

A randomized trial of methotrexate vs. azathioprine for severe atopic eczema: a critical appraisal

DOI: 10.1111/j.1365-2133.2012.10872.x

ORIGINAL ARTICLE: Schram ME, Roekevisch E, Leeflang MMG, Bos JD, Schmitt J, Spuls PI. A randomized trial of methotrexate versus azathioprine for severe atopic eczema. *J Allergy Clin Immunol* 2011; 128: 353–9

Summary

Aim Schram et al. aimed to compare the efficacy and safety of methotrexate vs. azathioprine in adults with severe atopic eczema.

Setting and design This single-blind, parallel-group (ratio 1 : 1), randomized controlled trial was conducted in a secondary care setting in the Netherlands between July 2009 and December 2010.

Study exposure Patients with severe atopic eczema were randomly assigned in a 1 : 1 ratio to receive either methotrexate (10–22.5 mg weekly) or azathioprine (1.5–2.5 mg kg⁻¹ daily) for 12 weeks, followed by a 12-week follow-up period.

Outcomes The outcome measures comprised various eczema severity measures including: SCORing of Atopic Dermatitis index (SCORAD); Investigator Global Assessment (IGA); Patient Global Assessment (PGA); Eczema Area and Severity Index (EASI); and Patient-Oriented Eczema Measurement (POEM). Further outcomes included a visual analogue scale of itch and sleeplessness; Skindex-17; serum levels of thymus and activation-regulated chemokine; quantity of topical corticosteroids used; and the number of courses of rescue medication (oral prednisolone) used.

Primary outcome measure The primary outcome was the mean change in SCORAD after 12 weeks of treatment.

Results Forty-five patients were screened and 42 of these were included in the trial. At week 12, patients in the methotrexate group had a mean $\pm$ SD relative reduction in SCORAD of $42 \pm 18\%$ compared with $39 \pm 25\%$ in the azathioprine group ($P = 0.52$). Proportions of patients achieving at least mild disease and reductions in impact on quality of life and symptoms were similar in both groups at weeks 12 and 24. No statistically significant differences were found in the number and severity of adverse events. Abnormalities in blood count were more common in the azathioprine group but no serious adverse events occurred in either group.

Conclusions Schram et al. conclude that both methotrexate and azathioprine achieved clinically relevant improvement and were safe in the short term. They also conclude that

both treatments are appropriate treatment options for severe atopic eczema in adult populations.

Comment

What is already known about this topic?

Eczema is a common chronic inflammatory skin disorder that affects both child and adult populations. There is a need for novel, medium to long term treatment options for patients with severe eczema.¹ Although effective, ciclosporin and prednisolone are best used as short-term treatments, the former being nephrotoxic and the latter predisposing to osteoporosis, hypertension and other side-effects.² As healthcare costs increase, dermatologists are keen to know which treatments achieve disease control safely and with minimum cost.¹ Methotrexate and azathioprine are both low-cost steroid-sparing systemic agents that are used to treat eczema, but both are unlicensed for this indication. There is a paucity of evidence for the use of methotrexate in eczema. No randomized controlled trial (RCT) of methotrexate in eczema exists and the only published evidence for its efficacy consists of case series and open-label studies.^{3–6} There is more evidence to support the use of azathioprine to treat eczema, although the only two published RCTs of azathioprine in eczema were placebo controlled.^{7,8} Uncontrolled studies on azathioprine also exist, showing similar results. An RCT comparing methotrexate and azathioprine in eczema has not been previously performed. Schram et al.¹ are to be congratulated for undertaking the first study comparing azathioprine and methotrexate in eczema.

Strengths of research

This study is the first head-to-head comparison of methotrexate against azathioprine in patients with severe atopic eczema. This study was not industry sponsored and the investigators had no relevant conflict of interest, making biased interpretation of results less likely. The study is well reported, closely following the CONSORT guidelines for reporting of RCTs.⁹ Markers of good reporting in this study include: clearly stated primary and secondary outcome measures; the use of a table showing the baseline characteristics of participants in each group and a flow diagram to show the fate of participants who entered the study; full reporting of side-effects; registration of the study on the Dutch Trials Register; and a thorough

consideration of the potential limitations and generalizability of the study.

In broad terms, the internal validity of an RCT can be assessed with respect to three main areas: randomization and its concealment, blinding and intention-to-treat analysis. These factors are addressed to a good level in this RCT, with the exception of the lack of blinding of the patients to treatment allocation (see details under 'Assessment of validity').

All participants were deemed to have severe atopic eczema as per Rajka and Langeland criteria, a validated severity measure.¹⁰ Numerous subjective and objective disease measures, many of which have been validated, were used to assess severity of disease and to measure disease improvement.

Assessment of validity

When evaluating RCTs, the first consideration is *internal validity*: did the authors take appropriate measures to minimize bias? In making treatment decisions, the second consideration is *external validity*: can the results of the RCT be reasonably applied to a particular group of patients in a particular clinical setting?¹¹

Internal validity

Randomization

Adequate randomization protects against selection bias. Selection bias is an error in creating intervention groups, such that they differ with respect to prognosis due to a variation in baseline characteristics resulting from the way in which they are selected. This will have an impact on internal validity.¹² Participants in the study of Schram *et al.*¹ were randomized in a 1 : 1 ratio using a computer program, which also allowed concealment of allocation from investigators. Treatment groups were well balanced at baseline, suggesting a reasonable randomization process.

Blinding

In this study, trained efficacy assessors were blinded for allocation. Statistical analysis was performed by the third author, who was also blinded for allocation. Patients and safety assessors were not blinded to allocation. Blinding participants would have been very difficult, as there is an obvious difference in the frequency of administration between azathioprine and methotrexate, as well as a different side-effect profile (methotrexate tending to cause more gastrointestinal upset). Furthermore, folic acid supplementation is only given in conjunction with methotrexate therapy. Nonblinding of participants can lead to ascertainment bias. Ascertainment bias is distortion of the results of a trial as a result of knowledge of the group assignment by the person assessing outcome, whether an investigator or the participants themselves.¹³

The only measure that could be adopted to blind the participants to allocation would be to use a double dummy system, such that participants in the methotrexate arm receive a placebo tablet daily in addition to the weekly methotrexate tablet, and vice versa in order to eliminate the differences in administration. Despite the lack of blinding, the participants were naïve to treatment with methotrexate or azathioprine. This should make it less likely for participants to know, based on any side-effects, which drug they were on. This will have helped to guard against any bias that may otherwise have occurred in the participants' evaluations of treatment efficacy.

Intention to treat

Data in this study were analysed on an intention-to-treat basis. In an intention-to-treat analysis, all patients are followed until death or the end of the trial, irrespective of whether they are still taking the treatment to which they were randomized. The most recent data available for any participants who drop out are used in this analysis. Performing an intention-to-treat analysis helps to guard against attrition bias, which may be caused by dropouts. For example, if a treatment is so effective that it completely clears a participant's skin disease, they may choose to leave the trial before its completion. Conversely, if the treatment does not work at all, participants may not wish to continue the study and may drop out early. Ignoring dropouts could make a treatment appear more or less effective than it truly is.

Power

Power calculations are conducted to determine the number of participants that will be required in a trial, adequately to test an intervention or hypothesis. Power is a measure of a study's ability to detect whether an effect actually occurred or whether it is simply due to an inaccurate observation or result of experimental error, sometimes referred to as an 'artefact'.^{14–17} It is often difficult for researchers to obtain large samples. When dealing with limited samples, the right amount of power is critical to ensure an adequate sample size is achieved to detect differences. The larger the sample size, the more power is associated with the study.¹⁸

A small sample size in each treatment arm, as in the study of Schram *et al.*, can lead to under powering, which will affect the internal validity of the study. The authors clearly acknowledge the issues regarding the power of the study. Power analysis showed that 42 participants were necessary for a study with 80% power and 5% significance.

The power analysis was based on the SCORAD score. The minimum clinically important difference of the SCORAD score is unknown, so the authors decided that a difference of at least 8 points would represent a clinical meaningful difference. This assumption was based on other studies^{19,20} and on personal experience with the scale. However, one of the references used by the authors to support this assumption does not specifically state that an 8-point difference would be clinically relevant.¹⁹

The authors do acknowledge that it is possible that a smaller SCORAD score difference could be clinically meaningful. In this case, the sample size would not have been adequate to detect a clinically meaningful difference between therapies.

External validity

Patients were recruited from an inpatient and outpatient setting at the Academic Medical Centre of Amsterdam, which is a referral centre for severe atopic eczema. Patients were also referred by regional dermatologists. These patients were unresponsive, contraindicated or intolerant to ciclosporin treatment. The mean age of participants was 40 years, with a mean duration of disease of 36 years, representing a population with very severe, chronic disease. Although the findings cannot readily be applied to the general population, most of the patients who are treated with these agents are likely to have a similar degree of disease severity and chronicity, so the external validity is likely to be good for similar healthcare settings. Obviously no paediatric participants were included and so no firm conclusions can be drawn about the efficacy of these agents in children.

The primary efficacy outcome in this study was measured at 12 weeks. In the clinical setting it is possible that clinicians would persevere with the initial treatment for a longer period of time, given that both treatments can take 3 or more months to reach maximum efficacy, so it could be argued that the primary efficacy outcome should have been measured later. The authors acknowledge that there are no international guidelines for the use of these treatments, so some arbitrary decisions had to be made regarding treatment regimens. It is also acknowledged that undertaking an RCT for longer periods of time can be difficult, and sometimes it can be necessary to use earlier time points to improve completeness of follow-up.

Other issues

The wash-out period for systemic treatment in this study was relatively short (2 weeks), so patients who had previously been on a systemic agent may have had a lower baseline SCORAD score, leading to a smaller difference between baseline and 12-week efficacy outcomes and underestimation of treatment efficacy.

At baseline, patients were allowed to start or continue to use topical corticosteroids. Also, six of the 42 participants received rescue medication in the form of oral corticosteroids. Both of these may have had some impact on the change in SCORAD score. Given the small size and potentially low power of the study, the effect of these concomitant medications may have been considerable.

One patient switched treatment from azathioprine to methotrexate after 12 weeks. Although the switching of a single patient is unlikely to have affected the overall outcome, the impact is again likely to be greater due to the small sample size.

Overall assessment (including any potential application of the research to clinical practice)

This study suggests that methotrexate and azathioprine may be equally effective in treating severe atopic eczema. This was an important, well-designed and well-reported study with a reasonable treatment period of 12 weeks followed by an adequate follow-up period of 12 weeks. However, the sample size was small and there are concerns regarding the power to detect clinically meaningful differences in efficacy between the two treatments. This study does not permit conclusions in paediatric populations, nor does it provide evidence regarding long-term efficacy or adverse effects. Thus, further research is required in children and with larger numbers of participants and longer duration of follow-up. Comparative studies with other therapies such as ciclosporin and oral corticosteroids would also be useful.

Department of Dermatology,
Queen's Medical Centre,
Nottingham NG7 2UH, U.K.

A.N. PATEL
S.M. LANGAN*
J.M. BATCHELOR

*Department of Non-Communicable
Disease Epidemiology, London School of
Hygiene and Tropical Medicine and
St John's Institute of Dermatology, London, U.K.
E-mail: dranandpatel@doctors.net.uk

Funding sources

None.

Conflicts of interest

None declared.

References

- Schram ME, Roekevisch E, Leeflang MMG et al. A randomized trial of methotrexate versus azathioprine for severe atopic eczema. *J Allergy Clin Immunol* 2011; **128**:353–9.
- Chakravarty K, McDonald H, Pullar T et al. BSR/BHPR guideline for disease-modifying anti-rheumatic drug (DMARD) therapy in consultation with the British Association of Dermatologists. *Rheumatology (Oxford)* 2008; **47**:924–5.
- Weatherhead SC, Wahie S, Reynolds NJ, Meggit SJ. An open-label, dose-ranging study of methotrexate for moderate-to-severe adult atopic eczema. *Br J Dermatol* 2007; **156**:346–51.
- Zoller L, Ramon M, Bergman R. Low dose methotrexate therapy is effective in late-onset atopic dermatitis and idiopathic eczema. *Isr Med Assoc J* 2008; **10**:413–14.
- Goujon C, Berard F, Dahel K et al. Methotrexate for the treatment of adult atopic dermatitis. *Eur J Dermatol* 2006; **16**:155–8.
- Shaffrali FC, Colver GB, Messenger AG, Gawkrödger DJ. Experience with low dose methotrexate for the treatment of eczema in the elderly. *J Am Acad Dermatol* 2003; **48**:417–19.
- Berth-Jones J, Takwale A, Tan E et al. Azathioprine in severe adult atopic dermatitis: a double-blind, placebo-controlled, crossover trial. *Br J Dermatol* 2002; **147**:324–30.

- 8 Meggitt SJ, Gray JC, Reynolds NJ. Azathioprine dosed by thiopurine methyltransferase activity for moderate-to-severe atopic eczema: a double-blind, randomised controlled trial. *Lancet* 2006; **367**:839–46.
- 9 Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; **340**:c332.
- 10 Rajka G, Langeland T. Grading of the severity of atopic dermatitis. *Acta Derm Venereol* (Stockh) 1989; **144**:13–14.
- 11 Rothwell PM. External validity of randomised controlled trials: 'to whom do the results of this trial apply?' *Lancet* 2005; **365**:82–93.
- 12 CONSORT. Selection Bias. Available at: <http://www.consort-statement.org/resources/glossary/q---z/selection-bias/> (last accessed 15 January 2012).
- 13 CONSORT. Ascertainment Bias. Available at: <http://www.consort-statement.org/resources/glossary/a---d/ascertainment-bias/> (last accessed 15 January 2012).
- 14 Hamilton L. *Data Analysis for Social Scientists: A First Course in Applied Statistics*. Belmont, CA: Duxbury Press, 1996.
- 15 Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates Inc., 1988.
- 16 Bland M. *An Introduction to Medical Statistics*, 3rd edn. Oxford: Oxford University Press, 2000.
- 17 Munro BH. *Statistical Methods for Health Care Research*, 5th edn. Philadelphia, PA: Lippincott Williams & Wilkins, 2005.
- 18 Siegel S, Castellan NJ. *Nonparametric Statistics for the Behavioral Sciences*. New York, NY: McGraw-Hill Book Company, 1988.
- 19 Paul C, Lahfa M, Bachelez H et al. A randomized controlled evaluator-blinded trial of intravenous immunoglobulin in adults with severe atopic dermatitis. *Br J Dermatol* 2002; **147**:518–22.
- 20 Schmitt J, Meurer M, Schwanebeck U et al. Treatment following an evidence-based algorithm versus individualised symptom-oriented treatment for atopic eczema. A randomised controlled trial. *Dermatology* 2008; **217**:299–308.

Response to a randomized trial of methotrexate vs. azathioprine for severe atopic eczema: a critical appraisal

DOI: 10.1111/j.1365-2133.2012.10874.x

Patel and colleagues¹ have written a critical appraisal of our study² 'A randomized trial of methotrexate versus azathioprine for severe atopic eczema'. The authors have made a very thorough appraisal of our study with many educational aspects and relevant comments concerning the internal and external validity. Furthermore, we thank the authors for drawing attention to our study in a dermatology-specific journal. This will benefit patients with severe atopic eczema.

Hereby we would like to take the opportunity to present our comments on some of the statements.

Power

In our study, we have performed our power analysis in the absence of a formally calculated minimal clinically important difference (MCID) and we decided that an MCID of 8 points

on the SCORAD was appropriate to use. It can be argued that the MCID could have been smaller. In that case our sample size would have been inadequate to detect a difference. However, we have just performed a clinimetric study assessing the MCIDs of different outcome parameters including the SCORAD.³ Here we found that the MCID of the SCORAD was 8.7 points. It is known that MCIDs may vary by population and context and thus a single MCID may be insufficient for all study applications. Nevertheless, it is an indication that our decision to use an MCID of 8 points was appropriate.

Blinding

We fully agree with the comments made about the blinding. Additionally, we stress that for this investigator-initiated, non-sponsored study it would have been very expensive to blind the patients. In randomized controlled trials, patients, physicians, the randomization, the outcome assessment (efficacy and safety) and the statistician performing the analyses could be blinded. In our study we met three of these five possibilities.

Wash-out

Wash-out was indeed relatively short and might have made the found treatment effect smaller. We were aware of that fact when we designed the study. Our consideration was the well-being of our patients. Atopic eczema has a fluctuating course and 4 weeks without potent topical corticosteroids and/or systemic therapy in severely affected patients was considered unethical.

Rescue medication and other issues

We agree with the statement that allowing patients to start or continue topical steroids and receive rescue medication could have had an impact on the change in SCORAD. Nevertheless, the total amount of topical steroids was low and equal in both groups and rescue medication was only allowed in the first 8 weeks.

Patients were allowed to stop or switch between study medications at week 12. This might have had an effect on our results at week 24. Our primary outcome measure was the SCORAD at week 12 and thus an effect on primary outcome is not expected.

Conclusion

Given the recently reported MCID in SCORAD³ the raised concern of limited statistical power of our study can be ruled out. We agree with the authors (as we stated in the discussion section of our initial article) that future studies should focus on children, long-term efficacy and safety and comparative studies.