

How should minimally important change scores for the Patient-Oriented Eczema Measure be interpreted?

A validation using varied methods*

L. Howells ¹, S. Ratib ¹, J.R. Chalmers ¹, L. Bradshaw² and K.S. Thomas ¹ on behalf of the CLOTHES trial team

¹Centre of Evidence Based Dermatology and ²Nottingham Clinical Trials Unit, School of Medicine, University of Nottingham, Nottingham, U.K.

Summary

Correspondence

Laura Howells.

E-mail: laura.howells1@nottingham.ac.uk

Accepted for publication

11 January 2018

Funding sources

This study was supported by the PhD studentship of L.H. awarded by the British Skin Foundation (ref: 8016). The CLOTHES trial was funded by the National Institute for Health Research Health Technology Assessment Programme (project number 11/65/01). The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Health Technology Assessment Programme, National Institute for Health Research (NIHR), National Health Service or the Department of Health. Initial development of the CLOTHES Trial was funded by the NIHR under its Programme Grants for Applied Research Programme (RP-PG-0407-10177).

Conflicts of interest

L.H., J.R.C. and K.S.T. are members of the HOME initiative. The Patient-Oriented Eczema Measure (POEM) was developed at the Centre of Evidence Based Dermatology, University of Nottingham.

*Plain language summary available online

DOI 10.1111/bjd.16367

Background The Patient-Oriented Eczema Measure (POEM), scored 0–28, is the core outcome instrument recommended for measuring patient-reported atopic eczema symptoms in clinical trials. To date, two published studies have broadly concurred that the minimally important change (MIC) of the POEM is three points. Further assessment of the MIC of POEM in different populations, and using a variety of methods, will improve interpretability of the POEM in research and clinical practice.

Objectives To calculate the smallest detectable change in the POEM and estimate the MIC of the POEM using a variety of methods in a trial dataset of children with moderate-to-severe atopic eczema.

Methods This study used distribution-based and anchor-based methods to calculate the MIC of the POEM in children with moderate-to-severe eczema.

Results Data were collected from 300 children. The smallest detectable change was 2.13. The MIC estimates were 1.07 (using 0.2 SD of baseline POEM scores) and 2.68 (using 0.5 SD of baseline POEM scores) based on distribution-based methods; were 3.09–6.13 based on patient-/parent-reported anchor-based methods; and were 3.23–5.38 based on investigator-reported anchor-based methods.

Conclusions We recommend the following thresholds be used to interpret changes in POEM scores: ≤ 2 , unlikely to be a change beyond measurement error; 2.1–2.9, a small change detected that is likely to be beyond measurement error but may not be clinically important; 3–3.9, probably a clinically important change; ≥ 4 , very likely to be a clinically important change.

What's already known about this topic?

- The Patient-Oriented Eczema Measure (POEM) is recommended as the core outcome instrument for measuring patient-reported symptoms in eczema clinical trials.
- Two previous studies have examined the minimally important change (MIC) of the POEM; one in children with mild eczema and another in adults with very severe eczema.
- These previous studies both concluded that the MIC in POEM is around three points.

What does this study add?

- This study explored the impact of different methodologies for calculating the MIC of the POEM in children with moderate-to-severe eczema.
- A change in POEM of less than two points is likely to be below the smallest detectable change (i.e. below measurement error) for the scale.
- The MIC varied considerably depending on the method used, but a change in POEM score below three points is unlikely to be a clinically important change.

What are the clinical implications of this work?

- This study aids sample size calculations for clinical trials and helps researchers, clinicians and patients to interpret changes in POEM scores in clinical trials and routine monitoring of eczema in clinical practice.

Minimally important change (MIC) has been defined as 'the smallest change in score in the construct to be measured which patients perceive as important'.¹ The Patient-Oriented Eczema Measure (POEM) measures patient-reported eczema severity, scored out of 28. Seven questions ask how frequently a symptom has occurred over the past week on a five-point scale from 'no days' to 'every day'.² The Harmonising Outcome Measures in Eczema (HOME) initiative, which is developing the core outcome set for atopic eczema clinical trials, has recommended all clinical trials measure patient-reported symptoms using the POEM.³ The National Institute for Health and Care Excellence guidance for clinical practice with children under 12 years old recommends use of the POEM.⁴ Understanding the MIC of the POEM can support sample size calculations for clinical trials and interpretation of trial results.⁵ MIC estimates also allow clinicians to interpret a patient's change in the POEM score in the clinical setting and aid decisions regarding whether a treatment alteration is required.⁵

There is a variety of terminology used to explain the concept of MIC, the most common alternatives being the minimal clinically important difference (MCID) and minimally important difference (MID). MIC tends to be used to refer to longitudinal within-person changes in scores, and MID for cross-sectional between-person differences.⁶ For consistency, we have used MIC throughout this study. There is an ongoing debate about whether the methods currently used are appropriate to estimate the MID.⁷

One of the major challenges for determining the MIC for a patient-reported outcome is that an MIC is not a fixed attribute, but is a variable concept that can be influenced by a number of factors including methods used to calculate it and baseline disease severity.^{5,8,9} There are broadly two main approaches that are used to calculate the MIC: anchor-based approaches and distribution-based approaches. In anchor-based approaches, a certain amount of change on an external criterion, which should ideally be a related and well interpretable outcome measure (the 'anchor'), is said to correspond to an MIC on the measuring instrument of interest.⁹ Distribution-based approaches are based on the distributional characteristics of the sample.⁹

The MIC of the POEM was previously estimated using a patient-reported anchor for the 'within-patient' score change method as 3.4 points in a study using datasets from two trials in adults with severe eczema.¹⁰ Subsequently, a study using

data from a trial in children with mild to moderate eczema aged between 1 month and 5 years from primary care used a combination of anchor-based and distribution-based methods to calculate the MIC. Results ranged from 2.5 to 4.27, which led the authors to conclude that the results broadly concurred with an MIC of three points.¹¹

It has been recommended that researchers use multiple methods and multiple datasets to triangulate MIC results.^{12–14} Although a multitude of MIC estimates could detract from the usefulness of a universal MIC threshold, it is important to explore how the MIC of the POEM may vary to ensure it is meaningful in the context used.

One concept that helps interpret change on a patient-reported outcome measure is the smallest detectable change. The smallest detectable change is defined as a change beyond measurement error.⁹ A change in POEM score that is greater than the smallest detectable change has 95% chance of being a true change in score rather than a random variation in the way people answer the POEM. The MIC of the POEM must be greater than the smallest detectable change to be useful. To date, the smallest detectable change of the POEM has not been assessed.¹⁵

Using the dataset from the CLOTHES trial, which was a 6-month randomized controlled trial to assess wearing silk clothing compared with usual care in children with moderate-to-severe eczema,¹⁶ the present study aimed to: (i) calculate the smallest detectable change in the POEM; (ii) estimate the MIC of the POEM by repeating methods used in previous studies and using additional methods; and (iii) assess whether using a patient/parent or investigator static global assessment as an anchor influences the anchor-based MIC estimates.

Patients and methods

The study protocol was registered on the Centre of Evidence Based Dermatology's Protocol Registration on 5 January 2017 prior to data analysis: <http://www.nottingham.ac.uk/research/groups/cebd/resources/protocol-registration.aspx>.

The CLOTHES trial

The CLOTHES trial is a parallel-group, randomized, controlled, observer-blind trial. Children aged 1–15 years were recruited from secondary care and the community and allocated to wear

either silk garments plus standard care or standard care only. At study entry, one of the eligibility criteria was that participants had either a moderate (9–11) or severe (12–15) score on the Nottingham Eczema Severity Score.^{16,17}

Measures

POEM measures patient-reported frequency of itch, sleep, bleeding, weeping/oozing, cracking, and flaking and dryness/roughness over the past week.² Each item is weighted equally and scored as 0 (no days), 1 (1–2 days), 2 (3–4 days), 3 (5–6 days) or 4 (every day).² This analysis used POEM scores at baseline and at 6 months from the CLOTHES dataset. A Patient's/Parent's Global Assessment (P/PGA) and Investigator's Global Assessment of Severity (IGA) were also used (these measures are described in Table 2).

Statistical analysis

The total POEM score was calculated by adding together the score from each item. If one item was missing the total score was still calculated, but total score was coded as missing if more than two items were missing.¹⁸ As only 9% of data collected at 6 months was missing, all analyses used complete case series. POEM scores from patients in both the treatment and control arm were not treated separately in this study. Except where stated otherwise, the statistical package Stata 14 was used to run analyses (StataCorp, College Station, TX, U.S.A.). No formal sample size was conducted, but the number of patients included in the analysis is likely to be sufficient. It has been recommended that validation studies contain at least 100 participants.⁹

Calculating the smallest detectable change

The smallest detectable change was calculated as $1.96 \times \sqrt{2} \times \text{SEM}$.⁹ The standard error of measurement (SEM) was calculated as $\text{SD}_{(\text{pooled})} \times \sqrt{1 - \text{ICC}}$.

$$\text{SD}_{(\text{pooled})} = \frac{\sqrt{\text{SD}(\text{baseline})^2 + \text{SD}(6\text{months})^2}}{2}$$

The intraclass correlation coefficient (ICC) for absolute agreement was derived from the test–retest reliability of the POEM, which was tested in the development of the POEM.² This was considered acceptable as there is similar variability in the CLOTHES data ($\text{SD}_{(\text{baseline})} = 5.36$) and the dataset used initially to validate the POEM by Charman *et al.* ($\text{SD}_{(\text{baseline})}$ is 7.73).²

Calculating the minimally important change (MIC)

Distribution-based methods Two distribution-based methods were used to estimate the MIC. An effect size is measured by the difference between the score at baseline and follow-up, divided by the standard deviation of the baseline score.¹⁹ 0.5

multiplied by the standard deviation of POEM scores at baseline was used as an estimate of the MIC.²⁰ It has been suggested that the MIC should correspond to a smaller effect size, therefore 0.2 multiplied by the standard deviation of POEM scores at baseline was also used.¹⁹

Anchor-based methods Four anchor-based methods were used to estimate the MIC. The IGA and P/PGA scores were transformed into a change score to provide an anchor: score at time point 1 (baseline) minus score at time point 2 (6 months). This creates change scores on the IGA and P/PGA that can range from –5 (worsened eczema severity) to 5 (improved eczema severity), that were used as anchors. The MIC was operationally defined as a positive one-point change on the P/PGA or the IGA anchor to indicate a minimal important improvement. This change indicates a change in severity banding on the P/PGA/IGA (e.g. from moderate to mild). This is also the cut-off used by Schram *et al.*, enabling comparisons with previous MIC estimates.¹⁰

For an anchor to be useful it must at least moderately correlate with the POEM change score ($r \geq 0.3$).¹³ This was assessed prior to the study with Pearson's *r* correlations between the POEM change scores and (i) the P/PGA change scores ($r = 0.55$) and (ii) the IGA change scores ($r = 0.46$), and both met this minimum criterion. The assumptions of Pearson's *r* correlations were met.

The first two anchor-based methods used mean change to analyse (i) 'within-patient' score change where the MIC is estimated as the mean change score of the smallest reported improvement (a positive one-point change in the anchor) and (ii) 'between-patient' score change where the MIC estimate is based on the relative change between the mean change score of the group with the smallest reported improvement on the anchor (a positive one-point change) and the mean change score of the group with no change on the anchor.

The third anchor-based method used the receiver operating characteristic (ROC) curve. The area under the curve of the ROC curve identifies the cut-off point on the POEM change scores that most optimally distinguishes between the anchor of IGA or P/PGA change scores ≤ 0 and IGA or P/PGA change scores ≥ 1 .¹⁹ The cut-off used to provide an MIC estimate will maximize the Youden *J* statistic: sensitivity – (1 – specificity).¹⁹ We used the statistical package R (R Foundation, Vienna, Austria) to use bootstrapping methods to allow us to calculate an MIC estimate with 95% confidence intervals (CIs).²¹

The fourth anchor-based method used predictive modelling. This method uses logistic regression to predict whether a patient belongs to the improved (≥ 1) or not improved group (≤ 0) on the IGA or P/PGA anchor using the change in POEM score as the predictor.²¹ The MIC is estimated using the equation $[\ln(\text{pre-odds}) - C]/B$, where *C* is the intercept and *B* is the regression coefficient for the change in POEM score from the logistic regression model. The pre-odds is calculated using the proportion improved based on the anchor divided by 1

minus the proportion improved based on the anchor.²¹ The Microsoft Excel spreadsheet designed to aid confidence interval calculations of predictive modelling MIC estimates provided in supplementary materials by Terluin *et al.* was used.²¹ It has been suggested that if the proportion improved on the anchor does not equal 0.5, an adjusted MIC may need to be calculated.²² As the proportion improved on the IGA anchor was 0.56 and the P/PGA anchor was 0.53, we have not reported the adjusted MIC here.

Results

A total of 300 children with atopic eczema were randomized into the CLOTHES trial and completed the POEM at baseline; 174 (58%) were female and the majority were of white ethnic origin ($n = 237$, 79%). At 6 months, 273 participants (91%) had an assessment visit, therefore 273 patients were included in the anchor-based MIC methods that required this time point. Table 1 summarizes the age, disease severity and POEM scores of the sample. Figure 1 provides the distribution of POEM scores at baseline.

Smallest detectable change

The smallest detectable change is a change beyond measurement error. The intraclass correlation coefficient was 0.9847 and the pooled SD was 6.21, therefore the smallest detectable change in the CLOTHES dataset was 2.13 points on POEM.

Minimally important change

The MIC of the POEM was analysed using two distribution-based methods and four anchor-based methods. Figure 2 summarizes the results of the MIC estimates for each method.

Distribution-based methods

As shown in Table 1, the SD of POEM scores at baseline was 5.36. Using 0.5 SD of baseline scores gave an MIC of 2.68 points (95% CI 2.5–2.89) and using 0.2 SD of baseline scores gave an MIC of 1.07 points (95% CI 1.00–1.16).

Table 1 Summary of participant demographics and Patient-Oriented Eczema Measure (POEM) scores

	n	Mean \pm SD	Minimum	Maximum
Age (years)	300	5.06 \pm 3.63	1	15
Nottingham Eczema Severity Score	300	13.13 \pm 1.62	9	15
POEM				
Baseline	300	16.95 \pm 5.36	4	28
6 months	273	12.16 \pm 6.95	0	28
Change scores	273	4.78 \pm 7.14	-21	24

Anchor-based methods

Table 2 shows how the IGA anchor and P/PGA anchor were created.

Mean change methods – the ‘within-patient’ score change and the ‘between-patient’ score change As presented in Table 3, when using the IGA as the anchor, the mean change score of the POEM was 2.02 for those with a change score of zero (no change on the IGA) and 5.25 for those with a score of one (defined as minimal improvement on the IGA). Therefore, for the ‘within-patient’ score change method the MIC is 5.25 (95% CI 4.04–6.46) and for the ‘between-patient’ score change method the MIC is 3.23 (95% CI 1.51–4.95).

As presented in Table 3, when using the P/PGA as the anchor, the mean change score of the POEM was 3.04 for those with a change score of zero (no change on the P/PGA) and 6.13 for those with a score of one (defined as minimal improvement on the P/PGA). Therefore, for the ‘within-patient’ score change method the MIC is 6.13 (95% CI 4.82–7.44) and for the ‘between-patient’ score change method the MIC is 3.09 (95% CI 1.25–4.94).

ROC curve method The MIC using the IGA anchor was 5.38 (95% CI 1.5–8.5) and the MIC using the P/PGA anchor was 4.75 (95% CI 3.5–6.5).

Predictive modelling method Table 4 contains the results of the logistic regression analyses that were used to calculate the MIC and confidence intervals. Using IGA as the anchor, the MIC estimate using predictive modelling was 4.43 (95% CI 2.21–6.74). Using the P/PGA anchor, the MIC estimate was 4.52 (95% CI 2.81–6.29).

Discussion

This study showed the smallest detectable change on the POEM is 2.13 points. Therefore, only MIC estimates above 2.13 can be considered a change beyond measurement error.

Using the CLOTHES trial dataset, the MIC estimates of this study ranged from 1.07 (using 0.2 SD of baseline POEM scores) to 6.13 (using an IGA as the anchor for the ‘within-patient’ score change method). The method used should be considered when interpreting published MIC values as it has clearly had an impact on estimates in this study. There is still debate over which method should be used, hence the pluralistic approach used in this study. There is a trade-off between convenient, hence widely used, distribution-based approaches and more theoretically appropriate anchor-based approaches.

Anchor-based approaches, unlike distribution-based approaches, include an explicit judgement of the importance of the change.²³ Nevertheless, it has been suggested MIC estimates using 0.5 SD corresponded well with anchor-based methods.²⁴ Within anchor-based approaches, the methods are evolving to become increasingly sophisticated, first with the development of the ROC curve method for use in this scenario

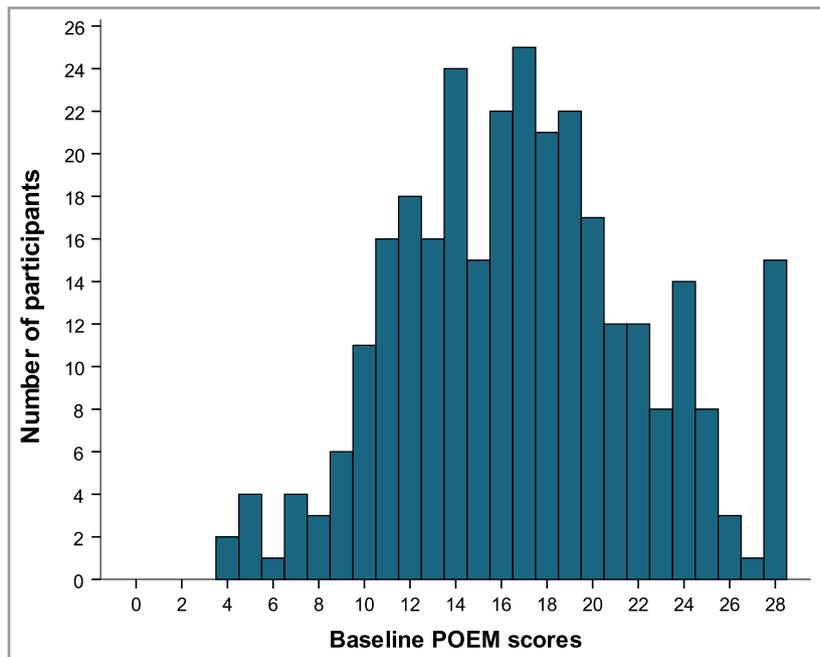


Fig 1. Distribution of baseline Patient-Oriented Eczema Measure (POEM) scores.

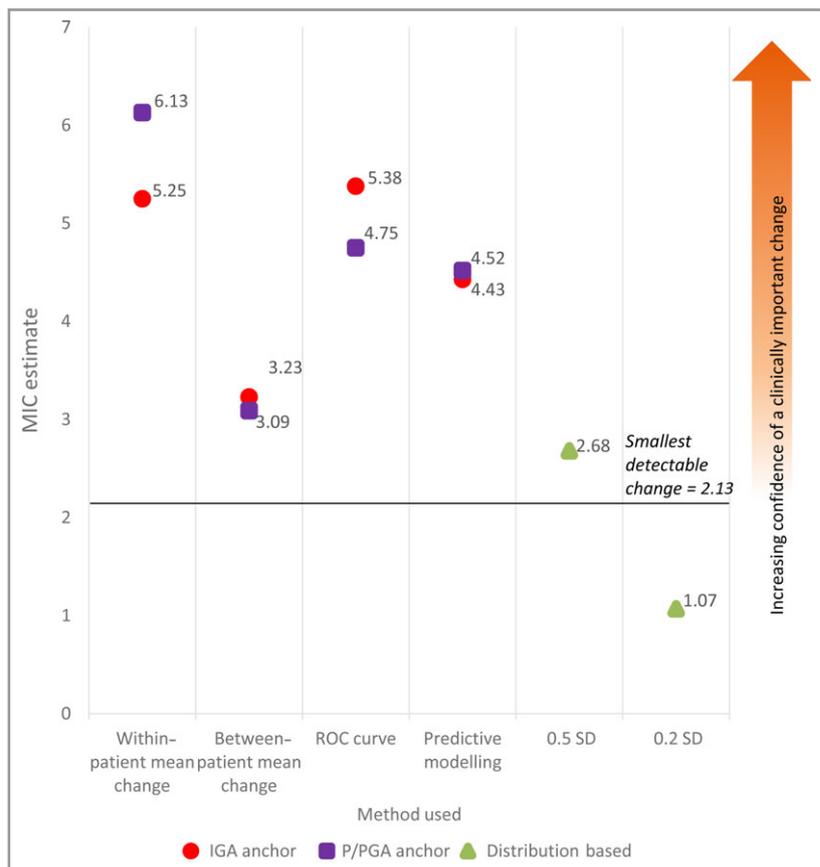


Fig 2. Summary of minimally important change (MIC) estimates by method used. IGA, Investigator’s Global Assessment; P/PPGA, Patient’s/Parent’s Global Assessment.

and more recently the development of the predictive modelling method. The two anchors, the IGA and the P/PGA, were in agreement for some methods, but not others. For example, the MIC estimates using the two anchors were very

close for the predictive modelling approach, but not for the ROC curve methods. We cannot provide firm conclusions as to why this is the case. It could be because of the ROC curve method being more sensitive to random sampling variation.²¹

Table 2 Measures used for anchors

Measure name	Question	Response options (tick one box)	Completed by:	Times collected to be used for anchors:
Static Patient's/Parent's Global Assessment (P/PGA)	How is your/your child's eczema today?	Clear Almost clear Mild Moderate Severe Very severe	Parent/legal guardian of child with eczema or child themselves if old enough (individual decision)	Baseline 6 months
Static Investigator's Global Assessment (IGA)	How is the child's eczema today?	Clear Almost clear Mild Moderate Severe Very severe	Research nurse (excluded measure when different nurse completed at different time points)	Baseline 6 months

Table 3 Mean Patient-Oriented Eczema Measure (POEM) change scores for participants classified according to change on anchors (n = 273)

Investigator's Global Assessment (IGA)			Patient's/Parent's Global Assessment (P/PGA)		
Change in score on IGA	n (%)	Mean POEM change score \pm SD	Change in score on P/PGA	n (%)	Mean POEM change score \pm SD
-5	0 (0)	n/a	-5	0 (0)	n/a
-4	0 (0)	n/a	-4	0 (0)	n/a
-3	0 (0)	n/a	-3	0 (0)	n/a
-2	1 (0.4)	-1 (n/a)	-2	9 (3.3)	-3 \pm 9.64
-1	24 (8.8)	1.54 \pm 6.77	-1	38 (13.9)	-0.79 \pm 4.82
0	95 (34.8)	2.02 \pm 6.01	0	81 (29.7)	3.04 \pm 5.88
1	111 (40.7)	5.25 \pm 6.43	1	92 (33.7)	6.13 \pm 6.33
2	33 (12.1)	10.67 \pm 6.01	2	38 (13.9)	9.05 \pm 5.92
3	9 (3.3)	15.89 \pm 7.17	3	10 (3.7)	11.8 \pm 4.57
	0 (0)	n/a	4	5 (1.8)	18.2 \pm 5.63
5	0 (0)	n/a	5	0 (0)	n/a

n/a, not applicable.

Table 4 Logistic regression results used for predictive modelling minimally important change estimates

	Investigator's Global Assessment (IGA)	Patient's/Parent's Global Assessment (P/PGA)
Pre-odds ^a	1.257	1.133
C \pm standard error	-0.293 \pm 0.159	-0.598 \pm 0.171
B \pm standard error	0.121 \pm 0.022	0.160 \pm 0.025
Correlation of C and B ^b	-0.570	-0.615

^aOdds of improvement based on anchor only; ^bthis was calculated using IBM SPSS Statistics 22 (IBM, Armonk, NY, U.S.A.) and was required as part of the calculation of 95% confidence intervals.

However, it could be because of the genuine differences in patient and investigator scoring of global severity.

The results of the distribution-based methods suggest that an effect size of 0.5 corresponds better to the anchor-based

methods than using an effect size of 0.2.^{20,25} As the MIC result from the 0.2 SD method was below the smallest detectable change we would not recommend this method be used in the future. However, both distribution-based methods produced lower MIC estimates than anchor-based methods. Either distribution-based methods are underestimating the MIC, or it is quite possible that the anchors used here were too broad to capture smaller yet important changes.

Gaunt *et al.* used multiple methods and reported a variation in the MIC ranging from 2.5 to 4.27 points in a sample of children with milder eczema than the sample in the current study.¹¹ Schram *et al.* used an anchor-based method and reported an MIC of 3.4 in adults with severe eczema.¹⁰ We have synthesized the findings from these previously published studies and the current study to provide recommendations of how to interpret changes on the POEM (Table 5).

There is a balance to be struck when estimating sample sizes for powering trials.²⁶ Powering to detect a nonclinically important change is unethical and wasteful, as it will result in overly large trials. However, an underpowered trial based on

Table 5 A guide to enhance interpretations of change on the Patient-Oriented Eczema Measure (POEM)

Change in the POEM	Suggested interpretation
≤ 2	Unlikely to be a change beyond measurement error
2.1–2.9	A small change detected that is likely to be beyond measurement error but unlikely to be clinically important
3–3.9	Probably a clinically important change
≥ 4	Very likely to be a clinically important change

detecting a large change in the POEM that provides inconclusive results and wide confidence intervals is also unacceptable.

The recommendations provided in this article should remain as guidelines. Rather than relying on fixed values to interpret the importance of a change on the POEM, researchers and clinicians should consider the context within which they are using the POEM. A small improvement in many individuals could result in a large reduction in burden at a societal level.

This study has some strengths and limitations. It estimated the MIC of the POEM using a broad array of methods but only in one dataset, which may limit the generalizability of the results beyond children in the U.K. with moderate-to-severe eczema. However, Gaunt *et al.* and Schram *et al.* used similar methods to calculate the MIC, which has allowed us to compare the results across these different populations.^{10,11} This study also included methods that have not previously been used to determine the MIC of the POEM.

The anchors used were determined by what was available in the CLOTHES dataset, which may not best conceptualize an MIC. The anchor did not ask about the importance of the change to the patients/parents or the investigator, which has been a criticism of anchors generally used to calculate MIC estimates.⁸ Furthermore, there is concern that the anchor-response categories may be too broad to capture the smallest amount of change that is important to patients, as MIC is equated to a change on severity banding in a global assessment, and it could be argued that smaller changes are clinically meaningful. The anchors were also static (asked about today), whereas POEM asks about the last week.

The ICC of the POEM has only been calculated in one study, and this was used to assess the smallest detectable change.² The sample age from the original study completing the test–retest reliability ranged from 12 months to 62 years, compared with the CLOTHES trial age range of 1–15 years. Further investigation of the test–retest agreement would be useful to increase the reliability of the ICC estimate.

This study should improve the interpretability of change scores for users of the POEM, be this for clinical trial sample size calculations and interpretation of findings or for use in clinical practice. However, there are still areas where understanding of the MIC could be further developed. If the ICC of the POEM used here is found to be inaccurate in subsequent studies, the smallest detectable change of the POEM should be

replicated. In the meantime, users of the POEM should be cautious in claiming any clinical importance of a change in POEM scores of two points or less. This study used one trial dataset of patients with eczema to look at the MIC of the POEM. Further work should assess how patient characteristics such as eczema severity, age, sex and ethnicity influence the MIC.

In conclusion, given the wide spread of MIC estimates generated using a variety of methods, the method used to calculate the MIC should be given careful consideration. Based on the evidence available for interpreting change in the POEM, we suggest that no change in score that is two points or under should be considered beyond measurement error and that changes in scores should usually be three points or above before the change is deemed to be clinically important.

Acknowledgments

We would like to thank Dr Berend Terluin and Dr Jessica de Boer for their advice and support with the statistical analysis. We would also like to thank the CLOTHES trial team for their support with this project.

References

- Mokkink LB, Terwee CB, Patrick DL *et al.* COSMIN Checklist Manual. Amsterdam: VU University Medical Centre, 2009.
- Charman CR, Venn AJ, Williams HC. The patient-oriented eczema measure: development and initial validation of a new tool for measuring atopic eczema severity from the patients' perspective. *Arch Dermatol* 2004; **140**:1513–19.
- Chalmers JR, Simpson E, Apfelbacher CJ *et al.* Report from the fourth international consensus meeting to harmonize core outcome measures for atopic eczema/dermatitis clinical trials (HOME initiative). *Br J Dermatol* 2016; **175**:69–79.
- National Institute for Health and Care Excellence. Atopic Eczema in Under 12s: Diagnosis and Management. London: NICE, 2007.
- Wright A, Hannon J, Hegedus EJ *et al.* Clinimetrics corner: a closer look at the minimal clinically important difference (MCID). *J Man Manip Ther* 2012; **20**:160–6.
- de Vet HCW, Ostelo RWJG, Terwee CB *et al.* Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res* 2006; **16**:131–42.
- Terwee CB, Roorda LD, Knol DL *et al.* Linking measurement error to minimal important change of patient-reported outcomes. *J Clin Epidemiol* 2009; **62**:1062–7.
- Terwee CB, Roorda LD, Dekker J *et al.* Mind the MIC: large variation among populations and methods. *J Clin Epidemiol* 2010; **63**:524–34.
- De Vet HC, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. New York, NY: Cambridge University Press, 2011.
- Schram ME, Spuls PI, Leeflang MM *et al.* EASI, (objective) SCORAD and POEM for atopic eczema: responsiveness and minimal clinically important difference. *Allergy* 2012; **67**:99–106.
- Gaunt DM, Metcalfe C, Ridd M. The Patient-Oriented Eczema Measure in young children: responsiveness and minimal clinically important difference. *Allergy* 2016; **71**:1620–5.
- Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol* 2002; **14**:109–14.

- 13 Revicki D, Hays RD, Cella D *et al.* Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; **61**:102–9.
- 14 Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003; **56**:395–407.
- 15 Gerbens LA, Prinsen CA, Chalmers JR *et al.* Evaluation of the measurement properties of symptom measurement instruments for atopic eczema: a systematic review. *Allergy* 2017; **72**:146–63.
- 16 Thomas KS, Bradshaw LE, Sach TH *et al.* Silk garments plus standard care compared with standard care for treating eczema in children: a randomised, controlled, observer-blind, pragmatic trial (CLOTHES Trial). *PLOS Med* 2017; **14**:e1002280.
- 17 Emerson RM, Charman CR, Williams HC. The Nottingham Eczema Severity Score: preliminary refinement of the Rajka and Langeland grading. *Br J Dermatol* 2000; **142**:288–97.
- 18 Centre of Evidence Based Dermatology, University of Nottingham. POEM — Patient Oriented Eczema Measure. Available at: <http://www.nottingham.ac.uk/research/groups/cebd/resources/poem.aspx> (last accessed 29 March 2018).
- 19 Copay AG, Subach BR, Glassman SD *et al.* Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007; **7**:541–6.
- 20 Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003; **41**:582–92.
- 21 Terluin B, Eekhout I, Terwee CB *et al.* Minimal important change (MIC) based on a predictive modeling approach was more precise than MIC based on ROC analysis. *J Clin Epidemiol* 2015; **68**:1388–96.
- 22 Terluin B, Eekhout I, Terwee CB. The anchor-based minimal important change (MIC), based on ROC analysis or predictive modeling, may need to be adjusted for the proportion of improved patients. *J Clin Epidemiol* 2017; **83**:90–100.
- 23 de Vet HC, Terwee CB, Ostelo RW *et al.* Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes* 2006; **4**:1–5.
- 24 Turner D, Schunemann HJ, Griffith LE *et al.* The minimal detectable change cannot reliably replace the minimal important difference. *J Clin Epidemiol* 2010; **63**:28–36.
- 25 Norman GR, Sloan JA, Wyrwich KW. The truly remarkable universality of half a standard deviation: confirmation through another look. *Expert Rev Pharmacoecon Outcomes Res* 2004; **4**:581–5.
- 26 Cook JA, Hislop J, Altman DG *et al.* Specifying the target difference in the primary outcome for a randomised controlled trial: guidance for researchers. *Trials* 2015; **16**:12.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Video S1 Author video.

Powerpoint S1 Journal Club Slide Set.