ORIGINAL ARTICLE

Outcome measures for clinical trials in paediatric IBD: an evidence-based, expert-driven practical statement paper of the paediatric ECCO committee

Frank M Ruemmele,1,2,3 Jeffrey S Hyams,4 Anthony Otley,5 Anne Griffiths,6 Kajia-Leena Kolho,7 Jorge Amil Dias,8 Arie Levine,9 Johanna C Escher,10 Jan Taminiou,11 Gabor Veres,12 Jean-Frederic Colombel,13 Séverine Vermeire,14 David C Wilson,15 Dan Turner16

ABSTRACT

Objective Although paediatric-onset IBD is becoming more common, few medications have a registered paediatric indication. There are multiple hurdles to performing clinical trials in children, emphasising the importance of choosing an appropriate outcome measure, which can facilitate enrolment, and thereby also drug approval. The aim of this consensus statement is to highlight paediatric specific issues and key factors critical for the optimal conduct of paediatric IBD trials.

Design The Paediatric European Crohn’s and Colitis Organisation (ECCO) committee has established an international expert panel to determine the best outcome measures in paediatric IBD, following a literature search and a modified Delphi process. All recommendations were endorsed by at least 80% agreement.

Results Recognising the importance of mucosal healing (MH), the panel defined steroid-free MH as primary outcome measure for all drugs of new category with one or two postintervention endoscopies per trial (at 8–12 weeks and/or 54 weeks). Since endoscopic evaluation is a barrier for recruitment in children, trials with medications already shown to induce MH in children or adults, could use paediatric-specific disease activity scores as primary outcome, including a modified Paediatric Crohn’s Disease Activity Index in Crohn’s disease and the Paediatric Ulcerative Colitis Activity Index in UC. Secondary outcomes should include safety issues, MR enterography-based damage and inflammatory scores (in Crohn’s disease), faecal calprotectin, quality of life scales, and a patient-reported outcome.

Conclusions It is crucial to perform paediatric trials early in the development of new drugs in order to reduce off-label use of IBD medication in children. The thoughtful choice of feasible and standardised outcome measures can help move us towards this goal.

INTRODUCTION

Paediatric forms of IBD are characterised by a more complicated disease course with marked inflammatory activity and subsequently frequent need for corticosteroids and immunosuppressive therapy compared with adult-onset IBD.1 2 However, paediatric clinical trials are most often initiated long after the approval of the drug for adults. Consequently, treatment strategies for paediatric...
IBD are often based on expert opinion and extrapolation from experience in adults rather than on level-1 evidence.

While consideration of the validity of a measure is important, in order to facilitate research in this vulnerable population, the design of clinical trials in children must also address barriers preventing the recruitment of children to the trials. The eligible population for recruitment is small given the lower number of incident and prevalent paediatric IBD cases compared with adult cases. Parents, concerned about potential side effects of therapy and the additional invasive tests, are more reluctant to have their children engaged in intervention trials than are adult patients. Many clinicians also express similar hesitancies in the face of invasive procedures. The fact that paediatric trials are often only confirmatory to similar larger adult trials should be used as an advantage to balance the challenging recruitment in children. These considerations are fundamental when designing clinical trials in children and when determining primary outcome measures in order to increase feasibility of paediatric trials and thus avoiding the current situation that many medications are given to children ‘off-label’.

On the other hand, well-designed clinical trials are mandatory to improve care for IBD in adults and children, particularly in industry-driven clinical trials of emerging therapies. The lack of correlation between intestinal mucosal inflammation and symptomatic scores is a challenge when designing clinical trials for Crohn’s disease (CD). In contrast, the correlation between intestinal inflammation and clinical disease indices is stronger in UC. Endoscopy facilitates direct evaluation of mucosal inflammation, but may be more difficult to repeatedly perform in children, whereas MR enterography (MRE) seems very promising. Clinical outcome measures should be able to predict predefined goals important to patients including long-term well-being, quality of life and prevention of damage, as well as tolerance and safety issues. Long-term safety is especially important in guiding translation of clinical trials towards clinical practice.

The aim of this consensus paediatric ECCO statement paper is to highlight age-specific considerations when selecting outcome measures in paediatric IBD research.

**METHODS**

The members of the paediatric committee of ECCO (FMR, K-LK, JAD, AL, JCE, GV, DT) established an international panel of experts in paediatric IBD and measurement methodology (JSH, AO, AG, JT, DCW) along with two adult experts (J-FC, SV). According to ECCO ‘GutCom’ guidelines a systematic approach was used: The panel was subdivided into working groups to address a predefined list of specific topics for paediatric UC and, separately, CD, following a literature search and a modified Delphi process. Representatives of the European Medicines Agency (EMA) and US Food and Drug Administration (FDA) were invited to participate in this process and representatives from EMA formed part of the discussion group. Upon literature review, relevant statements were proposed by each subcommittee and submitted to the entire group for extensive discussion followed by a first voting round on the proposed wording. Four of the initial 25 statements did not achieve the predefined 80% agreement cut-off (see online supplementary appendix 1). The panel extensively discussed the diverging aspects, and after rewording and modification all recommendations were endorsed by at least 80% agreement. After external peer review the total number of statements was reduced to 21 due to overlap of 4 statements.

**DISCUSSION**

**The choice of outcome measures**

It is widely accepted that a desired outcome of our therapies is not simply short-term improvement of symptoms but rather remission that predicts durable well-being. Therefore particular attention should be given to the definition of the primary outcome by which a clinical trial is judged to be positive or negative. To alleviate symptoms of a patient is always desirable; however, the efficacy of drugs should be measured by its potential to control inflammation. Improved medium-term and long-term outcomes have been associated with achievement of MH in UC8,9 and CD.6,7 Nonetheless, multiple secondary outcome

**General considerations for paediatric Crohn’s Disease and paediatric UC**

1. The ultimate treatment goal for all children/adolescents with IBD is to efficiently control inflammation reflected by disappearance of mucosal lesions. Steroid-free mucosal healing (MH) as assessed by endoscopy is recommended as the primary end point for all preauthorisation trials with medicines of new drug category (93% agreement)

2. Disease activity scores (weighted Paediatric Crohn’s Disease Activity Index (wPCDAI) or Paediatric Ulcerative Colitis Activity Index (PUCAI)) can be considered as primary end points in studies that test therapies already shown to induce MH (in adults) and if they do not represent a new drug category (93% agreement)

3. Design of clinical trials based on disease activity scores as primary outcome measures should include evaluation of endoscopic MH as a co-primary or secondary outcome measure in a subpopulation of patients (83% agreement)

4. Depending on the intervention under study, consideration should be given to include evaluation of efficacy for induction of remission, and then maintenance of remission within the same trial rather than designing separate clinical trials (100% agreement)

5. The primary end point for assessing remission in induction trials should be around Week 6–12 of therapy with some flexibility according to the mechanism of action of the intervention under study (100% agreement)

6. At least a 12-month-period is mandatory to evaluate maintenance of steroid-free remission (100% agreement)

7. Placebo-controlled trials are rarely acceptable in the design of clinical trials for the vulnerable population of children with IBD. A placebo may be acceptable if an adjunctive therapy is being studied, whereby both study groups (treatment and control) are receiving effective therapy (93% agreement)

8. For studies evaluating intestinal damage (eg, fistula) a minimum of 12 months may be required with a longer duration potentially being more informative (100% agreement)

9. Although superior to C reactive protein (CRP) and erythrocyte sedimentation rate (ESR) in reflecting the degree of inflammation, faecal calprotectin (FC) should be used as a secondary outcome measure only, because of the large variability in results leading to low precision, the moderate responsiveness to short-term change and lack of accepted cut-off values for categories of disease activity (90% agreement)

10. Other secondary outcomes should include safety assessment, patient-reported outcomes and health-related quality of life (as measured by the IMPACT questionnaire) (100% agreement)
measures are often included allowing extraction of valuable data relevant for the interpretation of the trial. To combine different outcome measures as primary (co-primary) end points is another elegant way to design clinical trials, however this often increases the number of patients needed. No gold standard exists to reflect the concept of ‘disease activity’, an assessment based on the conglomerate of symptoms, signs, radiographic appearance, presence and severity of inflammation (macroscopic, microscopic findings) and biomarkers. Of all of these measures, the macroscopic appearance of the intestinal mucosa is currently considered to be the most direct assessment of intestinal inflammation. Another major challenge is that there is as yet no validated definition of MH for CD. Defining MH as total absence of mucosal lesions may be too rigid a target, and not one that is achievable with current interventions. The Extend the safety and efficacy of adalimumab through endoscopic healing (EXTEND) study (in adult patients with CD) used absence of mucosal ulcerations in all segments of the ileocolon to define achievement of MH. However, with this restricted definition it is not clear how one would handle an incomplete endoscopy whereby, for example, the proximal colon and terminal ileum are not visualised. Since endoscopic evaluation can be subjective, central reading of endoscopic videos by trained assessors is important to improve consistency in assigning endoscopic scores.

There are situations when the clinical assessment better reflects the improved current state of a patient, while endoscopic improvement clearly lags behind this clinical improvement. Indeed, a recent post hoc analysis of the T-72 infliximab trial in children with UC showed that clinical remission (judged by the PUCAI) was not inferior to sigmoidoscopy in predicting 1-year outcomes as primary (co-primary) end points is relevant for the interpretation of the trial. To combine different objective measures of inflammation, such as FC or serum markers (CRP or ESR), or imaging should be associated as secondary outcome measures. The inclusion of new activity indices for use as primary or secondary outcome measures in IBD: patients themselves (ie, by patient reported outcome) or their physicians. The issue of the different perspectives of doctors and patients is not that one is right or wrong—both are equally valid, but failure to acknowledge the differences results in less effective evaluation of treatment effects. Although physicians obtain the clinical data from patients, complex processes are involved in hearing and interpreting patient reports, and in integrating these reports with other information to obtain a comprehensive picture of the disease status. In a prospective head-to-head comparison in UC, clinicians’ assessment of symptoms and signs were more closely correlated with the degree of colonoscopic inflammation in comparison with their patients, in children and adults. Patient assessments, physician assessments and direct measurement of the mucosa provide complementary information in clinical research. Primary outcome measures must reflect the degree of inflammation (ie, disease activity) which should be associated with long-term well-being. Multi-item indices that are aimed to proxy the level of inflammation (such as PUCAI and wPCDAI) should be developed and scored by clinicians, after obtaining a thorough history from the patients and considering the constellation of symptoms, signs and laboratory results. Patients should score directly quality of life and Patient Reported Outcome (PRO) instruments since physicians and parent proxies often underestimate functional disabilities and quality of life reported by their patients/children.

Use of placebo
Clinical trials must be feasible, ethical and able to evaluate the clinical end points incorporated into the study design. While a randomised, double blind parallel group comparison is considered the optimal study design for evaluating efficacy of a new therapy, this can lead to ethical and feasibility issues in the paediatric context. Whenever a standard therapy is available, most ethics review boards and paediatricians would consider an untreated placebo arm as unethical in children. The inclusion of a placebo arm in an intervention trial is possible when the tenet of clinical equipoise exists—meaning that investigators should have no decisive evidence that active therapy is superior to the control arm (potentially no therapy). Because to date virtually all paediatric IBD clinical trials have been conducted years after the active/study drug has been shown to be effective in adult patients, the critical factor for use of a placebo, clinical equipoise, no longer exists. Moreover, most parents will not consent to have their child treated with placebo alone if the active medication or an alternative treatment can be obtained outside the study. The panel unanimously agreed that placebo-controlled trials, without an active treatment in each treatment arm, are not appropriate for most paediatric trials. Not...
incorporating a placebo-controlled study design differs from the adult IBD EMA guidelines, which state: “Unless the study is aiming at demonstrating superiority, the trial should (when ethically justifiable) also include a placebo arm to provide internal validation of the study” (EMA: Pre-authorisation Evaluation of Medicines for Human Use. Guideline on the development of new medicinal products for the treatment of CD, 2008. Doc. Ref. CPMP/EWP/2284/99 Rev. 1).

Biological markers
Serum markers, most notably CRP and ESR, often reflect inflammation but not in all patients, especially in UC. Faecal markers are increasingly recognised as important markers to reflect mucosal inflammation. In UC, FC is superior to CRP and ESR in reflecting the degree of active inflammation.26 Bowel cleansing procedure may alter the observed level.27 A strictly normal FC (in many laboratories defined as <50–100 µg/g) is highly suggestive of MH,28 but higher values are more difficult to interpret also due to the large variance of the values (negative predictive value for the presence of active inflammation is 87–95% with lower positive predictive value28). In UC, the correlation between calprotectin and colonoscopic scores ranges between studies with r ~0.4–0.6.30–32 with one outlying r=0.83 as compared with clinical indices that correlate in a range of 0.65–0.8.34 In paediatric UC levels as high as 500 µg/g35 to 800 µg/g36 have been suggested as the best cut-off to facilitate recognition of preclinical relapse. On the other hand, FC normalises in a significant percentage of patients after effective treatment.37 38 In clinical CD-trials FC has been measured after induction therapy using exclusive enteral nutrition and it has been slowest to return to normal levels, implicating limited responsiveness to treatment.39 In adult IBD, FC has been shown to be useful to predict relapse by systematic review and meta-analysis with wide variability.40 Taken together the panel concludes that in CD the currently available serum and faecal biomarkers, and in UC, faecal biomarkers should be monitored as secondary rather than primary outcomes.

Health-related quality of life scores
The evaluation of quality of life is an important aspect of management of chronic disease in children and adolescents. The paediatric IBD disease-specific health-related quality of life (HRQOL) instrument—the IMPACT questionnaire—was developed by a multidisciplinary research team41 42 using extensive patient item generation and reduction, and has been evaluated widely.43 Based on this and other evaluation, IMPACT II and then IMPACT III were developed44 and validated internationally.45 Of note, non-disease-specific variables, particularly behavioural dysfunction have been shown to have an important impact on HRQOL in paediatric IBD.46 Given the importance of assessing the impact of therapeutic interventions on HRQOL, use of the IMPACT questionnaire as a secondary outcome measure in clinical trials has been recommended and of the versions available, IMPACT III is felt to be best suited to international multicentre trials, by virtue of its ease of translation and scoring.47 48 49 There are currently over 40 cross-cultural translations available of this questionnaire. IMPACT III has been used in the keynote paediatric CD randomised controlled trials (RCTs) of adalimumab46 47 and infliximab.49 The inclusion of this self-report outcome measure allows patients to express the impact of their disease on their daily life and quality of life. Since an important outcome measure is to improve the quality of life of children and adolescents with IBD, it is indispensable to evaluate the effect by a standardised and validated tool reflecting the patients’ perception.

Specific considerations for paediatric CD
1. For new drug categories and new medicines, it is recommended to design a preregistration study evaluating the percentage of patients with CD achieving steroid-free MH as the primary end point (93% agreement)
2. MH is best evaluated by ileocolonoscopy showing disappearance of mucosal ulcerations based on the Simple Endoscopic Score for Crohn’s Disease (SES-CD)<3 evaluated by central reading (92% agreement)
3. In other trials when endoscopy is waived (eg, the drug under study is not a new category), the PCDAI, a multi-item measure of disease activity, should serve as the primary outcome measure for induction and maintenance of remission trials. Of several versions available, the mathematically weighted index (wPCDAI) has shown the best performance (85% agreement)
4. wPCDAI cut-off scores for inactive disease and for grading of CD activity from mild to severe (see discussion and online supplementary appendix 2) can be used to screen inclusion criteria for eligibility (90% agreement)
5. The primary outcome for maintenance trials, which rely on clinical assessment, should be sustained corticosteroid-free remission (wPCDAI≤12.5) at Week 30 and Week 52, without any rescue treatment. Remission and response rates by the wPCDAI should be measured at all other study visits (100% agreement)
6. In maintenance of remission trials, an important paediatric-specific secondary outcome measure in children with ongoing linear growth potential is height velocity (100% agreement)

Mucosal healing and endoscopic evaluation
Endoscopic evaluation in CD is controversial. On one hand, the correlation between PCDAI and mucosal inflammation is poor but on the other hand paediatric-onset CD is characterised by significantly more extensive involvement of the GI tract, often beyond the reach of standard ileocolonoscopy.2 As a transmural pan-intestinal disease, inflammation in CD needs to be controlled not just in the mucosa visible during endoscopy, but also in all layers of the gut wall (as discussed in the recent ECCO-European Society of Pediatric Gastroenterology Hepatology and Nutrition (ESPGHAN) Guidelines on the management of paediatric CD).50 These shortcomings may be overcome by a MRE measure, but more evidence of validation is needed before using as primary outcome in clinical trials. This point was extensively discussed within the expert panel given the benefit of MRE evaluation (entire GI tract as well as transmural inflammation), a 100% agreement was reached that appropriate validation is indispensable to the introduction of MRE as a tool to assess MH. Once this validation is done, MRE might become the preferred tool to evaluate MH in the future. Video capsule endoscopy might be an attractive tool in children, but the lack of standardisation and validation as well as the current inability to assess the colon limits its use as a primary outcome in CD.51-53 Even if used as a secondary outcome, central reading and assessment of video capsule endoscopy recordings is indispensable for clinical trials. As in UC, the optimal time to assess MH in active CD has not been established. A number of trials have assessed MH at 10–12 weeks in adult CD trials (EXTEND 12 weeks, Mucosal Healing Study in Crohn’s Disease (MUSIC) 10 weeks),11 54 rather than the traditional outcome assessment of 8 weeks for a CD induction trial.

The Crohn’s Disease Endoscopic Index of Severity (CDEIS) records the presence and extent of preselected lesions in the rectum, sigmoid and left colon, transverse colon, right colon and
Inflammatory bowel disease

ileum. The SES-CD was developed as a more user-friendly method to assess disease activity with excellent correlation to the CDEIS. The measures and their respective weighting scores are shown in online supplementary appendix 2. A subsequent study has shown good correlation of the CDEIS and the SES-CD. A recent evaluation showed superior reliability of the SES-CD compared with the CDEIS. To allow accurate evaluation, all clinical trials should include video recording of the endoscopic procedure with central reading and assessment by an expert panel blinded to the clinical data and treatment arm.

One major challenge is to define MH in CD. Ideally, this would mean a SES-CD or CDEIS of 0. Since, this is not realistic, the expert panel suggests for CD the definition of ‘disappearance of ulcerations’ as MH after induction therapy (SES-CD <3 or CDEIS <3) in line with recent ECCO guidelines and also in keeping with recent RCTs in adult patients with CD (EXTEND). Similarly in maintenance trials, only the absence of ulcerations can be considered as MH (CDEIS <3 or SES-CD <3). Thus, only patients with mucosal lesions (ulcerations) prior treatment accessible to standard endoscopy should be included in clinical trials ensuring precise evaluation on follow-up endoscopy. This means that patients with a stenotic ileocaecal valve without signs of active inflammation/ulcerations should not be included in this type of trials. In case of incomplete endoscopic evaluation (failure to intubate ileocaecal valve) the CDEIS is still usable in patients with L2 or L3 disease presentation at baseline, since the total score is divided by the number of segments explored.

Disease activity scores

PCDAI is the oldest paediatric CD activity score, and also the most used and the most evaluated. Following its initial description, it has undergone prospective evaluation. Thereafter the cut-off scores were reassessed corresponding to response, remission and gradations of disease activity, plus the clinimetric properties of the PCDAI, using four prospectively collected data sets. The CDAI has been evaluated against PCDAI for paediatric CD, and PCDAI was shown to more accurately classify CD activity in children versus physician's global assessment. The expert panel suggests completing the CDAI (for physically mature adolescent patients who have achieved adult height) as additional secondary outcome measures to allow comparison with adult trials.

Despite the body of clinical and research evidence involving the PCDAI, there have been long-standing concerns about its sensitivity, such as the need for measurement of height velocity, which is not applicable for adolescents who have achieved their final adult height, and for venipuncture. An abbreviated PCDAI (abbrPCDAI) was therefore developed which omitted height measurement, presence of extraintestinal disease and the three laboratory values and then separately evaluated. A short form PCDAI (shPCDAI) was created by retaining and reweighting components of the PCDAI completed in more than 80% of visits in a large paediatric IBD registry. Lastly, a mathematically weighted version (wPCDAI) was developed and compared with PCDAI, abbrPCDAI and shPCDAI using four prospectively collected paediatric CD data sets and totalling 437 children. Key features of the wPCDAI included the removal of three items that have shown to be redundant on a multivariable analysis, two of low feasibility (height velocity and abdominal examination) and one blood measurement (haematocrit). The wPCDAI (see online supplementary appendix 2) had better performance than the PCDAI in construct validity and responsiveness, appears more feasible and it discriminated better between the disease activity categories (area under the receiver operating characteristic (ROC) curve of 0.97; 95% CI 0.95 to 0.99). The non-invasive versions (shPCDAI and abbrPCDAI) are more feasible, but their performance was inferior to the wPCDAI.

There is a paucity of data regarding integration of a disease activity score and biological markers. One paediatric study in 222 patients with 52 weeks of follow-up showed that normal CRP corticosteroid-free remission at Week 12 was predictive of sustained remission in comparison with clinical remission with an elevated CRP at Week 12. The fact that CRP is not elevated in all patients with active CD remains a limitation. Studies comparing or validating clinical and biological marker-based outcomes against long-term outcomes are currently a research gap.

Imaging techniques and measures in paediatric CD

In contrast with UC, where the inflamed organ can be easily examined by colonoscopy, in CD large parts of the small bowel potentially having lesions cannot be evaluated by endoscopy. Therefore, imaging techniques are taking an increasingly important place in the evaluation of patients with CD, and also allow the determination of transmural involvement. These techniques have included barium contrast radiography (small bowel series, enteroclysis), CT enterography, ultrasonography and MRE. Imaging modalities that use ionising radiation should be discouraged in the research setting, especially in children. Panés et al found in a meta-analysis of MRE a sensitivity and specificity of 78% (95% CI 67% to 84%) and 85% (95% CI 76% to 90%), respectively, for detecting mucosal inflammation in children and adults. Giles et al reviewed six paediatric studies and reported sensitivity and specificity of 0.84 (95% CI 0.77 to 0.90) and 0.97 (0.91 to 0.99), respectively.

The magnetic resonance index of activity (MaRIA) score, the most widely used inflammatory scoring system for MRE in adults, includes evaluation of wall thickness, relative contrast enhancement, oedema and ulcerations. A global score can be determined by adding the scores from the rectum, sigmoid, descending, transverse and ascending colon, and ileum. The global MaRIA score correlates well with CDEIS, the Harvey-Bradshaw score, and CRP in the adult cohort studied. For the evaluation of perianal disease there is a score based on T2 signal on MRI. A study to develop a paediatric MRI-based inflammation score is underway (ie, Paediatric Inflammatory Crohn’s MRE Index, termed PICMI). In adult patients with CD, the Lémann Score is a validated MRE-based score to measure intestinal damage, which is increasingly recognised as an important outcome measure to judge whether an intervention changes the natural history of the disease. The paediatric version, the Paediatric MRE-based Damage Index in Crohn’s disease (pMEDIC) is currently being validated.

Ultrasoundography is operator dependent and attempts to standardise ultrasonographic evaluation have not been widely accepted. However, new techniques including Doppler flow measurements may overcome these hurdles and seem promising for the future pending more paediatric validation data.

Assessment of growth as an outcome measure

Impaired linear growth is common in paediatric CD and reflects intestinal inflammation as well as poor nutrition. However, inclusion of linear growth in activity indices can be problematic, as discussed for the PCDAI. If participants have attained their final adult height before or during a clinical trial, and are
compared with participants of similar age who still have growth potential, this could impact their disease activity score despite there being no real difference between the disease activities of the two participants. For subjects who have not yet reached skeletal maturity during a clinical trial, linear growth as determined by growth velocity and stratified by Tanner stage/bone age should be determined as a secondary outcome. At least 6 months, and preferably 12 months, between data points is required for accurate assessment. Growth is not a valid end point in the context of a short-term induction of remission trial for paediatric CD.

It has been found in most studies that reduced bone mineral density may be seen in approximately half of patients with IBD with a 20–40% increased risk for fractures, especially in patients with CD. The growing skeleton may be particularly vulnerable to the detrimental effects of chronic inflammation and corticosteroids on bone formation. It is paramount to optimise bone density in childhood since the lifetime peak density is achieved by the age of ~20 years. Therefore, investigators are encouraged to include bone markers as secondary outcomes of trials, before and after the intervention.

Specific considerations for paediatric UC
1. For new drug categories that involve maintenance of remission, steroid-free MH should be the primary end point (evaluated by endoscopy (at enrolment and at 6–12 months)) (92% agreement)
2. MH is best evaluated by colonoscopy showing disappearance of mucosal ulcerations based on a MAYO score of 0 evaluated by central reading (100% agreement)
3. In trials when endoscopy is waived, the primary outcome measures should reflect the percentage of patients achieving or maintaining corticosteroid-free remission, defined as a PUCAI score of <10 points, with the treatment being evaluated versus the comparison arm (100% agreement)
4. The primary outcome of induction trials should be clinical remission (PUCAI<10 points) at Week 8–12 and the primary outcome of ‘maintenance’ trials should be sustained relapse-free corticosteroid-free remission (PUCAI<10 points) at Week 30 and Week 54. Remission and response rates by the PUCAI should be also measured at all other study visits (92% agreement)
5. In trials based on activity indices, a measure of MH as evaluated by endoscopy (Mayo endoscopic subscore) should be included as an important secondary outcome measure or co-primary outcome measure in a subgroup of patients, allowing comparisons with adult clinical trial data (91% agreement)

Disease activity scores
The ultimate goal in treating children with UC is to achieve sustained steroid-free clinical remission, and thus the primary outcome should reflect the percentage of children achieving remission (as opposed to ‘response’). Clinical disease activity indices are valuable tools to evaluate disease activity in UC. The PUCAI (see online supplementary appendix 3) has high clinical metric properties and good correlation with the presence of colonic inflammation on endoscopic evaluation. In addition, the PUCAI also performed well when completed directly by the patients. PUCAI cut-off scores for inactive disease and for grades of UC activity from mild to severe should be used to screen for eligibility into maintenance of remission or active treatment trials (see online supplementary appendix 3). Comparisons of improvement from baseline in PUCAI or of mean PUCAI in treatment versus control groups are unacceptable primary outcome measures. In clinical trials among hospitalised patients with acute severe UC, PUCAI scores must be monitored daily, with clear protocol instructions for altering therapy. The PUCAI can be accurately used on a day-by-day basis, given its high responsiveness to change and should be combined for outcome measure with no need for salvage therapy by 30 days.

The patient’s perspective is also important to record as secondary outcomes by means of quality of life scales and patient reported outcomes. The development of a patient reported PUCAI is underway, where symptoms are reported directly by the patient rather than being assessed by the physician. In addition, the IMPACT questionnaire is suitable for children ≥9 years of age and has been validated following translation in multiple languages.

Mucosal healing and endoscopic evaluation
The most widely used endoscopic scoring system in children is the Mayo subscore of 0–3 points. Of the other available scores, the Powell-Tuck index and the Rachmilewitz index included three and four endoscopic descriptors, respectively. There are, however, no clear instructions for scoring in these endoscopic scores and the site of disease severity evaluation, for example, the rectum or sigmoid or whether the evaluation should include the most inflamed segment has not been defined. The new scoring index Ulcerative Colitis Endoscopic Index of Severity (UCEIS) and the recently reported Ulcerative Colitis Colonoscopic Index of Severity (for the entire colon) are not yet validated in children (see online supplementary appendix 3). The UCEIS is based on assessment in the most severe lesion of sigmoidoscopy and includes three descriptors of vascular pattern, bleeding and ulceration graded in three or four levels (friability is excluded). Until these are validated in children, the panel recommends using the Mayo score with central readings. It is still debatable whether a score of 0 or 1 should be used to define MH and the data are very sparse to base informed recommendation. Nonetheless, the panel reached consensus that only a Mayo score of 0 should be regarded as MH (100% agreement), since anything less cannot be termed as such (a Mayo score of 1 means mild inflammation and this is not MH). While some studies did not differentiate between the two groups and showed favourable outcome to both, it is plausible to assume that if the degree of mucosal inflammation is associated with the disease outcome, the less inflammation the better the outcome. Indeed, correlation between macroscopic and microscopic inflammations is good in Mayo 0 but not Mayo 1 and histological resolution is associated with improved long-term outcomes. Post hoc analysis of the Active Ulcerative Colitis Trial (ACT) studies showed superiority of Mayo score 0 over 1 in some (but not all) of the outcomes. Finally, in the T72 trial of infliximab in paediatric UC, remission rates at Week 8 were 33% using the PUCAI, 33% using the endoscopic Mayo score 0, 40% using the total Mayo score and as high as 68% using an endoscopic score of 0/1 which was clearly an outlier in the assessment. The panel agreed that the bar should be set at complete MH, which evidently is an achievable goal in children (100% agreement). Histological scoring cannot currently be used as an outcome measure due to lack of clear validation studies, low reliability of subjective scoring and high sampling variability.

Imaging techniques and measures in paediatric UC
The use of imaging techniques, such as ultrasound and MRE, to reflect colonic inflammation may increase in the future, but currently there is insufficient evidence to use imaging as outcome measures in paediatric UC.
Inflammatory bowel disease

Timing of assessment of primary and secondary outcomes

Studies of acute severe colitis in hospitalised patients will, according to accepted guidelines for management of acute severe colitis, have daily PUCAI as part of the trial design, so that therapy can be altered early. For patients receiving ambulatory treatment of UC in a clinical trial setting, weekly completion of PUCAI based on telephone contact is sensible. Week 8 has been designated as the primary time of outcome assessment in induction trials. An earlier time point (<8 weeks) can be considered for clinical remission (PUCAI <10) in rapid-onset interventions. Scheduled endoscopic evaluation should be assessed at 8–12 weeks to allow adequate time for MH. Completely steroid-free clinical remission should not be assessed before 12 weeks.

For maintenance of remission trials, a period of at least 1 year is required to accurately assess the benefit of the study drug in prevention of relapse. Longer periods should be encouraged. Since the concept of a maintenance trial should be reflected in longitudinal disease assessment, more than one time point should be incorporated in the primary outcome measure. More than two time points may set the bar too high for approval of medications with moderate effect, and therefore we recommend relapse-free and steroid-free sustained clinical remission at Week 30 and Week 54, as proposed in recent paediatric trials. In medications with moderate effect, and therefore we recommend than two time points may set the bar too high for approval of longitudinal disease assessment, more than one time point is required to accurately assess the benefit of the study drug in prevention of relapse. Longer periods should be encouraged. Since the concept of a maintenance trial should be reflected in

SUMMARY

Though there are clear phenotypical differences between paediatric IBD and adult IBD, it is presumed that pathogenesis, mechanisms of symptom development and histopathological features are comparable. In general, response rates to previously evaluated therapies have been similar. Therefore, paediatric trials should take into account the data already accumulated in the larger adult trials and should not start from ‘scratch’. Extrapolating some of the data from previously conducted adult trials may justify paediatric studies of smaller sample size and with fewer barriers to recruitment, both appropriate goals in this vulnerable population. However, extrapolation should always be based upon and include pharmacokinetics data, pharmacodynamics, and evaluation of potential and real side effects/toxicities. It is not possible to predict growth and bone-related issues from adult studies. Thus, a thoughtful balance should be determined individually for each proposed trial. For instance, there is a consensus that controlling inflammation must be proven for all new drug categories (seen as complete MH on endoscopic evaluation). However, endoscopic evaluation is an invasive procedure requiring anaesthesia in children, and therefore its inclusion as a primary outcome measure might lessen the feasibility of clinical trials. It is thus recommended for trials with drugs already shown to induce MH in adults that assessment be based on disease activity indices as primary outcome measure combined with various secondary outcome measures, such as endoscopic evaluation in a subset of patients, imaging, inflammatory surrogate markers and quality of life questionnaires. Efficacy of the drugs/medicines or treatment strategy to test can be expressed as rate of steroid-free remission, time to relapse, changes in disease activity scores, imaging or endoscopic scores, as well as the normalisation of inflammatory measures. It is important to include paediatric-specific details into the strategy of drug development programmes, granting children with IBD access to new medications in a reasonable time frame, not years after adult patients with IBD, as is the usual case now.

REFERENCES


63 Schoepfer AM, Beglinger C, Strannan A, et al. Fecal calprotectin more accurately reflects endoscopic activity of ulcerative colitis than the Lichtiger Index, C-reactive
Inflammatory bowel disease


97 Rahmilewitz E. Mathemtical weighting of the pediatric Crohn’s disease activity index (PCDAI) and comparison with its other short versions. Inflamm Bowel Dis 2011;18:55–62.


Outcome measures for clinical trials in paediatric IBD: an evidence-based, expert-driven practical statement paper of the paediatric ECCO committee

Frank M Rueemmele, Jeffrey S Hyams, Anthony Otley, Anne Griffiths, Kajaa-Leena Kolho, Jorge Amil Dias, Arie Levine, Johanna C Escher, Jan Taminiau, Gabor Veres, Jean-Frederic Colombel, Séverine Vermeire, David C Wilson and Dan Turner

Gut 2015 64: 438-446 originally published online May 12, 2014
doi: 10.1136/gutjnl-2014-307008

Updated information and services can be found at:
http://gut.bmj.com/content/64/3/438

These include:

Supplementary Material
Supplementary material can be found at:
http://gut.bmj.com/content/suppl/2014/05/12/gutjnl-2014-307008.DC1.html

References
This article cites 96 articles, 11 of which you can access for free at:
http://gut.bmj.com/content/64/3/438#BBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

- Crohn's disease (907)
- Ulcerative colitis (1106)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/