therapy in very early polyarticular juvenile idiopathic arthritis (ACUTE-JIA): a multicentre randomised open-label clinical trial. *Ann Rheum Dis* 2011; 70(9):1605-1612.
 30. Vastert SJ, de JW, Noordman BJ, Holzinger D, Kuis W, Prakken BJ et al. Effectiveness of first line use of recombinant IL-1RA treatment in steroid naive systemic juvenile idiopathic arthritis: Results of a prospective cohort study. *Arthritis Rheum* 2013.
 31. Beukelman T, Patkar NM, Saag KG, Tolleson-Rinehart S, Cron RQ, DeWitt EM et al. 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: initiation and safety monitoring of therapeutic agents for the treatment of arthritis and systemic features. *Arthritis Care Res (Hoboken)* 2011; 63(4):465-482.
 32. Horneff G, Klein A, Klotsche J, Minden K, Huppertz HI, Weller-Heinemann F et al. Comparison of treatment response, remission rate and drug adherence in polyarticular juvenile idiopathic arthritis patients treated with etanercept, adalimumab or tocilizumab. *Arthritis Res Ther* 2016; 18(1):272.
 33. Lovell DJ, Giannini EH, Reiff A, Cawkwell GD, Silverman ED, Nocton JJ et al. Etanercept in children with polyarticular juvenile rheumatoid arthritis. Pediatric Rheumatology Collaborative Study Group. *N Engl J Med* 2000; 342(11):763-769.
 34. Brunner HI, Ruperto N, Tzaribachev N, Horneff G, Chasnyk VG, Panaviene V et al. Subcutaneous golimumab for children with active polyarticular-course juvenile

- idiopathic arthritis: results of a multicentre, double-blind, randomised-withdrawal trial. *Ann Rheum Dis* 2017.
35. Lovell DJ, Ruperto N, Goodman S, Reiff A, Jung L, Jarosova K et al. Adalimumab with or without methotrexate in juvenile rheumatoid arthritis. *N Engl J Med* 2008; 359(8):810-820.
 36. Lovell DJ, Reiff A, Ilowite NT, Wallace CA, Chon Y, Lin SL et al. Safety and efficacy of up to eight years of continuous etanercept therapy in patients with juvenile rheumatoid arthritis. *Arthritis Rheum* 2008; 58(5):1496-1504.
 37. Brunner HI, Ruperto N, Zuber Z, Keane C, Harari O, Kenwright A et al. Efficacy and safety of tocilizumab in patients with polyarticular-course juvenile idiopathic arthritis: results from a phase 3, randomised, double-blind withdrawal trial. *Ann Rheum Dis* 2015; 74(6):1110-1117.
 38. De Benedetti F, Brunner HI, Ruperto N, Kenwright A, Wright S, Calvo I et al. Randomized trial of tocilizumab in systemic juvenile idiopathic arthritis. *N Engl J Med* 2012; 367(25):2385-2395.
 39. Nimmrich S, Horneff G. Incidence of herpes zoster infections in juvenile idiopathic arthritis patients. *Rheumatol Int* 2015; 35(3):465-470.
 40. Becker I, Horneff G. Risk of Serious Infection in Juvenile Idiopathic Arthritis Patients Associated With Tumor Necrosis Factor Inhibitors and Disease Activity in the German Biologics in Pediatric Rheumatology Registry. *Arthritis Care Res (Hoboken)* 2017; 69(4):552-560.
 41. Horneff G, Klein A, Oommen PT, Hospach A, Foeldvari I, Feddersen I et al. Update on malignancies in children with juvenile idiopathic arthritis in the German BIKER Registry. *Clin Exp Rheumatol* 2016; 34(6):1113-1120.
 42. Windschall D, Muller T, Becker I, Horneff G. Safety and efficacy of etanercept in children with juvenile idiopathic arthritis below the age of 2 years. *Rheumatol Int* 2015; 35(4):613-618.
 43. Windschall D, Muller T, Becker I, Horneff G. Safety and efficacy of etanercept in children with the JIA categories extended oligoarthritis, enthesitis-related arthritis and psoriasis arthritis. *Clin Rheumatol* 2015; 34(1):61-69.
 44. Beukelman T, Xie F, Chen L, Baddley JW, Delzell E, Grijalva CG et al. Rates of hospitalized bacterial infection associated with juvenile idiopathic arthritis and its treatment. *Arthritis Rheum* 2012; 64(8):2773-2780.
 45. Beukelman T, Haynes K, Curtis JR, Xie F, Chen L, Bemrich-Stolz CJ et al. Rates of malignancy associated with juvenile idiopathic arthritis and its treatment. *Arthritis Rheum* 2012; 64(4):1263-1271.
 46. Davies R, Southwood TR, Kearsley-Fleet L, Lunt M, Hyrich KL. Medically Significant Infections Are Increased in Patients With Juvenile Idiopathic Arthritis Treated With Etanercept: Results From the British Society for Paediatric and Adolescent Rheumatology Etanercept Cohort Study. *Arthritis Rheumatol* 2015; 67(9):2487-2494.
 47. Wallis RS, Broder MS, Wong JY, Hanson ME, Beenhouwer DO. Granulomatous infectious diseases associated with tumor necrosis factor antagonists. *Clin Infect Dis* 2004; 38(9):1261-1265.
 48. Guzman J, Henrey A, Loughin T, Berard RA, Shiff NJ, Jurecnak R et al. Predicting Which Children with Juvenile Idiopathic Arthritis Will Have a Severe Disease Course: Results from the ReACCh-Out Cohort. *J Rheumatol* 2017; 44(2):230-240.
 49. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer; 2009.
 50. Wallace CA, Ruperto N, Giannini E. Preliminary criteria for clinical remission for select categories of juvenile idiopathic arthritis. *J Rheumatol* 2004; 31(11):2290-2294.
 51. Wallace CA, Huang B, Bandeira M, Ravelli A, Giannini EH. Patterns of clinical remission in select categories of juvenile idiopathic arthritis. *Arthritis Rheum* 2005; 52(11):3554-3562.

52. Wallace CA, Ravelli A, Huang B, Giannini EH. Preliminary validation of clinical remission criteria using the OMERACT filter for select categories of juvenile idiopathic arthritis. *J Rheumatol* 2006; 33(4):789-795.
53. Wallace CA, Giannini EH, Huang B, Irtter L, Ruperto N. American College of Rheumatology provisional criteria for defining clinical inactive disease in select categories of juvenile idiopathic arthritis. *Arthritis Care Res (Hoboken)* 2011; 63(7):929-936.
54. Fantini F, Gerloni V, Gattinara M, Cimaz R, Arnoldi C, Lupi E. Remission in juvenile chronic arthritis: A cohort study of 683 consecutive cases with a mean 10 year followup. *J Rheumatol* 2003; 30(3):579-584.
55. de Rotte MC, Luime JJ, Bulatovic M, Hazes JM, Wulffraat NM, de JR. Do snapshot statistics fool us in MTX pharmacogenetic studies in arthritis research? *Rheumatology (Oxford)* 2010; 49(6):1200-1201.
56. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; 73(1):13-22.
57. Lim LS, Pullenayegum E, Moineddin R, Gladman DD, Silverman ED, Feldman BM. Methods for analyzing observational longitudinal prognosis studies for rheumatic diseases: a review & worked example using a clinic-based cohort of juvenile dermatomyositis patients. *Pediatr Rheumatol Online J* 2017; 15(1):18.
58. Molenberghs G, Verbeke G. Models for discrete longitudinal data. New York: Springer; 2005.
59. Cox DR. Regression models and life-tables (with discussion). *J Royal Stat Soc B* 1972; 34:182-220.
60. Amorim LD, Cai J. Modelling recurrent events: a tutorial for analysis in epidemiology. *Int J Epidemiol* 2015; 44(1):324-333.
61. Karin N, Wildbaum G. The role of chemokines in adjusting the balance between CD4+ effector T cell subsets and FOXP3-negative regulatory T cells. *Int Immunopharmacol* 2015; 28(2):829-835.
62. Karin N, Wildbaum G, Thelen M. Biased signaling pathways via CXCR3 control the development and function of CD4+ T cell subsets. *J Leukoc Biol* 2016; 99(6):857-862.
63. Van Raemdonck K, Van den Steen PE, Liekens S, Van Damme J, Struyf S. CXCR3 ligands in disease and therapy. *Cytokine Growth Factor Rev* 2015; 26(3):311-327.
64. Lacotte S, Brun S, Muller S, Dumortier H. CXCR3, inflammation, and autoimmune diseases. *Ann N Y Acad Sci* 2009; 1173:310-317.
65. Lin YT, Wang CT, Gershwin ME, Chiang BL. The pathogenesis of oligoarticular/polyarticular vs systemic juvenile idiopathic arthritis. *Autoimmun Rev* 2011; 10(8):482-489.
66. Put K, Avau A, Brisse E, Mitera T, Put S, Proost P et al. Cytokines in systemic juvenile idiopathic arthritis and haemophagocytic lymphohistiocytosis: tipping the balance between interleukin-18 and interferon-gamma. *Rheumatology (Oxford)* 2015; 54(8):1507-1517.
67. Takahara T, Shimizu M, Nakagishi Y, Kinjo N, Yachie A. Serum IL-18 as a potential specific marker for differentiating systemic juvenile idiopathic arthritis from incomplete Kawasaki disease. *Rheumatol Int* 2015; 35(1):81-84.
68. Jelusic M, Lukic IK, Tambic-Bukovac L, Dubravcic K, Malcic I, Rudan I et al. Interleukin-18 as a mediator of systemic juvenile idiopathic arthritis. *Clin Rheumatol* 2007; 26(8):1332-1334.
69. Holzinger D, Frosch M, Kastrup A, Prince FH, Otten MH, van Suijlekom-Smit LW et al. The Toll-like receptor 4 agonist MRP8/14 protein complex is a sensitive indicator for disease activity and predicts relapses in systemic-onset juvenile idiopathic arthritis. *Ann Rheum Dis* 2012; 71(6):974-980.
70. Frosch M, Ahlmann M, Vogl T, Wittkowski H, Wulffraat N, Foell D et al. The myeloid-related proteins 8 and 14 complex, a novel ligand of toll-like receptor 4, and interleukin-1beta form a positive feedback mechanism in systemic-onset juvenile idiopathic arthritis. *Arthritis Rheum* 2009; 60(3):883-891.

71. Wittkowski H, Frosch M, Wulffraat N, Goldbach-Mansky R, Kallinich T, Kuemmerle-Deschner J et al. S100A12 is a novel molecular marker differentiating systemic-onset juvenile idiopathic arthritis from other causes of fever of unknown origin. *Arthritis Rheum* 2008; 58(12):3924-3931.
72. Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. *Nat Rev Genet* 2012; 13(4):260-270.
73. Verwoerd A, Ter Haar NM, de RS, Vastert SJ, Bogaert D. The human microbiome and juvenile idiopathic arthritis. *Pediatr Rheumatol Online J* 2016; 14(1):55.
74. Kamada N, Seo SU, Chen GY, Nunez G. Role of the gut microbiota in immunity and inflammatory disease. *Nat Rev Immunol* 2013; 13(5):321-335.
75. Ivanov II, Frutos RL, Manel N, Yoshinaga K, Rifkin DB, Sartor RB et al. Specific microbiota direct the differentiation of IL-17-producing T-helper cells in the mucosa of the small intestine. *Cell Host Microbe* 2008; 4(4):337-349.
76. Arvonen M, Virta LJ, Pokka T, Kroger L, Vahasalo P. Repeated exposure to antibiotics in infancy: a predisposing factor for juvenile idiopathic arthritis or a sign of this group's greater susceptibility to infections? *J Rheumatol* 2015; 42(3):521-526.
77. Horton DB, Scott FI, Haynes K, Putt ME, Rose CD, Lewis JD et al. Antibiotic Exposure and Juvenile Idiopathic Arthritis: A Case-Control Study. *Pediatrics* 2015; 136(2):e333-e343.
78. Kindgren E, Fredrikson M, Ludvigsson J. Early feeding and risk of Juvenile idiopathic arthritis: a case control study in a prospective birth cohort. *Pediatr Rheumatol Online J* 2017; 15(1):46.
79. Conradi S, Malzahn U, Paul F, Quill S, Harms L, Then BF et al. Breastfeeding is associated with lower risk for multiple sclerosis. *Mult Scler* 2013; 19(5):553-558.
80. Bonifacio E, Warncke K, Winkler C, Wallner M, Ziegler AG. Cesarean section and interferon-induced helicase gene polymorphisms combine to increase childhood type 1 diabetes risk. *Diabetes* 2011; 60(12):3300-3306.
81. Chmiel R, Beyerlein A, Knopff A, Hummel S, Ziegler AG, Winkler C. Early infant feeding and risk of developing islet autoimmunity and type 1 diabetes. *Acta Diabetol* 2015; 52(3):621-624.
82. Moayyedi P, Surette MG, Kim PT, Libertucci J, Wolfe M, Onischi C et al. Fecal Microbiota Transplantation Induces Remission in Patients With Active Ulcerative Colitis in a Randomized Controlled Trial. *Gastroenterology* 2015; 149(1):102-109.
83. Rossen NG, Fuentes S, van der Spek MJ, Tijssen JG, Hartman JH, Dufloou A et al. Findings From a Randomized Controlled Trial of Fecal Transplantation for Patients With Ulcerative Colitis. *Gastroenterology* 2015; 149(1):110-118.
84. Vermeire S, Joossens M, Verbeke K, Wang J, Machiels K, Sabino J et al. Donor Species Richness Determines Faecal Microbiota Transplantation Success in Inflammatory Bowel Disease. *J Crohns Colitis* 2016; 10(4):387-394.
85. Mattner J, Schmidt F, Siegmund B. Faecal microbiota transplantation-A clinical view. *Int J Med Microbiol* 2016; 306(5):310-315.
86. Bohan A, Peter JB. Polymyositis and dermatomyositis (second of two parts). *N Engl J Med* 1975; 292(8):403-407.
87. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). *N Engl J Med* 1975; 292(7):344-347.
88. Rider LG, Katz JD, Jones OY. Developments in the classification and treatment of the juvenile idiopathic inflammatory myopathies. *Rheum Dis Clin North Am* 2013; 39(4):877-904.
89. Lovell DJ, Lindsley CB, Rennebohm RM, Ballinger SH, Bowyer SL, Giannini EH et al. Development of validated disease activity and damage indices for the juvenile idiopathic inflammatory myopathies. II. The Childhood Myositis Assessment Scale (CMAS): a quantitative tool for the evaluation of muscle function. The Juvenile Dermatomyositis Disease Activity Collaborative Study Group. *Arthritis Rheum* 1999; 42(10):2213-2219.
90. Rider LG, Werth VP, Huber AM, Alexanderson H, Rao AP, Ruperto N

- et al. Measures of adult and juvenile dermatomyositis, polymyositis, and inclusion body myositis: Physician and Patient/Parent Global Activity, Manual Muscle Testing (MMT), Health Assessment Questionnaire (HAQ)/Childhood Health Assessment Questionnaire (C-HAQ), Childhood Myositis Assessment Scale (CMAS), Myositis Disease Activity Assessment Tool (MDAAT), Disease Activity Score (DAS), Short Form 36 (SF-36), Child Health Questionnaire (CHQ), physician global damage, Myositis Damage Index (MDI), Quantitative Muscle Testing (QMT), Myositis Functional Index-2 (FI-2), Myositis Activities Profile (MAP), Inclusion Body Myositis Functional Rating Scale (IBMFRS), Cutaneous Dermatomyositis Disease Area and Severity Index (CDASI), Cutaneous Assessment Tool (CAT), Dermatomyositis Skin Severity Index (DSSI), Skindex, and Dermatology Life Quality Index (DLQI). *Arthritis Care Res (Hoboken)* 2011; 63 Suppl 11:S118-S157.
91. Lazarevic D, Pistorio A, Palmisani E, Miettunen P, Ravelli A, Pilkington C et al. The PRINTO criteria for clinically inactive disease in juvenile dermatomyositis. *Ann Rheum Dis* 2013; 72(5):686-693.
 92. Almeida B, Campanilho-Marques R, Arnold K, Pilkington CA, Wedderburn LR, Nistala K. Analysis of Published Criteria for Clinically Inactive Disease in a Large Juvenile Dermatomyositis Cohort Shows That Skin Disease Is Underestimated. *Arthritis Rheumatol* 2015; 67(9):2495-2502.
 93. Mullahy J. Specification and testing of some modified count data models. *J Econom* 1986; 33:341-365.
 94. Pinheiro JC, Bates DM. Mixed-effects models in S and S-PLUS. New York: Springer-Verlag; 2000.
 95. Li Q, Pan J, Belcher J. Bayesian inference for joint modelling of longitudinal continuous, binary and ordinal events. *Stat Methods Med Res* 2016; 25(6):2521-2540.
 96. Bellutti Enders F, Bader-Meunier B, Baildam E, Constantin T, Dolezalova P, Feldman BM et al. Consensus-based recommendations for the management of juvenile dermatomyositis. *Ann Rheum Dis* 2017; 76(2):329-340.
 97. Campanilho-Marques R, Almeida B, Deakin C, Arnold K, Gallot N, De Iorio M et al. Comparison of the Utility and Validity of Three Scoring Tools to Measure Skin Involvement in Patients With Juvenile Dermatomyositis. *Arthritis Care Res (Hoboken)* 2016; 68(10):1514-1521.
 98. Moncrieffe H, Ursu S, Holzinger D, Patrick F, Kassoumeri L, Wade A et al. A subgroup of juvenile idiopathic arthritis patients who respond well to methotrexate are identified by the serum biomarker MRP8/14 protein. *Rheumatology (Oxford)* 2013; 52(8):1467-1476.
 99. Bulatovic Calasan M, Den Boer E, de Rotte MC, Vastert SJ, Kamphuis S, De Jonge R et al. Methotrexate polyglutamates in erythrocytes are associated with lower disease activity in juvenile idiopathic arthritis patients. *Ann Rheum Dis* 2013.
 100. Masi L, Ricci L, Zulian F, Del Monte F, Simonini G, Capannini S et al. Serum osteopontin as a predictive marker of responsiveness to methotrexate in juvenile idiopathic arthritis. *J Rheumatol* 2009; 36(10):2308-2313.
 101. Bulatovic M, Heijstek MW, Van Dijkhuizen EHP, Wulffraat NM, Pluijm SM, de JR. Prediction of clinical non-response to methotrexate treatment in juvenile idiopathic arthritis. *Ann Rheum Dis* 2012; 71(9):1484-1489.
 102. de Rotte MC, Bulatovic M, Heijstek MW, Jansen G, Heil SG, van Schaik RH et al. ABCB1 and ABCC3 gene polymorphisms are associated with first-year response to methotrexate in juvenile idiopathic arthritis. *J Rheumatol* 2012; 39(10):2032-2040.
 103. Moncrieffe H, Hinks A, Ursu S, Kassoumeri L, Etheridge A, Hubank M et al. Generation of novel pharmacogenomic candidates in response to methotrexate in juvenile idiopathic arthritis: correlation between gene expression and genotype. *Pharmacogenet Genomics* 2010; 20(11):665-676.
 104. Cobb J, Cule E, Moncrieffe H, Hinks A, Ursu S, Patrick F et al. Genome-wide data reveal novel genes for methotrexate response

- in a large cohort of juvenile idiopathic arthritis cases. *Pharmacogenomics J* 2014.
105. Moncrieffe H, Bennett MF, Tsoras M, Luyrink LK, Johnson AL, Xu H et al. Transcriptional profiles of JIA patient blood with subsequent poor response to methotrexate. *Rheumatology (Oxford)* 2017.
 106. Ramanan AV, Whitworth P, Baildam EM. Use of methotrexate in juvenile idiopathic arthritis. *Arch Dis Child* 2003; 88(3):197-200.
 107. Hashkes PJ, Laxer RM. Medical treatment of juvenile idiopathic arthritis. *JAMA* 2005; 294(13):1671-1684.
 108. Ortiz-Alvarez O, Morishita K, Avery G, Green J, Petty RE, Tucker LB et al. Guidelines for blood test monitoring of methotrexate toxicity in juvenile idiopathic arthritis. *J Rheumatol* 2004; 31(12):2501-2506.
 109. van der Meer A, Wulffraat NM, Prakken BJ, Gijsbers B, Rademaker CM, Sinnema G. Psychological side effects of MTX treatment in juvenile idiopathic arthritis: a pilot study. *Clin Exp Rheumatol* 2007; 25(3):480-485.
 110. Murray KJ, Lovell DJ. Advanced therapy for juvenile arthritis. *Best Pract Res Clin Rheumatol* 2002; 16(3):361-378.
 111. Brunner HI, Johnson AL, Barron AC, Passo MH, Griffin TA, Graham TB et al. Gastrointestinal symptoms and their association with health-related quality of life of children with juvenile rheumatoid arthritis: validation of a gastrointestinal symptom questionnaire. *J Clin Rheumatol* 2005; 11(4):194-204.
 112. Alsufyani K, Ortiz-Alvarez O, Cabral DA, Tucker LB, Petty RE, Malleson PN. The role of subcutaneous administration of methotrexate in children with juvenile idiopathic arthritis who have failed oral methotrexate. *J Rheumatol* 2004; 31(1):179-182.
 113. Figueroa-Moseley C, Jean-Pierre P, Roscoe JA, Ryan JL, Kohli S, Palesh OG et al. Behavioral interventions in treating anticipatory nausea and vomiting. *J Natl Compr Canc Netw* 2007; 5(1):44-50.
 114. Franova J, Fingerhutova S, Kobrova K, Srp R, Nemcova D, Hoza J et al. Methotrexate efficacy, but not its intolerance, is associated with the dose and route of administration. *Pediatr Rheumatol Online J* 2016; 14(1):36.
 115. Dupont-Lucas C, Grandjean-Blanchet C, Leduc B, Tripcovici M, Larocque C, Gervais F et al. Prevalence and Risk Factors for Symptoms of Methotrexate Intolerance in Pediatric Inflammatory Bowel Disease. *Inflamm Bowel Dis* 2017; 23(2):298-303.
 116. Fatimah N, Salim B, Nasim A, Hussain K, Gul H, Niazi S. Frequency of methotrexate intolerance in rheumatoid arthritis patients using methotrexate intolerance severity score (MISS questionnaire). *Clin Rheumatol* 2016; 35(5):1341-1345.
 117. Bulatovic Calasan M, van den Bosch OF, Creemers MC, Custers M, Heurkens AH, van Woerkom JM et al. Prevalence of methotrexate intolerance in rheumatoid arthritis and psoriatic arthritis. *Arthritis Res Ther* 2013; 15(6):R217.
 118. Patil P, Parker RA, Rawcliffe C, Olaleye A, Moore S, Daly N et al. Methotrexate-induced nausea and vomiting in adolescent and young adult patients. *Clin Rheumatol* 2014; 33(3):403-407.
 119. Consolaro A, Giancane G, Schiappapietra B, Davi S, Calandra S, Lanni S et al. Clinical outcome measures in juvenile idiopathic arthritis. *Pediatr Rheumatol Online J* 2016; 14(1):23.
 120. Hamilton RA, Kremer JM. Why intramuscular methotrexate may be more efficacious than oral dosing in patients with rheumatoid arthritis. *Br J Rheumatol* 1997; 36(1):86-90.
 121. Teresi ME, Crom WR, Choi KE, Mirro J, Evans WE. Methotrexate bioavailability after oral and intramuscular administration in children. *J Pediatr* 1987; 110(5):788-792.
 122. Tukova J, Chladek J, Nemcova D, Chladkova J, Dolezalova P. Methotrexate bioavailability after oral and subcutaneous administration in children with juvenile idiopathic arthritis. *Clin Exp Rheumatol* 2009; 27(6):1047-1053.
 123. Klein A, Kaul I, Foeldvari I, Ganser G, Urban A, Horneff G. Efficacy and safety of oral and parenteral methotrexate therapy in children with juvenile

- idiopathic arthritis: an observational study with patients from the German Methotrexate Registry. *Arthritis Care Res (Hoboken)* 2012; 64(9):1349-1356.
124. Scheuern A, Tyrrell PN, Haas JP, Hugle B. Countermeasures against methotrexate intolerance in juvenile idiopathic arthritis instituted by parents show no effect. *Rheumatology (Oxford)* 2017; 56(6):901-906.
125. Petrelli A, Prakken BJ, Rosenblum ND. Developing translational medicine professionals: the Marie Skłodowska-Curie action model. *J Transl Med* 2016; 14(1):329.
126. Breiman L. Manual on setting up, using and understanding random forests v3.1. www.stat.berkeley.edu/~breiman/Using_random_forests_V3.1.pdf . 2002. Berkeley, University of California. 26-6-2017. Ref Type: Online Source
127. Box GEP, Draper NR. Empirical model-building and response surfaces. New York: John Wiley & Sons; 1987.

Chapter

12

SUMMARY
NEDERLANDSE SAMENVATTING
DANKWOORD
CURRICULUM VITAE
LIST OF PUBLICATIONS

SUMMARY

Juvenile idiopathic arthritis (JIA) is an umbrella-term covering a group of autoimmune disorders characterised by chronic arthritis of unknown cause, lasting at least six weeks and with an onset before 16 years of age. It is currently impossible to predict the prognosis of children with JIA leading to difficulties in the determination of the right treatment for each patient.

In **Part I, Chapters 2 and 3** we set out to find predictors for the prognosis of children with JIA. **Chapter 2** describes a systematic literature review, summarising known predictors for four different disease outcomes, i.e., disease activity, joint damage, functional ability and quality of life. Most articles in the literature were retrospective and considered clinical and routine laboratory variables only. Many candidate predictors were assessed in only one study and were not validated in an independent one. Nevertheless, a longer diagnostic delay predicted a worse prognosis in more than one study. Likewise, the polyarticular categories of JIA were associated with worse outcomes in multiple studies and symmetric joint involvement predicted more joint damage and disability in two studies. Higher disease activity at baseline was not unequivocally associated with higher disease activity during follow up. Involvement of the hips, ankles and sacroiliac joints, as well as enthesitis were risk factors for persistent disease activity in one study each. We concluded that predictors lacked validation and that prospective studies assessing other categories of predictors were needed.

We performed such a prospective, observational study of a well-defined cohort of 152 treatment-naïve JIA patients, enrolled within six months after the onset of symptoms, to predict disease activity in the first two years of disease. The results are presented in **Chapter 3**, showing that no accurate model could be developed in the whole cohort taken together (area under the curve [AUC] of the receiver operating characteristics [ROC] curve <0.65 in test data). However, prediction accuracy improved upon considering more homogeneous subgroups separately. For oligoarticular patients, lower odds of inactive disease were predicted by higher juvenile arthritis disease activity score (JADAS)-71 and higher relative abundance of the operational taxonomic unit (OTU) *Mogibacteriaceae* at baseline (AUC of the model in test data 0.69). For polyarticular rheumatoid factor (RF) negative patients, lower odds of inactive disease were predicted by longer duration of morning stiffness, lower haemoglobin levels and higher levels of the chemokine CXCL-9 at baseline (AUC 0.69 in test data). Finally, for antinuclear antibody (ANA) positive patients, lower odds of inactive disease were predicted by longer duration of morning stiffness and lower haemoglobin levels at baseline (AUC 0.72 in test data). We concluded that the heterogeneity of JIA was hampering the accurate prediction of its prognosis. However, nonetheless, the predictors for the more homogeneous subgroups merited further investigations into their potential role in the pathogenesis of JIA and could form a basis for future prediction models.

Like its prognosis, the cause of JIA including potential environmental predisposing factors, is still unknown. Previous reports suggested that gut microbiota may be involved. **Chapter 4** shows the results of the comparison of the gut microbiota composition between JIA patients at baseline ($n = 99$), in inactive disease ($n = 44$) and during persistently active disease ($n = 25$) and geography-matched healthy controls ($n = 107$) of similar age as the patients. Results showed clear differences between Italian patients at baseline and healthy controls, both when comparing OTUs individually (nine OTUs were differently abundant between patients and controls at the $p < 0.05$ level after false discovery rate [FDR] correction) and when classifying patients and controls using random forest classification (AUC > 0.99 in test data, indicating almost perfect classification). Fifteen OTUs were differently abundant between Dutch patients at baseline and healthy controls ($p < 0.05$), but no evidence of differences ($p > 0.05$) remained after FDR correction. The gut microbiota composition was also associated with the age of the subjects, but not with disease activity status. Hence, our study provided evidence for alterations in the gut microbiota composition in JIA patients; however the relationship between gut microbiota composition and JIA was complex. In future studies, the results should be replicated and a potential causal role of the gut microbiota in the pathogenesis of JIA should be investigated.

Part II focused on juvenile dermatomyositis (JDM). This is an autoimmune pathology involving mainly skeletal muscles and the skin, but other organ systems might be involved as well. Like in JIA, predictors for the prognosis of JDM are largely unknown.

Using data available from a large multicentre cohort of JDM patients in the United Kingdom, we developed a model of disease activity using four parameters of the recently developed inactive disease criteria, being the creatine kinase (CK) level, childhood myositis assessment scale (CMAS), manual muscle testing of 8 muscle groups (MMT8) and the physician's global assessment of disease activity (PGA). **Chapter 5** describes the statistical background of the model, related to the peculiarities in the distributions of the outcome parameters (non-normal with a peak of best-possible values and a long tail towards pathological values), dependencies among multiple visits of the same patient and correlations among the four different outcome parameters. These challenges were resolved using a hurdle-type approach, in which the best-possible value on the respective scales of the outcome parameters was modelled separately from the remainder of values. Interdependencies among visits were resolved using a by-subject random intercept and random slopes for the time elapsed since diagnosis. Finally, the correlations among the outcome parameters were modelled explicitly by specifying a multivariate normal distribution for the subject-specific random intercepts of the four outcome parameters.

These efforts resulted in a well-fitting model that followed the observed patterns over time. Out-of-sample predictions were accurate, but showed wide prediction intervals due to uncertainties in the predictions, parameter estimation and imputation of missing values.

Chapter 6 discusses the clinical interpretation of the results. Myalgia and dysphonia were associated with worse disease activity. Many cutaneous symptoms were associated with the PGA, but, remarkably, periorbital rash and nail fold changes were also associated with strictly muscular outcomes. CMAS and MMT8 were lower in the presence of contractures, whereas calcinosis was associated with higher CMAS and higher PGA values. Arthritis at diagnosis, the only parameter at diagnosis associated with any of the outcome parameters, predicted higher CMAS and MMT8 scores, whereas joint swelling at the visit was associated with lower CMAS and higher PGA values. The model indicated furthermore that, of the four outcome parameters, CMAS and MMT8 were correlated, whereas PGA captured mainly unique information. These results can be used to build prediction models for patients with JDM and diversify treatment for different categories of patients. Furthermore, information about the correlation among the outcomes is useful for the determination of the right set of disease activity parameters in JDM. CMAS and MMT8 might be summarised into a single instrument, whereas PGA should be maintained as outcome parameter in order to capture cutaneous disease activity adequately.

Finally, **Part III** deals with methotrexate (MTX) a cornerstone in the treatment of JIA and also frequently used for JDM. In **Chapter 7** we review the literature systematically to find out what is known about predictors for MTX efficacy and adverse effects. Many well-performed studies were found with respect to MTX efficacy. Predictors for poor response to MTX were higher childhood health assessment questionnaire (CHAQ) score, bilateral wrist involvement, ANA negativity, lower MRP8/14 (S100A8/A9) levels, lower levels of long-chain MTX polyglutamates, higher osteopontin levels and lower haemoglobin levels, as well as single nucleotide polymorphisms (SNPs) in some genes involved in the transport of MTX in and out of the cell. Many of these predictors, however, had not been validated in independent studies. Much less was known about predictors for MTX adverse events. Potential predictors were higher levels of alanine aminotransferase (ALT) and SNPs in genes involved in MTX polyglutamate breakdown and folate metabolism.

Next, we focused on MTX adverse events, more specifically MTX intolerance. MTX intolerance is defined as gastrointestinal side effects occurring not only after the intake of MTX, but also before intake (anticipatory), or when being reminded of MTX (associative). In **Chapter 8** we discuss the development and validation of a questionnaire to measure MTX intolerance, the methotrexate intolerance severity score (MISS). The questionnaire was validated in a cross-sectional cohort of 86 JIA patients and found to distinguish well between tolerant and intolerant patients. A cut-off score of ≥ 6 (with at least one point due to anticipatory, associative or behavioural complaints) had the best diagnostic value with a sensitivity of 0.88 and a specificity of 0.80. Using this cut-off, we determined in a cross-sectional cohort of 297 patients that MTX intolerance occurred in 150 patients (50.5%), more frequently on subcutaneous (SC) MTX than oral MTX (67.5% *versus* 44.5%, $p = 0.001$).

These results were corroborated by our study of a cross-sectional cohort of 179 German JIA patients (**Chapter 9**), showing that 40.8% of patients were intolerant. Patients taking MTX exclusively SC were more frequently intolerant than patients taking MTX exclusively orally (43% *versus* 29%). There was no evidence of differences in the frequency of MTX intolerance between patients taking MTX SC and those taking MTX orally in univariate analysis ($p = 0.10$), however there were some confounders. After adjustment for those confounders, multivariable analysis confirmed higher odds of MTX intolerance for patients taking MTX SC (odds ratio 3.4, 95% confidence interval 1.2-10.0, $p = 0.02$).

In **Chapter 10**, we attempted to predict the occurrence of MTX intolerance at 6 or 12 months after the start of MTX in a prospective, observational cohort of 152 patients. The final prediction model consisted of JIA category, ANA positivity, parent/patient assessment of pain, JADAS-27, thrombocyte count, ALT level and creatinine level and showed an AUC of 0.67 (corrected for optimism by bootstrap). Introduction of SNPs in genes involved in MTX metabolism did not improve the predictive performance of the model. Thus, the results of **Chapters 8, 9 and 10** combined indicated that MTX intolerance was frequent indeed, and difficult to predict, warranting regular surveillance using the MISS to monitor its occurrence.

Taken together, this thesis provides instruments to clinicians for tailor-made treatment. In JIA, this proved challenging due to the heterogeneity of the disease. More homogeneous subgroups should be defined, ideally by identifying causes or predisposing factors of the disease, such as, potentially, the gut microbiota. Since JIA prognosis cannot yet be predicted accurately, prediction of treatment efficacy is offered as an alternative approach. Attention should be paid to potential side effects of MTX, especially MTX intolerance, which occurs frequently and which should be monitored regularly using the MISS. In JDM, various signs and symptoms were associated with increased disease activity, providing opportunities to initiate delineating more homogeneous subgroups of patients and provide tailor-made treatment.

NEDERLANDSE SAMENVATTING

Jeugdreuma (juvenile idiopathische artritis, JIA) komt voor bij ongeveer 1 op de 1000 kinderen en is daarmee de meest voorkomende kinderreumatologische aandoening. De term omvat een groep van ziekten die gekenmerkt wordt door chronische gewrichtsontsteking (artritis), met een duur van ten minste zes weken, welke ontstaat voor de leeftijd van 16 jaar (juвениel). Deze gewrichtsontsteking uit zich door gewrichtszwelling, pijn en bewegingsbeperking, met name 's morgens bij het wakker worden of na een lange periode stil gezeten te hebben. De oorzaak van JIA is onbekend (idiopathisch), dat wil zeggen dat de diagnose pas gesteld kan worden als alle bekende oorzaken van chronische gewrichtsontsteking uitgesloten zijn.

Het beloop van de ziekte bij kinderen met JIA is erg wisselend. Zo zijn er kinderen bij wie één knie ontstoken is, die goed reageert op een lokale therapie (een gewrichtsinjectie met corticosteroiden) en daarna nooit meer ontstoken raakt. Aan de andere kant van het spectrum zijn er ook kinderen met hevige ontsteking van veel gewrichten, die zelfs niet goed reageert op de nieuwste klasse geneesmiddelen (de zogenaamde biologicals). Het is op dit moment onmogelijk om vooraf te voorspellen welk kind welk beloop gaat krijgen. Dat is een groot gemis, omdat een juiste inschatting van het beloop er toe kan leiden dat we veel eerder effectieve therapie zouden kunnen geven aan wie dat nodig heeft. Op die manier zouden we de kinderen eerder ontstekingsvrij kunnen krijgen, waardoor ze dus minder last hebben, minder uitval op school en minder risico op langetermijngevolgen zoals gewrichtsschade. Daarom zijn er gegevens nodig die al op het moment van de diagnose laten zien welke kant een kind op zal gaan: of het een mild beloop van de ziekte zal hebben, dan wel of het kind een agressievere vorm van de ziekte heeft.

In **Deel I** van deze thesis proberen we hier achter te komen. **Hoofdstuk 2** beschrijft een onderzoek waarin we in de literatuur systematisch nagezocht hebben wat er al over dit onderwerp bekend was. Het bleek dat er wel veel studies gedaan waren, maar dat deze over het algemeen van matige kwaliteit waren. Bovendien waren er nog maar weinig factoren bekend die in meerdere studies lieten zien dat ze het beloop van de ziekte konden voorspellen. Daarom hebben we een nieuwe studie opgezet (**Hoofdstuk 3**) waarin we in een grote groep van 152 kinderen met JIA uit Nederland en Italië naar deze vraag gekeken hebben. Uit dit onderzoek bleek dat het voorspellen van het beloop van kinderen met JIA erg lastig is. Toch is het gelukt om enkele gegevens te vinden die het beloop voorspelden, door alle kinderen onder te verdelen in kleinere groepen die meer op elkaar leken (bijvoorbeeld door alle kinderen te nemen die maximaal vier ontstoken gewrichten hadden). De volgende gegevens voorspelden een lagere kans op het volledig ontstekingsvrij worden (en zijn dus slecht om te hebben): een langere duur van de ochtendstijfheid, meer bloedarmoede (ofwel een lagere concentratie hemoglobine), een hogere concentratie van een ontstekingseiwit in het bloed (CXCL-9), meer bacteriën van een bepaalde soort (*Mogibacteriaceae*) in de darmbacteriën en meer

ziekteactiviteit op het moment van de diagnose. Met deze gegevens was de voorspelling van het beloop van de ziekte nog niet perfect. In de toekomst zal dus meer onderzoek moeten worden gedaan om onze resultaten te bevestigen en ook om nieuwe gegevens te vinden voor het voorspellen. Maar vooral moeten kinderen met jeugdreuma beter gegroepeerd worden, in het ideale geval door achter de oorzaak van de aandoening te komen.

Zoals hierboven beschreven zijn deze oorzaken van jeugdreuma onbekend. Er wordt gedacht dat de aandoening ontstaat door de invloed van bepaalde omgevingsfactoren in kinderen die genetisch extra gevoelig zijn voor het ontwikkelen van de ziekte. Welke omgevingsfactoren dat dan zouden zijn is nog onbekend. De laatste tijd is er echter veel aandacht gekomen voor de miljoenen bacteriën die in de darm zitten. Het lijkt er op dat mensen met chronische ontstekingsziekten, zoals reumatoïde artritis, een andere samenstelling van deze darmbacteriën hebben dan gezonde mensen. In **Hoofdstuk 4** bekijken wij of de samenstelling van de darmbacteriën ook verschillend is tussen kinderen met JIA en gezonde kinderen. Het blijkt dat er forse verschillen zaten tussen Italiaanse kinderen met JIA en Italiaanse gezonde kinderen, terwijl er minder duidelijke verschillen waren tussen Nederlandse kinderen met JIA en Nederlandse gezonde kinderen. Dit zou kunnen komen door het feit dat de groep Italiaanse kinderen groter was dan de groep Nederlandse kinderen, waardoor verschillen makkelijker te zien zijn. Ook blijkt in **Hoofdstuk 4** dat de samenstelling van de darmbacteriën verandert met de leeftijd, terwijl het moment waarop het onderzoek gedaan werd (dus: op het moment van diagnose, op het moment dat kinderen volledig ontstekingsvrij zijn en op het moment dat de kinderen een terugval hebben) niet veel uitmaakt. Dit zijn belangrijke bevindingen, want het zou dus inderdaad zo kunnen zijn dat deze darmbacteriën betrokken zijn bij het ontstaan van jeugdreuma. Het is ook nog mogelijk dat jeugdreuma er juist voor zorgt dat de samenstelling anders wordt of dat de twee dingen niets met elkaar te maken hebben. Dat moet in toekomstige studies uitgezocht worden.

In **Deel II** van de thesis focussen we op een andere ziekte, namelijk juveniele dermatomyositis (JDM). Dit is een veel zeldzamere aandoening, die elk jaar bij ongeveer 2-4 op de 1 miljoen kinderen gediagnosticeerd wordt. Deze ziekte kenmerkt zich door ontsteking van de spieren, wat zich uit in spierzwakte. Daarnaast treedt ontsteking van de huid op, wat te zien is als karakteristieke roodheid in het gezicht en op de handen. Net als JIA heeft JDM een wisselend beloop waarbij een grote groep binnen twee jaar goed reageert op medicijnen, terwijl een andere grote groep telkens een terugval blijft krijgen. Ongeveer 2-3% van de kinderen overlijden zelfs aan de ziekte.

Ook hier hebben we geprobeerd om gegevens te vinden die voorspellend zijn voor het beloop. Het probleem bij JDM is alleen dat er veel minder over bekend is dan over JIA. Er is niet eens overeenstemming over de meetinstrumenten waarmee de ziekteactiviteit gemeten moet worden. Wij hebben gekozen (**Hoofdstuk 5**) voor een viertal instrumenten dat gebruikt is in de definitie voor inactieve ziekte zoals die voorgesteld is door een

internationale groep onderzoekers (PRINTO). Dit zijn: een eiwit dat vrijkomt uit spieren zodra ze ontstoken zijn (creatinekinase, CK); een serie kleine opdrachten die het kind moet uitvoeren om het uithoudingsvermogen van de spieren te testen (*childhood myositis assessment scale*, CMAS); een handmatige test van de spiersterkte in 8 spieren (*manual muscle testing of 8 muscle groups*, MMT8); en een cijfer voor de mate van ziekteactiviteit gegeven door de arts (*physician's global assessment of disease activity*, PGA). In **Hoofdstuk 5** bespreken we de uitdagingen die we tegenkwamen bij het maken van het model, en onze oplossingen daarvoor. Eén van de problemen was dat alle kinderen in de studie meerdere keren gemeten waren, soms zelfs een paar keer per jaar over een periode van meerdere jaren. We wilden al deze informatie gebruiken in het model, maar veel modellen kunnen niet op een goede manier rekening houden met het feit dat er meerdere metingen van hetzelfde kind zijn. Een ander probleem was dat veel kinderen in de studie op een bepaald moment ontstekingsvrij werden en daardoor perfecte scores voor alle vier de meetinstrumenten lieten zien, terwijl er daarnaast toch nog een aanzienlijke groep was die allerlei andere scores hadden. Deze zogenaamde scheve verdeling maakt het maken van een model lastig. Ten slotte waren de vier meetinstrumenten ook niet onafhankelijk van elkaar: als de score voor één van de instrumenten (bijvoorbeeld de CMAS) verslechtert, dan is de kans aanzienlijk dat de score van een ander instrument (bijvoorbeeld de MMT8) ook verslechtert. Daar moesten we rekening mee houden bij het maken van het model. We laten zien in **Hoofdstuk 5** dat onze oplossingen voor deze problemen werken, omdat het resultaat een model was dat de ziekteactiviteit van de kinderen in de studie vrijwel perfect voorspelde.

In **Hoofdstuk 6** bespreken we vervolgens het klinische nut van het model. Het bleek dat meerdere symptomen van de ziekte samenhangen met ernstigere ziekteactiviteit, zoals spierpijn en stemveranderingen op basis van ontsteking van de spieren van de stembanden. Ook sommige vormen van huiduitslag die veel voorkomen bij JDM hingen samen met meer ziekteactiviteit en het interessante was dat deze tekenen van huidontsteking ook een verband hadden met meetinstrumenten voor ziekteactiviteit in de spieren, zoals de CMAS en de MMT8. Dit betekent dat ziekteactiviteit in de huid vaak gepaard gaat met ziekteactiviteit in de spieren. Artsen moeten daar dus bedacht op zijn. Net als bij JIA, kunnen deze gegevens leiden tot een onderverdeling van kinderen met meer of minder ziekteactiviteit en kan de behandeling daarop aangepast worden. Bovendien bleek uit **Hoofdstuk 6** dat de CMAS en de MMT8 nauw met elkaar samenhangen. Het lijkt dus mogelijk om deze meetinstrumenten te vereenvoudigen en samen te voegen, om daardoor tijd te besparen. Aan de andere kant was de mening van de arts over de ziekteactiviteit (de PGA) onafhankelijk van de andere meetinstrumenten en is het dus belangrijk om dit instrument volledig mee te laten tellen in het beoordelen van de ziekteactiviteit.

Tot slot richten we ons in **Deel III** op één van de belangrijkste medicijnen in de behandeling van zowel JIA als JDM: methotrexaat (MTX). Dit middel is oorspronkelijk

ontwikkeld als behandeling voor kinderen met leukemie, maar enkele decennia geleden bleek het in veel lagere dosering ook goed te werken als krachtige ontstekingsremmer. Het is een veilig en effectief middel, maar toch geldt dat ongeveer 30% van de kinderen er helemaal niet op reageren. Ook komen er bijwerkingen voor, met name een milde vorm van leverirritatie (die snel over gaat als het middel tijdelijk wordt gestopt) en klachten van misselijkheid en overgeven. Nog steeds met hetzelfde doel voor ogen om van te voren te kunnen bepalen welke therapie het beste is voor welk kind, bekijken we in **Hoofdstuk 7** wat er bekend is in de literatuur over gegevens die de effectiviteit van MTX en het optreden van bijwerkingen voorspellen. Voor effectiviteit bleken er veel studies gedaan te zijn, die ook van goede kwaliteit zijn. Er zijn dus veel gegevens bekend die samenhangen met verminderde of vermeerderde effectiviteit van MTX, alleen de meeste van deze gegevens zijn slechts onderzocht in één studie. Om zeker te zijn dat deze gegevens de effectiviteit van MTX voorspellen, moeten ze nog in andere studies gecontroleerd worden. Er waren weinig studies gedaan naar gegevens die het optreden van bijwerkingen bij MTX voorspellen. Dit hangt wellicht samen met de gedachte dat MTX een veilig geneesmiddel is met weinig bijwerkingen. Wij laten in **Hoofdstuk 8 en 9** zien dat dit niet helemaal geldt. Er zijn kinderen die bijwerkingen zoals misselijkheid, overgeven en buikpijn krijgen, niet alleen na het innemen van MTX, maar ook ervoor, of op het moment dat ze aan de MTX herinnerd worden. Als deze bijwerkingen bij elkaar voorkomen, dan noemen we dat MTX intolerantie. MTX intolerantie gaat ook vaak gepaard met gedragsproblemen, zoals weigering om het middel in te nemen, of een slecht humeur op de dag dat het ingenomen moet worden. In **Hoofdstuk 8** ontwikkelen we een vragenlijst die deze bijwerking meet (*methotrexate intolerance severity score*, MISS). Met deze vragenlijst bleek dat van een grote groep van 297 kinderen met JIA, maar liefst de helft (50.5%) klachten had van MTX intolerantie. Verrassend genoeg kwam deze bijwerking vaker voor bij kinderen die de MTX met een injectie onder de huid kregen (67.5%) dan bij kinderen die de MTX als tablet innamen (44.5%). Er werd altijd gedacht dat de tablet meer klachten zoals misselijkheid veroorzaakte, maar dit blijkt dus niet zo te zijn. Ook in een grote groep Duitse kinderen met JIA kwam MTX intolerantie vaak voor, en vaker bij kinderen die de MTX als injectie kregen dan bij kinderen die de MTX als tablet namen (**Hoofdstuk 9**).

Omdat deze bijwerking zo vaak voorkomt en zo belangrijk is, probeerden we het optreden ervan ook te voorspellen (**Hoofdstuk 10**). Het bleek alleen dat dit bijna niet mogelijk was met de gegevens die wij tot onze beschikking hadden. Het is daarom belangrijk dat de arts op de hoogte is van het grote risico op MTX intolerantie, vooral bij kinderen die de MTX als injectie krijgen, zodat hij goed controleert of de bijwerking daadwerkelijk optreedt. Een behandeling is er nog niet, waardoor op dit moment de enige optie lijkt om MTX te stoppen en een ander middel te starten in het geval van MTX intolerantie.

Alles samenvattend, de studies in deze thesis richten zich met name op het voorspellen van ziekteactiviteit en het optreden van MTX intolerantie. Het uiteindelijke doel daarvan is om vroegtijdig te bepalen wat de juiste therapie is voor elk individueel kind. Voor JIA bleek dit moeilijk te gaan. Toch hebben we een aantal gegevens gevonden die aan de voorspelling van het beloop van de ziekte kunnen bijdragen. Verder moet er onderzoek naar deze ziekte komen dat gericht is op het ontrafelen van de oorzaken van de aandoening. Mogelijk dat de darmbacteriën hierbij een rol spelen. Omdat het beloop van JIA nog niet accuraat kan worden voorspeld, is het zaak om de effectiviteit van medicatie, met name MTX, te voorspellen. Hierbij moet ook gelet worden op mogelijke bijwerkingen van MTX, zoals MTX intolerantie. Hiervan hebben wij laten zien dat het vaak voorkomt, maar dat het lastig te voorspellen is. Daarom moet er frequent bepaald worden of MTX intolerantie optreedt met behulp van de MISS. In JDM hingen verschillende ziekteverschijnselen samen met ziekteactiviteit. Deze verschijnselen kunnen worden gebruikt om patiënten onder te verdelen in groepen en een op de patiënt toegespitste behandeling te starten.

DANKWOORD

What though I trace each herb and flower,
That drink the morning dew,
Did I not own Jehovah's power,
How vain were all I knew.
Say what's the rest but empty boast,
The pedant's idle claim,
Who having all the substance lost
Attempts to grasp a name.
G.F. Händel, Solomon

Als ik deze woorden schrijf, is het bijna Kerst 2017. Het feest waarop de christelijke wereld de geboorte van Jezus Christus op aarde gedenkt. Deze gebeurtenis was in het Oude Testament al door veel profeten voorspeld. Zo ook door Maleachi, in hoofdstuk 4 vers 2: "Ulieden daarentegen die Mijn Naam vreest, zal de Zon der gerechtigheid opgaan, en er zal genezing zijn onder Zijn vleugelen" (Statenvertaling). De Zon der gerechtigheid, Hij is de *Sol iustitiae* naar wie ook in het logo van de Universiteit Utrecht verwezen wordt: *Sol iustitiae, illustra nos!* Aan deze universiteit heb ik gedurende meer dan tien jaar mogen studeren en promoveren. Iets wat zonder de komst van deze *Sol iustitiae* op aarde onmogelijk zou zijn geweest. Vanaf deze plaats wens ik de universiteit en al haar medewerkers van harte de bescherming, bewaring en verlichting door Hem toe.

12

Graag wil ik iedereen bedanken die, op welke wijze dan ook, heeft bijgedragen aan mijn promotie.

DANKWOORD

Professor Wulffraat, beste Nico. Al vroeg in mijn ruim tienjarige loopbaan aan de UU kruisten onze wegen elkaar. Ik weet nog goed hoe ik als tweedejaarsstudent bij je binnenkwam. Heel erg bedankt voor alle begeleiding die je me hebt gegeven. Jouw enthousiasme en vele ideeën zijn altijd een enorme motivatie voor mij geweest. Nu moet je me ook nog leren hoe je er toch in slaagt zo enorm veel grants binnen te slepen.

Professore Martini, caro Alberto. Grazie di cuore di avermi accettato nel Suo equipe a Genova. Per me è stato un vero onore e una esperienza unica che ha arricchito molto il mio percorso di dottorato. Non solo, ovviamente, per via dell'esperienza lavorativa nell'ospedale Gaslini, ma anche per via dell'esperienza di Italia 'da dentro', piuttosto che come turista. Grazie davvero e spero che continueremo ad incontrarci spesso!

Dottoressa Malattia, carissima Clara. Tu sei proprio unica, come persona, medico e ricercatrice. Ammiro la tua creatività nel disegnare studi, il tuo ottimismo e la tua capacità

di tirare fuori qualcosa di utile dai dati scarsi che avevamo a disposizione, talvolta. Sei una grande ricercatrice e una bravissima dottoressa – i tuoi pazienti ti adorano, e tu adori loro. Ma soprattutto sei una persona unica. Grazie della collaborazione e spero tanto che la possiamo continuare nel futuro!

Professor Prakken, beste Berent. Weliswaar ben je niet mijn promotor, maar op deze plaats toch hartelijk bedankt dat ik mocht meedoen met EUtrain. Het is een ontzettend leerzame periode geweest, zowel op professioneel als op persoonlijk vlak, die mij de kansen heeft geboden om in contact te komen met vrijwel al de mensen die ik in dit dankwoord noem. Zonder EUtrain was dit een compleet ander proefschrift geworden! Heel erg bedankt!

De leden van de leescommissie, Prof. dr. J.M. van Laar, Prof. dr. E.E.S. Nieuwenhuis, Prof. dr. A.B.J. Prakken, Prof. dr. M.A.F.J. van de Laar en dr. S.M.F. Pluijm, wil ik bedanken voor de tijd en moeite die zij hebben geïnvesteerd in het lezen en beoordelen van mijn proefschrift.

All co-authors and collaborators in MD-Paedegree, thanks for the nice collaboration we have had. It has been a pleasure working with you and I am looking forward to continuing this collaboration in future. We have been able to publish some nice papers already and I am sure many more will follow.

Professor Wedderburn, dear Lucy, thanks for inviting me to London to work on the exciting JDRG database. It has been a period of hard work with a fantastic result (two papers in my thesis!) and it has been a real pleasure working with you. I am sure we will meet again in future. Dear Maria, I can hardly think of a time I learned more than during the time I worked with you. Thanks for introducing me to the wondrous world of Bayesian statistics! Claire, thank you so much for thinking along and critically revising my manuscripts. You are the co-author with the shortest response time by far.

Maja en Marloes, een dankwoord zonder jullie kan niet bestaan. Jullie zijn de eersten geweest die mij op het pad der wetenschap begeleid hebben en zonder jullie zou heel deel 3 van mijn proefschrift niet hebben bestaan! Heel erg bedankt voor de fijne en goede begeleiding! Ook Saskia en Robert en alle andere coauteurs op de MTX-stukken wil ik bedanken voor de goede samenwerking.

Tutti i genovesi, dove dovrei iniziare e finire per ringraziarvi? Grazie a tutti voi del reparto della Pediatria II: Angelo, non mi dimentico mai del lavoro che abbiamo fatto per il trial AIFA, nè del successo che abbiamo avuto! Stefania, grazie di tutte le favole che sapevi raccontare, anche se all'inizio non ne capivo «un tubo» (cit.). Nicola, Paolo, Marco,

dottressa Alpigiani, Francesca, Roberta, Gabriella e Alessia: grazie a tutti per tutto quello che mi avete insegnato. Grazie anche a tutti i collaboratori di PRINTO: Chiara, Eugenia, Francesca, Irene, Marco, Maria-Angela, Roberto. Simona, eri sempre a disposizione per dare una mano, soprattutto quando dovevano essere compilati documenti lunghi, noiosi e burocratici di cui nemmeno l'impiegato del comune capiva niente. E' stato un vero piacere essere da voi e mi mancate sempre!

Alessandro, grazie non solo della collaborazione, ma anche dell'amicizia, e qua devono essere nominate anche Maria-Chiara e la piccola Alice! Mi ricordo sempre dei tuffi a Punta Chiappa e il ristorante meraviglioso che è lì. Sono proprio contento per te che hai potuto stabilirti a Genova, perchè sei un ottimo medico e ricercatore. Stiamo disegnando un nuovo progetto proprio in questo momento e non vedo l'ora a continuare la nostra collaborazione!

Stefano, grazie per la collaborazione e la tua pazienza con me. Mi hai insegnato tante cose rispetto all'uso dell'ecografia in AIG. Spero che tu abbia trovato, o che troverai presto, un posto dove stabilirti.

Ringrazio tutti gli altri collaboratori: del reparto di fisioterapia: Anna Ronchetti, Andrea Ravaschio, Christian Savio; della radiologia, soprattutto il dottore Magnano, Francesca Magnaguagno e Francesca Maiuri; del laboratorio di immunologia, soprattutto Claudia Pastorino e Sabrina Chiesa; tutti gli specializzandi che sono passati in questi anni. Grazie a tutti!

12

DANKWOORD

Grazie Genova, per avermi accolto. Mi manchi. «Genova mia di Sturla, / che ancora nel sangue mi urla. // [...] // Genova di tutta la vita. / Mia litania infinita.» (cit. Giorgio Caproni, *Litania*, 1954).

Bas, heel erg bedankt voor je begeleiding in al deze jaren! Dank ook voor de positieve blik waarmee jij kijkt en waarmee je me regelmatig geholpen hebt als ik dingen weer eens wat te somber inzag. Dat we nog vaak goede discussies over statistiek mogen hebben! Joost, ook dank voor de goede samenwerking. Mooi dat we het treat-to-target artikel zo mooi hebben kunnen wegzetten. Succes met de laatste loodjes van je eigen promotie! Nienke, ook jou wil ik heel hartelijk danken voor de fijne samenwerking. Een aantal jaren geleden begonnen wij min of meer tegelijk als beginnende onderzoekertjes (dat zijn we nog steeds, eigenlijk) en nu eindigen we ook vrijwel gelijktijdig. Ook jij nog heel veel succes met het afronden van je promotie en met je verdere carrière.

De Loosdrecht-Vastert groep wil ik bedanken voor het feit dat jullie mij in korte tijd zoveel inzicht gegeven hebben in al het labwerk. Ik hoop en vrees dat ik een aardige luis in de

pels geweest ben op statistisch gebied. Jorg en Bas, jullie hebben een ontzettend mooie groep met veel potentie op het gebied van *translational medicine*. Ik hoop en verwacht dat jullie daar veel munt uit kunnen slaan.

Mijn roommates uit Utrecht: Henk, Michiel, Sytze en Wilco. Het was kort maar krachtig en altijd gezellig bij jullie op de kamer. Bedankt daarvoor en voor de goede gesprekken en discussies die we regelmatig hadden. Ik hoop dat *little Laura* en *Rexthor, the dog-bearer* nog lang bij jullie voortleven!

Mijn vrienden uit het Symi-groepje en de stichting vrienden van Symi. Wilco, Anneke, Jonathan, Christine, Bram en Anne-Linde, dank voor jullie vriendschap. Ik hoop dat die nog lang mag blijven bestaan! Leuk dat jullie me kwamen opzoeken in Italië. Is het niet hoog tijd voor weer een reisje?

Mijn paranimfen wil ik hartelijk bedanken. Johan, je bent een echte vriend om samen mee te praten, te discussiëren en te lachen. Ik kan veel leren van je spontaniteit en openheid. Ik hoop dat ik je ook wat terug kan geven. Henri, zwager en nu ook mijn paranimf, en Dianne, bedankt voor alles. Ik heb de vakanties die jullie bij mij in Italië doorbrachten altijd geweldig gevonden. Misschien moeten we dat nog maar eens doen.

Lieve vader en moeder, dit proefschrift is vanzelfsprekend aan jullie opgedragen. Van jongs af aan hebben jullie me gesteund en gestimuleerd om mijn weg te gaan, zelfs toen bleek dat ik voor lange tijd naar Italië zou vertrekken. Hierboven heb ik diverse mensen bedankt zonder wie delen van dit proefschrift er niet zouden zijn geweest, maar zonder jullie zou er überhaupt nooit een letter op papier gekomen zijn en zou ik niet zijn wie ik nu ben. Bedankt voor alles!

CURRICULUM VITAE

Pieter van Dijkhuizen, MD was born on January 9th, 1988 in Amersfoort, The Netherlands. He studied medicine at Utrecht University from 2006 to 2012. During medical school he joined the group of Prof. Dr. A.B.J. Prakken and Prof. Dr. N.M. Wulfraat as part of an extra-curricular honours programme and performed clinical research related to methotrexate efficacy and intolerance in children with juvenile idiopathic arthritis.

After graduating from medical school in 2012, he became an early-stage researcher in the European Translational training for Autoimmunity & Immune manipulation Network (EUTRAIN). As part of this programme he moved to Genoa (Italy) where he started his PhD research in June 2013 under supervision of Prof. Dr. N.M. Wulfraat, Prof. Dr. A. Martini and Dr. C. Malattia. Results of this research are presented in this thesis. For his work on gut microbiota in JIA, he received the gold medal Kourir award of the French organisation for children affected by chronic arthritis for the best research work that could provide new insights into the development of treatment of juvenile idiopathic arthritis.

In September 2017, he started working as a resident paediatrics (not in training) at the Meander Medisch Centrum in Amersfoort, The Netherlands.

LIST OF PUBLICATIONS

This thesis

Dijkhuizen EHP van, Aidonopoulos O, Ter Haar NM, Pires Marafon D, Magni-Manzoni S, Ioannidis YE, Putignani L, Vastert SJ, Malattia C, De Benedetti F, Martini A, for the MD-Paedigree consortium. Prediction of inactive disease in juvenile idiopathic arthritis: A multicenter, observational cohort study. [Submitted].

Dijkhuizen EHP van, Del Chierico F, Malattia C, Russo A, Pires Marafon D, Ter Haar NM, Magni-Manzoni S, Vastert SJ, Dallapiccola B, Prakken B, Martini A, De Benedetti F, Putignani L, for the MD-Paedigree consortium. Microbiome analytics of the gut microbiota in juvenile idiopathic arthritis patients: An observational, longitudinal cohort study. [Submitted].

Dijkhuizen EHP van, Deakin CT, De Iorio M, Wedderburn LR. Clinical signs and symptoms in a joint model of four disease activity parameters in juvenile dermatomyositis: A prospective, longitudinal, multicenter cohort study. [Submitted].

Dijkhuizen EHP van, Deakin CT, Wedderburn LR, De Iorio M. Modelling disease activity in juvenile dermatomyositis: A Bayesian approach. *Stat Methods Med Res* 2017 Jan 1 [Epub ahead of print].

Dijkhuizen EHP van, Pouw JN, Scheuern A, Hügler B, Hardt S, Ganser G, Kümmerle-Deschner JB, Horneff G, Holzinger D, Bulatović Čalasan B, Wulffraat NM. Methotrexate intolerance in oral and subcutaneous administration in patients with juvenile idiopathic arthritis: a cross-sectional, observational study. *Clin Exp Rheumatol* 2016 Jan-Feb;34(1):148-54.

Dijkhuizen EHP van, Bulatović Čalasan M, Pluijm SMF, Rotte MCFJ de, Vastert SJ, Kamphuis S, Jonge R de, Wulffraat NM. Prediction of methotrexate intolerance in juvenile idiopathic arthritis: a prospective, observational cohort study, *Pediatric Rheumatol Online J.* 2015 Feb 18;13:5.

Dijkhuizen EHP van, Wulffraat NM. Prediction of methotrexate efficacy and adverse events in patients with juvenile idiopathic arthritis: a systematic literature review. *Pediatric Rheumatol Online J.* 2014;12:51.

Dijkhuizen EHP van, Wulffraat NM. Early predictors of prognosis in juvenile idiopathic arthritis: A systematic literature review, *Ann Rheum Dis.* 2015 Nov;74(11):1996-2005. doi 10.1136/annrheumdis-2014-205265.

Bulatovic M, Heijstek MW, **Dijkhuizen EHP van**, Wulffraat NM, Pluijm SMF, Jonge R de. Prediction of clinical non-response to methotrexate treatment in juvenile idiopathic arthritis, *Ann Rheum Dis*. 2012 Sep;71(9):1484-9. doi 10.1136/annrheumdis-2011-200942.

Bulatovic M, Heijstek MW, Verkaaik M, **Dijkhuizen EHP van**, Armbrust W, Hoppenreijjs EPA *et al*. High prevalence of methotrexate intolerance in juvenile idiopathic arthritis determined with a validated methotrexate intolerance severity score, *Arthritis Rheum*. 2011 Jul;63(7):2007-13. doi: 10.1002/art.30367

Other publications

Dijkhuizen EHP van, Vanoni F, Magnano G, Magnaguagno F, Rosendahl K, Van Rossum MA, Martini A, Malattia C, OMERACT MRI in JIA working group. Impact of the inclusion of the metacarpophalangeal joints on the metric properties of the wrist MRI scoring system in JIA. [Submitted].

Swart JF, **Dijkhuizen EHP van**, Wulffraat NM, De Roock S. Clinical juvenile arthritis disease activity score proven to be ready for use in treat-to-target therapy in juvenile idiopathic arthritis. *Ann Rheum Dis* 2017 Nov 14. doi 10.1136/annrheumdis-2017-212104 [Epub ahead of print].

Lanni S, **Dijkhuizen EHP van**, Vanoni F, Viola S, Magnaguagno F, Magnano GM, Gandolfo C, Ravelli A, Malattia C. Ultrasound changes in synovial abnormalities induced by start of treatment in juvenile idiopathic arthritis. *Clin Exp Rheumatol* 2017 Nov 28 [Epub ahead of print].

Ravelli A, Davi S, Bracciolini G, Pistorio A, Consolaro A, **Dijkhuizen EHP van**, *et al*. Intra-articular corticosteroids versus intra-articular corticosteroids plus methotrexate in oligoarticular juvenile idiopathic arthritis: a multicentre, prospective, randomized, open-label trial. *Lancet* 2017 Mar 4;389(10072):909-16.

Prinold JAI, Mazzà C, Di Marco R, Hannah I, Malattia C, Magni-Manzoni S, Petrarca M, Ronchetti AB, Tanturri De Horatio L, **Dijkhuizen EHP van**, Wesarg S, Viceconti M. A patient-specific foot model for the estimate of ankle joint forces in patients with juvenile idiopathic arthritis. *Ann Biomed Eng* 2016 Jan;44(1):247-57.

Dijkhuizen EHP van, Carpay JA, [A bad aftertaste caused by Pine Nuts], *Ned Tijdschr Geneeskd* 2011;155:A2844

Appendix

PREDICTION OF CLINICAL NON-RESPONSE TO METHOTREXATE TREATMENT IN JUVENILE IDIOPATHIC ARTHRITIS

Maja Bulatović ,* Marloes W. Heijstek,*
E.H. Pieter van Dijkhuizen, Nico M. Wulffraat,
Saskia M.F. Pluijm and Robert de Jonge

* Contributed equally

A

ABSTRACT

Objectives

Methotrexate (MTX) is a cheap and efficacious drug in juvenile idiopathic arthritis (JIA) treatment. If JIA patients are unresponsive to MTX, early and effective combination treatment with biologicals is required to prevent joint damage. The authors developed a prediction model to identify JIA patients not responding to MTX.

Methods

In a cohort of 183 JIA patients, clinical variables and single nucleotide polymorphisms (SNPs) in genes involved in the mechanism of action of MTX were determined at the start of MTX treatment. These variables were used to construct a prediction model for non-response to MTX treatment during the first year of treatment. Non-response to MTX was defined according the American College of Rheumatology paediatric 70 criteria. The prediction model was validated in a cohort of 104 JIA patients.

Results

The prediction model included: erythrocyte sedimentation rate and SNPs in genes coding for methionine synthase reductase, multidrug resistance 1 (MDR-1/ABCB1), multidrug resistance protein 1 (MRP-1/ABCC1) and proton-coupled folate transporter (PCFT). The area under the receiver operating characteristics curve (AUC) was 0.72 (95% CI: 0.63 to 0.81). In the validation cohort, the AUC was 0.65 (95% CI: 0.54 to 0.77). The prediction model was transformed into a total risk score (range 0–11). At a cut-off of ≥ 3 , sensitivity was 78%, specificity 49%, positive predictive value was 83% and negative predictive value 41%.

Conclusions

The prediction model that we developed and validated combines clinical and genetic variables to identify JIA patients not responding to MTX treatment. This model could assist clinicians in making individualized treatment decisions.

A

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is one of the most common chronic rheumatic diseases in childhood with a reported prevalence between 16 and 159 per 100,000.¹ In the treatment of JIA, methotrexate (MTX) is the cornerstone disease-modifying anti-rheumatic drug. MTX is efficacious in 30%–70% of patients, depending on the JIA subtype.^{2,3} Patients who do not respond or partially respond to MTX are given biologicals such as tumour necrosis factor α (TNF α) inhibitors, interleukin 1 (IL-1) receptor blockers or IL-6 blockers alone or in combination with MTX. The high efficacy of these combination therapies^{4–8} is leading to a tendency to apply biologicals early in the treatment of JIA, even before knowing the patient's response to MTX monotherapy.^{9–11} This is consistent with the need for early effective treatment of JIA, crucial for preventing irreversible joint destruction and long-term disabilities.^{1,4,12} However, combination therapy is unnecessary in those patients who could respond to MTX monotherapy, given that the long-term adverse effects of biologicals, particularly TNF α blockers, are largely unknown and could include development of autoimmune phenomena such as inflammatory bowel disease and malignancies such as leukaemia and lymphoma.^{12–16} To ensure that only patients unresponsive to MTX receive early additional treatment with biologicals and those responsive to MTX are spared costly drugs with potentially serious adverse effects, it is crucial to predict those patients who will be unresponsive to MTX monotherapy.

A prediction model for MTX efficacy was successfully constructed in rheumatoid arthritis (RA).¹⁷ However, to date no model has been constructed to predict MTX non-response in JIA. The aim of this study was to develop and validate such a prediction model, using clinical and genetic predictors.

METHODS

Study design and patients

Two observational cohort studies were performed at the Wilhelmina Children's Hospital, University Medical Center Utrecht. The derivation cohort, consisting of retrospectively collected patients who had started MTX monotherapy between 1990 and 2006, was used to develop the prediction model. The validation cohort, consisting of prospectively collected patients who had started MTX monotherapy between January 2007 and June 2010, was used to test the external validity of the model.

Patients, aged 1–18 years, with a confirmed JIA according to the International League of Associations for Rheumatology criteria¹⁸ and an available blood sample were eligible for inclusion. Patients were excluded if longitudinal data after start of MTX treatment could not be retrieved and blood samples could not be used to determine the SNPs. Their clinical data on disease characteristics, disease activity and medication use were

collected from medical charts at the moment of MTX start and at 3, 6 and 12 months after MTX start. This study was approved by the University Medical Center Utrecht Medical Ethics Committee.

Assessment of MTX clinical response

Clinical response to MTX in the first year of treatment was determined using the American College of Rheumatology paediatric 70 (ACR70) criteria for disease activity.¹⁹ The validated core-set criteria²⁰ for disease activity were: (1) Physician's global assessment of disease activity on a 10 cm visual analogue scale; (2) Number of active joints, defined by joint swelling or limitation of movement accompanied by pain and tenderness; (3) Number of joints with limitation of movement; (4) Physical functional ability, measured with the Childhood Health Assessment questionnaire (CHAQ) disability on a 0–3 scale;²¹ (5) Parent or patient assessment of patient's well-being on a 10 cm visual analogue scale; and (6) Erythrocyte sedimentation rate (ESR). Good clinical response to MTX according to ACR70 criteria means at least 70% improvement in at least three of the six core-set criteria, with no more than 30% worsening in more than one of the remaining criteria.

MTX non-responders were defined as patients who did not satisfy the ACR70 criteria in at least two out of three visits during the first year of MTX treatment. This definition was used since clinical response to MTX is known to fluctuate in a large proportion of patients between different time points in the first year of treatment.²² MTX non-responders also included patients discontinuing MTX and/or switching to anti-TNF α therapy or other biologicals due to insufficient effect of MTX.

A

Clinical and genetic variables

At baseline, JIA was divided into three subtype categories: oligoarticular JIA, polyarticular JIA and other subtypes including systemic, psoriatic and enthesitis-related JIA (table 1). Other disease characteristics, core-set criteria and information on medication use are shown in table 1.

The genetic variables, single nucleotide polymorphism (SNPs), were selected based on their involvement in the MTX metabolic pathways, their high polymorphic allele frequency and documented functional effects. DNA for SNP analysis was obtained from whole blood or isolated peripheral blood mononuclear cells. Genomic DNA was isolated using the QIAmp DNA Mini Blood Kit (Qiagen, Venlo, The Netherlands). The following SNPs were determined using real-time PCR with Taqman technique according to protocols provided by the manufacturer (Taqman, Applied Biosystems, Foster City, California, USA): methylenetetrahydrofolate reductase (*MTHFR* rs1801133 and rs1801131), reduced folate carrier (*RFC/SLC19A1* rs1051266), methionine synthase reductase (*MTRR* rs1801394), inosine triphosphatase (*ITPA* rs1127354), adenosine monophosphate deaminase (*AMPD1* rs17602729), 5-aminoimidazole-4-carboxamide ribonucleotide transformylase

Table 1. Prevalence, univariate OR (95% CI) for potential predictors of ACR70 MTX non-response for derivation and validation cohorts at baseline

Variables	Derivation cohort (n=183)		Validation cohort (n=104)	
	Frequency (%)	OR (95% CI)	Frequency (%)	OR (95% CI)
Female	122 (66.7)	1.10 (0.53-2.30)	61 (58.7)	1.72 (0.76-3.90)
JIA subtype				
Oligoarticular (persistent/extended)	66 (36.1)	Reference	45 (43.3)	Reference
Polyarticular (RF negative/positive)	68 (37.2)	0.81 (0.36-1.81)	39 (37.5)	2.92 (1.13-7.52)
Other (systemic-onset/psoriatic/enthesitis-related)	49 (26.8)	1.20 (0.47-3.04)	20 (19.2)	2.04 (0.66-6.27)
Disease characteristics				
ANA positive	91 (49.7)	0.99 (0.49-1.99)	55 (52.9)	0.25 (0.10-0.65)
Age at onset	85 (46.4)	0.95 (0.47-1.91)	59 (56.7)	0.76 (0.33-1.73)
Age at MTX start	129 (70.5)	0.75 (0.34-1.67)	74 (74.0)	1.43 (0.58-3.53)
Disease duration*	113 (61.7)	1.86 (0.92-3.78)	44 (42.3)	1.77 (0.77-4.12)
Core set criteria				
Active joints	110 (60.1)	0.77 (0.37-1.59)	55 (52.9)	0.93 (0.41-2.12)
Limited joints*	126 (68.9)	0.59 (0.26-1.33)	59 (56.7)	0.73 (0.31-1.69)
PGA	150 (82.0)	0.59 (0.21-1.63)	55 (52.9)	1.32 (0.58-2.99)
ESR*	114 (62.3)	0.38 (0.16-0.93)	66 (63.5)	0.70 (0.28-1.77)
Medication†				
MTX dose	64 (35.0)	0.66 (0.32-1.36)	45 (43.3)	1.13 (0.50-2.57)
Folic acid	101 (55.2)	0.78 (0.38-1.59)	96 (92.3)	0.61 (0.12-3.18)
Single nucleotide polymorphisms				
<i>MTHFR</i> rs1801133 C>T	94 (51.3)	1.45 (0.72-2.93)	49 (47.2)	1.67 (0.73-3.80)
<i>MTHFR</i> rs1801131A>C	97 (53.0)	0.79 (0.39-1.61)	56 (53.9)	0.37 (0.16-0.87)
<i>MTRR</i> rs1801394 A>G*	147 (80.3)	0.28 (0.08-0.96)	78 (75.0)	0.62 (0.23-1.66)
<i>AMPD1</i> rs17602729 G>A	42 (22.9)	1.03 (0.45-2.39)	24 (24.0)	0.74 (0.29-1.86)
<i>ATIC</i> rs2372536 C>G	105 (57.4)	0.97 (0.48-1.96)	72 (69.2)	0.65 (0.26-1.61)
<i>ADORA2A</i> rs5751876 C>T	109 (59.6)	0.87 (0.42-1.79)	73 (70.2)	1.29 (0.54-3.08)
<i>MDR-1/ABCB1</i> rs1128503 G>A	43 (23.5)	0.39 (0.18-0.84)	32 (30.8)	0.47 (0.20-1.10)



Table 1. (continued)

Variables	Derivation cohort (n=183)		Validation cohort (n=104)	
	Frequency (%)	OR (95% CI)	Frequency (%)	OR (95% CI)
<i>MDR-1/ABCB1</i> rs1045642 G>A*	AA vs GA/GG 30 (16.4)	0.57 (0.24-1.38)	17 (16.3)	0.71 (0.25-2.07)
<i>MRP-1/ABCC1</i> rs35592 T>C*	TC/CC vs TT 76 (41.6)	0.50 (0.24-1.01)	36 (34.6)	0.75 (0.32-1.72)
<i>MRP-1/ABCC1</i> rs3784862 A>G	AG/GG vs AA 83 (47.0)	0.97 (0.48-1.97)	47 (45.2)	1.28 (0.56-2.90)
<i>MRP-2/ABCC2</i> rs171620 C>T	CT/TT vs CC 66 (36.1)	0.92 (0.45-1.91)	38 (36.6)	0.84 (0.36-1.93)
<i>MRP-3/ABCC3</i> rs4793665 T>C	TC/CC vs TT 130 (72.1)	1.54 (0.73-3.27)	65 (62.5)	2.24 (0.97-5.14)
<i>MRP-3/ABCC3</i> rs3785911 A>C	AC/CC vs AA 101 (55.2)	1.01 (0.50-2.04)	49 (47.1)	1.18 (0.52-2.65)
<i>MRP-5/ABCC5</i> rs2139560 G>A	GA/AA vs GG 127 (69.4)	1.49 (0.72-3.12)	61 (58.7)	1.57 (0.69-3.57)
<i>BCRP/ABCG2</i> rs13120400 T>C	TC/CC vs TT 85 (46.5)	1.23 (0.61-2.49)	48 (46.1)	1.85 (0.80-4.27)
<i>GGH</i> rs10106587A>C	AC/CC vs AA 90 (49.1)	0.65 (0.32-1.32)	58 (53.9)	1.14 (0.50-2.58)
<i>GGH</i> rs3758149 G>A	GA/AA vs GG 84 (45.9)	1.38 (0.68-2.81)	58 (53.9)	1.36 (0.60-3.07)
<i>PCFT</i> rs2239907 C>T*	CT/TT vs CC 125 (68.3)	0.56 (0.25-1.26)	69 (66.4)	1.44 (0.51-4.11)

Variables with >10% missing values and/or observed frequencies of <5 with respect to MTX clinical response in the derivation or the validation cohort were excluded from the univariate logistic analysis: RF, CHAQ disability, CHAQ well-being, *RFCS/SLC19A1* rs1051266 C>T, *ITPA* rs1127354 G>T, *ADA* rs73598374 C>T, *MDR-1/ABCB1* rs2032582 C>AT, *MRP-4/ABCC4* rs868853 T>C, *MRP4/ABCC4* rs2274407 C>A, *BCRP/ABCG2* rs2231142 G>T. Variable *MRP2/ABCC2* rs4148396 C>T was excluded from the univariate logistic analysis because it correlated with *MRP-2/ABCC2* rs717620 C>T.

*Variables significantly associated with MTX non-response ($p \leq 0.20$) were included in the multivariate logistic regression analysis. p values – disease duration >1 year: $p=0.09$; limited joints >1: $p=0.20$; ESR: $p=0.03$; *MTRR* rs1801394 A>G; $p=0.04$; *MDR-1/ABCB1* rs1128503 C>T; $p=0.02$; *MRP-1/ABCC1* rs35592 T>C; $p=0.05$; *PCFT* rs2239907 A>G; $p=0.16$.

†In the derivation cohort, 170 (92.9%) patients were on non-steroidal anti-inflammatory drugs (NSAIDs) and two (1.1%) patients were on sulphasalazine at MTX start. In the validation cohort, 78 (75.0%) patients were on NSAIDs and six (5.8%) patients were on sulphasalazine at MTX start.

OR, odds ratios; CI, confidence interval; ACR70, American College of Rheumatology paediatric 70; ANA, antinuclear antibody; CHAQ, Childhood Health Assessment questionnaire; ESR, erythrocyte sedimentation rate; JIA, juvenile idiopathic arthritis; MTX, methotrexate; PGA, physician global assessment; RF, rheumatoid factor.

(*ATIC* rs2372536), adenosine-deaminase (*ADA* rs73598374), adenosine A2A receptor (*ADORA2A* rs5751876), multidrug resistance 1 (*MDR-1/ABCB1* rs1128503, rs1045642, rs2032582), multidrug resistance protein 1-5 (*MRP-1/ABCC1* rs35592, rs3784862; *MRP-2/ABCC2* rs4148396, rs717620; *MRP-3/ABCC3* rs4793665, rs3785911; *MRP-4/ABCC4* rs868853, rs2274407; *MRP-5/ABCC5* rs2139560), breast cancer resistance protein (*BCRP/ABCG2* rs13120400, rs2231142), γ glutamyl hydrolase (*GGH* rs10106587, rs3758149) and proton-coupled folate transporter (*PCFT* rs2239907).

Statistical analysis

To construct a risk model to predict non-responders to MTX, backward logistic regression analysis was performed in several stages. First, all continuous clinical variables were dichotomized to facilitate the use of the model in daily clinical practice. Second, univariate ORs with 95% CI were calculated (table 1). If two potential predictors correlated (Spearman's $r \geq 0.40$), the clinically more relevant or the more significant variable in the univariate analysis was given preference. Third, to obtain the final prediction model, clinical and genetic variables with a p value of ≤ 0.20 on the log-likelihood test were combined in the multivariate logistic regression analysis.

To calculate predicted probabilities of being an MTX non-responder, we used the following formula:

$$P_{\text{MTXoutcome}} = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)}}.$$

where P is the predicted probability of being an MTX non-responder, β_0 is the constant and β_1 , β_2 and β_p represent the regression coefficients for each of the predictors x_1 , x_2 and x_p . To evaluate the predictive power of the model, we used the predicted probabilities for MTX non-response to construct a receiver operating characteristic (ROC) curve. The area under the ROC curve (AUC) measures the concordance of predictive values with actual outcomes, with an AUC of 0.5 reflecting no predictive power and an AUC of 1.0 reflecting perfect prediction. To assess whether the models fit the data well, we employed the Hosmer-Lemeshow test.

To compute the risk score of being an MTX non-responder for individual patients, the regression coefficients β of the predictors in the final model were transformed into simple scores that sum up to a total risk score (table 3). Within the total risk score, sensitivity, specificity, positive predictive value and negative predictive value were calculated for several cut-off scores.

The prediction model was externally validated in the validation cohort. To do this, we entered the regression coefficients of the predictors obtained from the derivation

cohort into the abovementioned formula. This was used to construct a ROC curve for the validation cohort. All statistical analyses were carried out with SPSS V.15.0.0 (SPSS, Chicago, Illinois, USA).

Table 2. ACR70 MTX non-response frequency (%)

Time point	Derivation cohort (n=183)	Validation cohort (n=104)
3 months	175 (95.6)	83 (79.8)
6 months	140 (76.5)	64 (61.5)
12 months	103 (56.3)	52 (50.0)
First year of treatment*	143 (78.1)	68 (65.4)

*According to definition: ACR70 non-responder in at least two out of three time points during the first year of treatment.

ACR70, American College of Rheumatology paediatric 70; MTX, methotrexate.

Table 3. Prediction model and scores for ACR70 MTX non-response

Predictors		β	Score	OR (95%CI)	p value
Clinical					
ESR	>12 mm/h	-0.820	-2	0.44 (0.17-1.12)	0.09
Genetic					
<i>MTRR</i> rs1801394 A>G	AG/GG	-1.172	-3	0.31 (0.09-1.11)	0.07
<i>MDR-1/ABCB1</i> rs1045642 G>A	AA	-0.714	-2	0.49 (0.22-1.11)	0.09
<i>MRP-1/ABCC1</i> rs35592 T>C	TC/CC	-0.793	-2	0.45 (0.21-0.98)	0.04
<i>PCFT</i> rs2239907 C>T	CT/TT	-0.569	-2	0.57 (0.24-1.36)	0.20
Constant		3.758	11		
AUC derivation cohort (95% CI)		0.72 (0.63-0.81)			
AUC validation cohort (95% CI)		0.65 (0.54-0.77)			
Hosmer-Lemeshow test (p value)		0.91			

Example: Risk score of a JIA patient having all predictors is calculated as follows: Add up the constant (11) to scores of individual predictors, namely 11+(-2)+(-3)+ (-2)+(-2)+(-2), which equals 0 points.

ACR70, American College of Rheumatology paediatric 70; ESR, erythrocyte sedimentation rate; MTX, methotrexate.

A

RESULTS

Patient characteristics

183 Patients were included in the derivation cohort after removal of five patients due to missing longitudinal data. Upon eliminating three patients receiving IL-1 receptor blockers at MTX start, 104 patients were included in the validation cohort.

Baseline characteristics (table 1) did not differ significantly between the cohorts, besides disease duration before MTX start, which was longer in the derivation (median: 1.9 years, IQR: 0.3–7.6) than in the validation cohort (median: 0.8 years, IQR: 0.3–4.5) ($p=0.001$). MTX starting dose was comparable between the two cohorts, namely 9.4 mg/m²/week in the derivation and 9.8 mg/m²/week in the validation cohort. Within the cohorts, MTX starting dose was equivalent in future responders and non-responders. However, in the derivation, but not in the validation cohort, MTX dose was significantly higher in non-responders compared with responders at 6 months (12.3 mg/m²/week vs 9.3 mg/m²/week) and at 12 months (10.9 mg/m²/week vs 7.1 mg/m²/week) after MTX start.

In the derivation cohort, after 1 year of treatment, 149 patients (81.4%) were still on MTX, and 27 patients (14.8%) had stopped MTX due to insufficient effect ($n=5$), disease remission ($n=18$), gastrointestinal intolerance ($n=3$) or hepatotoxicity ($n=1$). In the validation cohort, 99 (92.5%) patients were still receiving MTX after 1 year and eight patients (7.7%) had stopped MTX due to insufficient effect ($n=3$), disease remission ($n=2$) and gastrointestinal intolerance ($n=3$).

During the first year of treatment, 143 patients (78.1%) in the derivation and 68 patients (65.4%) in the validation cohort were ACR70 non-responders (table 2), while 114 (62.3%) patients in the derivation and 52 (50%) patients in validation cohort were ACR50 non-responders (data not shown). These frequencies corresponded to the frequencies at 6 months after MTX start, which is a commonly used time point to establish MTX efficacy. The ACR70 non-responder frequencies in the validation cohort were similar to those found earlier;³ however, ACR70 frequencies in the derivation cohort were higher, possibly due to significantly longer disease duration before MTX start²³ in this cohort.

Prediction model for MTX non-responders according to ACR70

The following variables, univariately associated ($p\leq 0.20$) with MTX non-response, were included in the multivariate logistic regression: disease duration, limited joints, ESR, *MTRR* rs1801394, *MDR-1/ABCB1* rs1045642, *MRP-1/ABCC1* rs35592 and *PCFT* rs2239907 (table 1). Variables of the final prediction model consisted of ESR and *MTRR* rs1801394, *MDR-1/ABCB1* rs1045642, *MRP-1/ABCC1* rs35592 and *PCFT* rs2239907 (table 2). The AUC of the prediction model was 0.72 (95% CI: 0.63 to 0.81), indicating that it classified 72% of patients correctly (table 3). The Hosmer-Lemeshow goodness-of-fit test was not statistically significant ($p=0.91$), indicating that the model fit the data well.

These predictors were used to test the model in a validation cohort. The AUC of the validation cohort was 0.65 (95% CI: 0.54 to 0.77), indicating that 65% of patients were classified correctly (table 3).

To enable healthcare professionals to easily use the model, the regression coefficients β of the model's predictors, transformed into simple scores, were used to compute an individual risk score for being an MTX non-responder (table 3). This score ranged from 0 to 11 points with a higher score reflecting a higher probability of non-response. The risk score of a patient that has all predictors of the final model is calculated by adding up the constant to the simple scores, assigned to individual predictors: 11 (the constant)+(-2)+(-3)+(-2)+(-2)+(-2), which results in a risk score of 0. If all predictors are present, the probability of non-response is 0.42. On the other hand, the risk score of a patient having no predictors would be equal to the constant of 11. If no predictors are present, the probability of non-response is 0.98. Within the 0-11 range, the diagnostic accuracy of different cut-offs for predicting the risk of being an MTX non-responder was evaluated by computing the corresponding sensitivity, specificity, positive predictive value and negative predictive value (table 4).

Our goal was to correctly identify as many future MTX non-responders as possible (high sensitivity), while attempting to avoid misidentification of MTX responders as MTX non-responders as much as possible (reasonable specificity). In the derivation cohort, this was reached at the cut-off ≥ 3 , where 98 of 125 (78%) MTX non-responders and 19 of 39 (49%) MTX responders were identified correctly; 27 non-responders were classified as responders (false negatives) and 20 responders were classified as non-responders (false positives) (table 4). Similarly, in the validation cohort, at the cut-off ≥ 3 , 48 (79%) of 61 MTX non-responders were identified correctly, whereas nine (26%) of 34 MTX responders were identified correctly; 13 non-responders were classified as responders (false negatives) and 25 responders were classified as non-responders (false positives).

A

DISCUSSION

We developed and validated a prediction model for clinical non-response in two large JIA cohorts consisting of ESR and four SNPs in the *MTRR*, *MDR-1/ABCB1*, *MRP-1/ABCC1* and *PCFT* genes. The model classified 72% of patients correctly in the derivation and 65% in the validation cohort.

To our knowledge, no previous studies have constructed a model to predict MTX non-response in JIA. Several studies did report associations of MTX non-response in JIA with polyarticular disease, longer disease duration, ANA negativity and a higher level of disability.^{24,25} In our study, longer disease duration and ANA negativity were univariately associated with MTX non-response, although not significantly. Moreover, extended oligoarticular JIA subtype was associated with MTX response.²⁵

Table 4. Diagnostic parameters for various risk score cut-offs predicting ACR70 MTX non-response

Cohort	Cut-off	Probability of MTX				
		non-response	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Derivation	≥1	0.42	97	13	78	55
	≥3	0.60	78	49	83	41
	≥5	0.77	50	82	90	34
	≥6	0.88	44	87	92	33
	≥7	0.90	24	95	94	28
	≥8	0.93	16	97	95	27
Validation	≥1	0.62	97	9	66	60
	≥3	0.77	79	26	66	41
	≥5	0.88	46	68	72	41
	≥6	0.91	40	71	71	40
	≥7	0.91	21	88	76	38
	≥8	0.91	11	91	70	36

Risk scores were calculated in 164 patients in the derivation cohort (n=19 baseline erythrocyte sedimentation rate (ESR) values were missing) and in 95 patients in the validation cohort (n=8 ESR values were missing and n=1 failed genotyping of *PCFT* gene).

ACR70, American College of Rheumatology paediatric 70; MTX, methotrexate; NPV, negative predictive value; PPV, positive predictive value.

However, we and others^{26;27} observed equal MTX response rates among different JIA subtypes. Therefore, in the present study JIA subtype was not a predictor of MTX non-response. Furthermore, no effect modification was detected upon restricting the analysis to the more prevalent oligoarthritis and polyarthritis subtypes since the prediction model preserved its predictive power (AUC: 0.72, 95% CI: 0.61 to 0.82).

An MTX efficacy prediction model was constructed in RA, classifying 85% of patients correctly.¹⁷ This model contained four clinical variables and four SNPs encoding *AMPD1*, *ATIC*, *ITPA* and *MTHFD1* (methylentetrahydrofolate dehydrogenase) enzymes. Despite differences in definitions of response and in demographics of RA and JIA patients, inclusion of SNPs was essential for adequate prediction of MTX non-response in both models. Our prediction model with ESR only yielded a poor AUC of 0.59 (95% CI: 0.49 to 0.69), whereas the addition of SNPs raised the AUC to 0.72. Therefore, SNPs were crucial for a good prediction of MTX non-responders in JIA.

The goal of our model is to correctly identify future non-responders who can be given early additional treatment with biologicals, and simultaneously to keep misidentification of future responders as non-responders to a minimum. This goal stems from the following important changes in treatment mentality of paediatric rheumatologists over the past years, prompted by the need to establish early disease control to prevent irreversible joint damage. First, paediatric rheumatologists no longer consider MTX response according to ACR30 or ACR50 sufficient, but judge it to be good only if patients satisfy the more stringent ACR70 criteria.²⁸⁻³⁰ Furthermore, they consider patients MTX responders if they

satisfy these criteria already within 3 months after MTX start. These changes in treatment mentality have resulted in a lower threshold to start early combination treatment with biologicals. Although very effective, biologicals potentially carry a heightened risk of malignancies and inflammatory bowel disease.¹²⁻¹⁶ To address these risks, while considering it crucial to adequately treat MTX non-responders as early as possible with biologicals and at the same time restrict their use to those patients who really need them, we selected a cut-off ≥ 3 as the optimal score. Using this cut-off in the derivation cohort would allow 98 (78%) of 125 non-responders to receive early additional treatment with biologicals, and spare 19 (49%) of 39 patients, identified as responders, from receiving them. In the validation cohort, 79% of non-responders would be given timely biological treatment, whereas 26% of patients identified as responders would be spared from receiving them (table 4). Although the sensitivity at this cut-off was the same for both cohorts, the specificity was considerably lower in the validation cohort (49% vs 26%), which is due to its relatively small size.

The choice of a cut-off, however, depends on the clinical goal. A cut-off ≥ 6 could be chosen, if clinicians use the prediction model primarily to select as many responders as possible, while avoiding misidentification of non-responders as responders. At this cut-off, 34 of 39 (87%) MTX responders were identified correctly, while 55 of 125 (44%) MTX non-responders were identified correctly. Similar diagnostic parameters were obtained in the validation cohort (Table 4).

Our model was constructed for ACR70 non-responders in at least two of three visits during the first year of treatment, due to known fluctuations in MTX (non-)response during the first year.²² Nevertheless, the model had an equally strong predictive power for ACR70 non-responders (AUC=0.71, 95% CI: 0.62 to 0.80) at 6 months after MTX start. Depending on the clinician's preference, the model could also be applied for a less stringent ACR50 non-response, since its predictive power was strong both in the first year of treatment (AUC=0.70, 95% CI: 0.61 to 0.77) and at 6 months after MTX start (AUC=0.72, 95% CI: 0.63 to 0.80).

Further studies are needed to evaluate the effect of these SNPs on enzyme activity and transporter function. As we and others have shown, the non-synonymous rs1045642 SNP in the *MDR-1/ABCB1* efflux transporter gene was associated with a higher probability of good clinical response to MTX.³¹ The synonymous rs35592 SNP in another *MRP-1/ABCC1* efflux transporter gene has been associated with higher risk of MTX non-response in psoriasis patients,³² whereas here this SNP was associated with a lower risk of non-response. The synonymous *PCFT* rs2239907 SNP, whose protein is an influx transporter, has not been described earlier in relation to MTX efficacy in arthritis. Finally, the non-synonymous *MTRR* rs1801394 SNP was associated with decreased MTX sensitivity in acute lymphoblastic leukaemia,³³ whereas in our JIA cohorts it conferred a decreased risk of MTX non-response.

A limitation of the model is its moderate predictive power of 65% in the relatively small validation cohort. This can impede its direct clinical use, indicating the need for

further refinement. Therefore, to confirm the model's clinical applicability, validation will be performed in a large international cohort prior to its implementation in daily clinical practice. Pharmacogenetic testing may also challenge the model's application in daily clinical practice. Nevertheless, we show that SNPs are indispensable to adequately predict MTX non-responders in our JIA cohorts. Furthermore, such testing is becoming routinely available and less expensive.

Our model predicted and validated MTX non-response in two JIA cohorts by combining clinical and genetic variables. The model offers the promise of personalised treatment in JIA where patients unresponsive to MTX monotherapy will promptly receive additional treatment with biologicals and those destined to be MTX responders will not. Therefore, we will implement the model in daily clinical practice to establish whether its use will result in reduction of disease activity and better disease control in JIA patients.

FUNDING

Funding provided by the Dutch Arthritis Association and MEDAC GmbH, Germany.

REFERENCE LIST

1. Ravelli A, Martini A. Juvenile idiopathic arthritis. *Lancet* 2007; 369(9563):767-778.
2. Foell D, Wulffraat N, Wedderburn LR, Wittkowski H, Frosch M, Gerss J et al. Methotrexate withdrawal at 6 vs 12 months in juvenile idiopathic arthritis in remission: a randomized clinical trial. *JAMA* 2010; 303(13):1266-1273.
3. Ruperto N, Murray KJ, Gerloni V, Wulffraat N, de Oliveira SK, Falcini F et al. A randomized trial of parenteral methotrexate comparing an intermediate dose with a higher dose in children with juvenile idiopathic arthritis who failed to respond to standard doses of methotrexate. *Arthritis Rheum* 2004; 50(7):2191-2201.
4. Hashkes PJ, Laxer RM. Medical treatment of juvenile idiopathic arthritis. *JAMA* 2005; 294(13):1671-1684.
5. Lovell DJ, Giannini EH, Reiff A, Cawkwell GD, Silverman ED, Nocton JJ et al. Etanercept in children with polyarticular juvenile rheumatoid arthritis. Pediatric Rheumatology Collaborative Study Group. *N Engl J Med* 2000; 342(11):763-769.
6. Ruperto N, Lovell DJ, Cuttica R, Wilkinson N, Woo P, Espada G et al. A randomized, placebo-controlled trial of infliximab plus methotrexate for the treatment of polyarticular-course juvenile rheumatoid arthritis. *Arthritis Rheum* 2007; 56(9):3096-3106.
7. Swart JF, Barug D, Mohlmann M, Wulffraat NM. The efficacy and safety of interleukin-1-receptor antagonist anakinra in the treatment of systemic juvenile idiopathic arthritis. *Expert Opin Biol Ther* 2010; 10(12):1743-1752.
8. Prince FH, Twilt M, Ten CR, Van Rossum MA, Armbrust W, Hoppenreijns EP et al. Long-term follow-up on effectiveness and safety of etanercept in juvenile idiopathic arthritis: the Dutch national register. *Ann Rheum Dis* 2009; 68(5):635-641.
9. Beresford MW, Baildam EM. New advances in the management of juvenile idiopathic arthritis--2: the era of biologicals. *Arch Dis Child Educ Pract Ed* 2009; 94(5):151-156.
10. Mircic M, Kavanaugh A. Inhibition of IL6 in rheumatoid arthritis and juvenile idiopathic arthritis. *Exp Cell Res* 2011; 317(9):1286-1292.
11. Tynjala P, Vahasalo P, Tarkiainen M, Kroger L, Aalto K, Malin M et al. Aggressive combination drug therapy in very early polyarticular juvenile idiopathic arthritis (ACUTE-JIA): a multicentre randomised open-label clinical trial. *Ann Rheum Dis* 2011; 70(9):1605-1612.
12. Hashkes PJ, Uziel Y, Laxer RM. The safety profile of biologic therapies for juvenile idiopathic arthritis. *Nat Rev Rheumatol* 2010; 6(10):561-571.
13. Dallochio A, Canioni D, Ruemmele F, Duquesne A, Scoazec JY, Bouvier R et al. Occurrence of inflammatory bowel disease during treatment of juvenile idiopathic arthritis with etanercept: a French retrospective study. *Rheumatology (Oxford)* 2010; 49(9):1694-1698.
14. Diak P, Siegel J, La GL, Choi L, Lemery S, McMahon A. Tumor necrosis factor alpha blockers and malignancy in children: forty-eight cases reported to the Food and Drug Administration. *Arthritis Rheum* 2010; 62(8):2517-2524.
15. Horneff G, De BF, Foeldvari I, Girschick HJ, Michels H, Moebius D et al. Safety and efficacy of combination of etanercept and methotrexate compared to treatment with etanercept only in patients with juvenile idiopathic arthritis (JIA): preliminary data from the German JIA Registry. *Ann Rheum Dis* 2009; 68(4):519-525.
16. Ruperto N, Martini A. Pediatric rheumatology: JIA, treatment and possible risk of malignancies. *Nat Rev Rheumatol* 2011; 7(1):6-7.
17. Wessels JA, van der Kooij SM, le CS, Kievit W, Barerra P, Allaart CF et al. A clinical pharmacogenetic model to predict the efficacy of methotrexate monotherapy in recent-onset rheumatoid arthritis. *Arthritis Rheum* 2007; 56(6):1765-1775.
18. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J et al. International

- League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004; 31(2):390-392.
19. Ruperto N, Ravelli A, Falcini F, Lepore L, De SR, Zulian F et al. Performance of the preliminary definition of improvement in juvenile chronic arthritis patients treated with methotrexate. Italian Pediatric Rheumatology Study Group. *Ann Rheum Dis* 1998; 57(1):38-41.
 20. Giannini EH, Ruperto N, Ravelli A, Lovell DJ, Felson DT, Martini A. Preliminary definition of improvement in juvenile arthritis. *Arthritis Rheum* 1997; 40(7):1202-1209.
 21. Wulffraat N, van der Net JJ, Ruperto N, Kamphuis S, Prakken BJ, Ten CR et al. The Dutch version of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ). *Clin Exp Rheumatol* 2001; 19(4 Suppl 23):S111-S115.
 22. de Rotte MC, Luime JJ, Bulatovic M, Hazes JM, Wulffraat NM, de JR. Do snapshot statistics fool us in MTX pharmacogenetic studies in arthritis research? *Rheumatology (Oxford)* 2010; 49(6):1200-1201.
 23. Albers HM, Wessels JA, van der Straaten RJ, Brinkman DM, Suijlekom-Smit LW, Kamphuis SS et al. Time to treatment as an important factor for the response to methotrexate in juvenile idiopathic arthritis. *Arthritis Rheum* 2009; 61(1):46-51.
 24. Vilca I, Munitis PG, Pistorio A, Ravelli A, Buoncompagni A, Bica B et al. Predictors of poor response to methotrexate in polyarticular-course juvenile idiopathic arthritis: analysis of the PRINTO methotrexate trial. *Ann Rheum Dis* 2010; 69(8):1479-1483.
 25. Ravelli A, Viola S, Migliavacca D, Ruperto N, Pistorio A, Martini A. The extended oligoarticular subtype is the best predictor of methotrexate efficacy in juvenile idiopathic arthritis. *J Pediatr* 1999; 135(3):316-320.
 26. Woo P, Southwood TR, Prieur AM, Dore CJ, Grainger J, David J et al. Randomized, placebo-controlled, crossover trial of low-dose oral methotrexate in children with extended oligoarticular or systemic arthritis. *Arthritis Rheum* 2000; 43(8):1849-1857.
 27. Giannini EH, Brewer EJ, Kuzmina N, Shaikov A, Maximov A, Vorontsov I et al. Methotrexate in resistant juvenile rheumatoid arthritis. Results of the U.S.A.-U.S.S.R. double-blind, placebo-controlled trial. The Pediatric Rheumatology Collaborative Study Group and The Cooperative Children's Study Group. *N Engl J Med* 1992; 326(16):1043-1049.
 28. Ravelli A, Ramenghi B, Di Fuccia G, Ruperto N, Zonta L, Martini A. Factors associated with response to methotrexate in systemic-onset juvenile chronic arthritis. *Acta Paediatr* 1994; 83(4):428-432.
 29. Bartoli M, Taro M, Magni-Manzoni S, Pistorio A, Traverso F, Viola S et al. The magnitude of early response to methotrexate therapy predicts long-term outcome of patients with juvenile idiopathic arthritis. *Ann Rheum Dis* 2008; 67(3):370-374.
 30. Albers HM, Brinkman DMC, Kamphuis SSM, van Suijlekom-Smit LWA, van Rossum MAJ, Hoppenreijns EPAH et al. Clinical course and prognostic value of disease activity in the first two years in different subtypes of juvenile idiopathic arthritis. *Arthritis Care Res* 2010; 62(2):204-212.
 31. Pawlik A, Wrzesniewska J, Fiedorowicz-Fabrycy I, Gawronska-Szklarz B. The MDR1 3435 polymorphism in patients with rheumatoid arthritis. *Int J Clin Pharmacol Ther* 2004; 42(9):496-503.
 32. Warren RB, Smith RL, Campalani E, Eyre S, Smith CH, Barker JN et al. Genetic variation in efflux transporters influences outcome to methotrexate therapy in patients with psoriasis. *J Invest Dermatol* 2008; 128(8):1925-1929.
 33. de Jonge R, Hooijberg JH, van Zelst BD, Jansen G, van Zantwijk CH, Kaspers GJ et al. Effect of polymorphisms in folate-related genes on in vitro methotrexate sensitivity in pediatric acute lymphoblastic leukemia. *Blood* 2005; 106(2):717-720.