Deliverable number D3.1

“Evidence- and consensus-based guidance for the design, conduct and reporting of paediatric CTs (4)”

“Selection and measurement of outcomes in paediatric clinical trials”

Lead Beneficiary
Sick Kids

Author(s)
M Kapadia, P Thurairajah, M Offringa

Revision date
31 August 2015

Dissemination Level
Public

WP3 – Research tools to facilitate interoperability in paediatric research

Task 3.01 – Develop standardized sets of relevant outcome measures in paediatric research.

This project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement n. 261060
## Index

1 List of authors ........................................................................................................... 3
2 Abstract ...................................................................................................................... 4
3 Receivers of the document ......................................................................................... 6
4 Full Description of tool .............................................................................................. 7
   4.1 Introduction ............................................................................................................ 7
   4.2 Problem Statement .............................................................................................. 8
   4.3 Scope of GRiP T3.01 .......................................................................................... 8
   4.4 Audience ............................................................................................................. 9
   4.5 Methods ............................................................................................................... 9
      4.5.1 Scoping literature review ............................................................................. 10
      4.5.2 Development of draft tool ....................................................................... 10
      4.5.3 Qualitative feedback from industry partners ............................................. 10
      4.5.4 Survey with academic clinical trialists ....................................................... 10
      4.5.5 Final GRiP deliverable ............................................................................. 10
4.6 Results ..................................................................................................................... 11
   4.6.1 Scoping literature review ............................................................................. 11
   4.6.2 Themes identified for selection, measurement and reporting of outcomes ... 12
      Why: Rationale for selecting the outcome ............................................................ 12
      What: Qualification - Relevance of the outcome for the target population ...... 12
      How: Determination of how the outcome will be measured ................. 13
      Who: Determination of the most appropriate source of information ......... 14
      Where: Facilities at the location where the outcome is measured ............. 14
      When: Specification of time points of measuring outcome .................... 15
      Outcome reporting and Core Outcome Sets ............................................... 15
   4.6.3 Qualitative feedback from industry partners ............................................. 16
   4.6.4 Survey of clinical trialists ............................................................................ 16
   4.6.5 Final GRiP deliverable ............................................................................. 16
4.7 Knowledge translation plan ................................................................................... 22
4.8 Limitations and indications for future research ................................................. 22
   4.8.1 Limitations .................................................................................................... 22
   4.8.2 Future Research ......................................................................................... 23
      Empirical Evidence Gathering ..................................................................... 23
      Checklist evaluation against AGREE II guideline ..................................... 24
5 Document History .................................................................................................... 25
6 References ................................................................................................................. 25
7 Appendices .................................................................................................................. 28
   7.1 Appendix I: InSPECT (Instrument for Selection of Pediatric Endpoints in Clinical Trials – Exploratory and Confirmatory Trials) .................................................. 28
   7.2 Appendix II: Glossary of terms ..................................................................... 33
   7.3 Appendix III: Invitation to industry partners .............................................. 35
   7.4 Appendix IV: Survey – clinical trialists ....................................................... 36
# 1 List of authors

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mufiza Kapadia: <a href="mailto:mufiza.kapadia@sickkids.ca">mufiza.kapadia@sickkids.ca</a></td>
<td>Sick Kids</td>
<td>W, C, R</td>
</tr>
<tr>
<td>Martin Offringa: <a href="mailto:martin.offringa@sickkids.ca">martin.offringa@sickkids.ca</a></td>
<td>Sick Kids</td>
<td>W, C, R</td>
</tr>
<tr>
<td>Winnie Chan: <a href="mailto:winnie.chan@sickkids.ca">winnie.chan@sickkids.ca</a></td>
<td>Sick Kids</td>
<td>W, C, R</td>
</tr>
<tr>
<td>Pravheen Thurairajah: <a href="mailto:pravheen.thurairajah@sickkids.ca">pravheen.thurairajah@sickkids.ca</a></td>
<td>Sick Kids</td>
<td>W</td>
</tr>
</tbody>
</table>

**Reviewers**

<table>
<thead>
<tr>
<th>Name</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mark Turner: <a href="mailto:Mark.Turner@liverpool.ac.uk">Mark.Turner@liverpool.ac.uk</a></td>
<td>C, R</td>
</tr>
<tr>
<td>Elin Haf Davies: <a href="mailto:elin@elinhafdavies.co.uk">elin@elinhafdavies.co.uk</a></td>
<td>C, R</td>
</tr>
</tbody>
</table>

*W= research on topic, writing guideline  
C= participating in conference calls  
R= reviewing drafts and final deliverable*
2 Abstract

Introduction: Evidence shows that children's responses to medical treatment differs significantly from the adult response due to their developing anatomical, physiological and pharmacological make up. With the current encouragement of more trials designed and conducted separately in children [1-5], it is increasingly being recognized that the health outcomes in children are fundamentally different from outcomes in adults. Therefore, the methodology to select and measure outcomes is distinctly different from adults [6]. Studies show that a large proportion of successful phase II drug trials fail in phase III. A major contributor to this high failure rate is an “inadequate” phase II trial design including the selection of outcome(s) that provide limited or misleading information regarding the true efficacy/safety balance of the agent. Therefore, selecting the appropriate outcomes in phase II drug trials in order to increase the probability of success in phase III drug trial is a challenging yet critical step in pediatric drug development. Qualification and validation, key elements of the justification of the selected outcomes phase II-III drug trials, needs to be transparent. Unfortunately, the quality of the description of outcomes in current paediatric clinical trial protocols and reports is remarkably poor. To date, very little work has been done on the guidance of appropriate outcome selection and measurement in paediatric drug trials.

In this document we present the methods used to develop a GRIP tool for selecting, measuring and reporting outcomes in paediatric phase II and III clinical trials at the study design and reporting stage. We present a preliminary checklist developed following systematic evidence gathering, and feedback provided by end-users. The purpose is to assist clinical trialists to systematically select, measure and report outcomes for phase II and III drug trials in paediatric populations.

Aims and Objectives: The aim of the GRIP Task 3.01 is to contribute to the greater GRIP goal to standardize outcome selection and measurement at the design level of paediatric phase II and III drug trials, and outcome reporting at the time of publication of its results. GRIP T3.01 aims to inform recommendations for rigorous selection of outcomes in paediatric phase II and III drug trials, both at the protocol and reporting stage.

Methods: After formulation of the research question, a sensitive search for existing guidelines and recommendations on outcome selection in phase II and III paediatric drug trials was done in three steps: we 1) identified all existing guidelines from regulators (FDA and EMA) and WHO on the choice of outcomes in adult drug trials; 2) conducted a targeted Google search using the following search terms: 'outcome measure in clinical trial', 'guideline for outcome measure', ' paediatric outcome measure', and 'guideline for paediatric outcome measure'; 3) conducted a targeted literature search in MEDLINE using the Ovid interface (1946 to November Week 2 2014) for guidelines for outcome selection specific to phase II and III trials to ensure no key documents were missed. From the yield of these 3 gatherings, key themes were then identified on outcome selection in paediatric phase II and III drug trials. A draft checklist containing recommendations and justifications for outcome selection was formulated, and pilot tested with pharmaceutical industry partners and pediatric clinical trialists.

Results: A total of 17 guidelines or recommendation documents on outcome selection were identified. None of these generic guidelines contained specific guidance regarding the choice of outcomes in phase II and phase III in paediatric drug trials. To develop such guidance, six items/themes on outcome selection were identified from existing guidelines and the literature: i) Why: Rationale for selecting the outcome; ii) What: Outcome’s qualification and validity for the health condition, the age specific sub-population, and the intervention; iii) How: Determination of how the outcome will be measured; iv) Who: Determination of the most appropriate source of information; v) When: Timing and responsiveness of measurement, and vi) Where: Description of the location where the outcome measurement
will occur. These six themes are presented in a 14 item preliminary GRiP InSPECT (Instrument for Selecting Pediatric Endpoints in Clinical Trials) checklist.

**Conclusion:** Guidance for the selection of outcomes in phase II and III paediatric drug trials is needed. Our checklist informs justified and transparent outcome selection. The next steps could involve gathering of empirical evidence (e.g. from Pediatric Investigation Plans) and expert panel evaluation of the proposed checklist (e.g. against AGREEII criteria) in preparation for adoption.
### 3 Receivers of the document

<table>
<thead>
<tr>
<th>User group</th>
<th>Identified need(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRIP Beneficiaries</td>
<td>Being informed about the development of an evidence- and consensus based tool for the design, conduct and reporting of paediatric CTs in WP3</td>
</tr>
<tr>
<td>European Commission</td>
<td>Being informed about the development of an evidence- and consensus based tool for the design, conduct and reporting of paediatric CTs in WP3</td>
</tr>
<tr>
<td>Clinical investigators (e.g. paediatricians, methodologists, residents, PhD students)</td>
<td>Awareness of the tool, signposting it, stimulation of peers and junior colleagues to work in line with the tool when designing and conducting phase II-III drug trials.</td>
</tr>
<tr>
<td>Industry</td>
<td>Application of the tool in the design of phase II-III drug trials</td>
</tr>
<tr>
<td>Educators</td>
<td>Awareness of the tool, endorsement and implementation</td>
</tr>
<tr>
<td>Funders</td>
<td>Evaluation of the submitted grant applications for the comprehensive, transparent reporting of outcome selection and measurement for pediatric phase II-III drug trials</td>
</tr>
<tr>
<td>Journal editors</td>
<td>Evaluation of the submitted research articles for the comprehensive, transparent reporting of outcome selection and measurement</td>
</tr>
<tr>
<td>Parents organisation</td>
<td>Information that helps them to decide whether the proposed trial methodology is sound, especially whether the outcomes selected are qualified and acceptable</td>
</tr>
</tbody>
</table>
4 Full Description of tool

4.1 Introduction

Evidence shows that children’s responses to medical treatment differs significantly from the adult response due to their developing anatomical, physiological and pharmacological make up [7]. Children’s response to treatment differs based on their age, developmental stage, developmental size and other factors [1-5]. With the current encouragement of more trials designed and conducted separately in children [1-5], it is increasingly being recognized that the health outcomes in children are fundamentally different from outcomes in adults. Therefore, the methodology to select and measure outcomes is distinctly different from adults [6]. Due to the lack of age-appropriate and validated outcome measures, clinical trials performed in children have led to inconclusive results, leaving children, families, physicians, regulators and industry uncertain about the efficacy and safety profiles of many drugs. This is now recognized as an unethical and unacceptable situation.

There is a further complexity in relation to phase II and III drug trials in children. The therapeutic exploratory phases of new drug trials (phase II) explore their safety and potential clinical benefit in humans. Among other roles, the results of phase II studies provide the basis for confirmatory study design including selection of study outcomes for the next phase of the trial (EMA CPMP/ICH/291/95). The drug’s final approval through a regulator’s market authorization and the adoption of new therapeutic regimens by the clinical community are usually based on results from these phase III confirmatory trials.

Studies show that a large proportion of successful phase II drug trials fail in phase III [8]. This is frequently based on insufficient attention to dose-finding, selection of an inappropriate target population or safety concerns that arise during or after the study. About half of failures to move from Phase II to Phase III are due to problems with “efficacy”. This may mean that the drug may not actually have the intended effect. An important contributor to this high failure rate is likely to be “inadequate” attention to the selection of outcomes during Phase II -providing limited or even misleading information regarding the efficacy of the test agent. There are also increasing challenges as the trial moves from demonstrating “efficacy” using surrogate endpoints (that justify the “pivotal” trials) towards demonstrating “efficacy” using clinically important endpoints (that justify licensing/marketing authorisation). Selecting the appropriate outcomes in phase II drug trials to ensure its success in phase III drug trial is therefore a challenging yet critical step.

Moreover, the quality of description of outcomes in paediatric clinical trial protocols and trial reports is remarkably poor. To date, very little work has been done to address the appropriate choice of outcomes in clinical trials in children.

A systematic review of phase II and III studies of advanced solid malignancies found that approximately 60% of the drug trials in oncology that were proven “successful” in phase II failed in phase III [1]. A major contributor to this high failure rate was an “inadequate” phase II trial design including the selection of outcomes that provides limited or even misleading information regarding the efficacy of the test agent, e.g., a tumour regression endpoint commonly used in oncology phase II trial may not translate to clinical benefit, and positive response to other markers of progression of the disease, such as progression-free survival (PFS) and/or time to progression (TTP) in phase II predicts the success of phase III trials (2).

Some outcomes in exploratory trials are about “go – no go” decisions, using outcomes to identify studies that are worthy of further investment. While surrogate outcomes utilized in exploratory trials present specific advantages in this phase of drug development, they are,
unfortunately, not always validated for their value in predicting clinical benefit, and therefore often do not translate into clinical outcomes that are relevant to the patient and the decision makers like national competent authorities who grant marketing authorizations, health care providers and payers.

However, surrogate outcomes are not the only outcomes in early phase clinical trials (see section 3.1.3.2 of ICH E8). Selection of dosage regimens etc. may be based on outcomes that are not intended to be reflected in pivotal trials and are not intended to reflect clinically meaningful outcomes.

Phase II and III drug trials must be designed, performed and reported to allow accurate interpretation of results and to obtain the best quality data to facilitate and inform unbiased decisions regarding the subsequent development of the drug(s) under study in the phase III setting. This go/no go decision is particularly important since phase III trials involve large number of patients and are expensive and time consuming. Therefore, selecting the appropriate outcomes in phase II drug trials in order to increase the probability of success in phase III drug trial is a challenging yet critical step in pediatric drug development.

In this document we present the methods used to develop a GRiP tool for selecting, measuring and reporting outcomes in paediatric phase II and III clinical trials at the study design and reporting stage. We present a new checklist developed following systematic evidence gathering, and feedback provided by end-users. The purpose of the checklist is to assist clinical trialists to systematically select, measure and report qualified and validated outcomes for phase II and III drug trials in paediatric sub-populations.

4.2 Problem Statement

Although guidance exists for specific elements of phase II and III trial design such as when to incorporate randomisation, there is little information to assist trialists and clinicians in choosing qualified, relevant and valid outcomes for paediatric clinical trials.

T 3.01 Research Questions

1) What frameworks, guidelines or tools are currently available for selection, measurement and reporting of qualified and validated outcomes in the design and reporting of phase II and III drug trials?
   a. Is there a method for sharing or harmonizing outcomes between exploratory and confirmatory drug trials?
   b. Is there a focus on children?
   c. If not, can these initiatives be adopted for phase II and III trials in children?

If there aren’t any current paediatric frameworks available for practical use then:

2) What stepwise approach could enhance selecting, measuring and reporting qualified and validated outcomes in the design and conduct of paediatric phase II and III drug trials?

4.3 Scope of GRiP T3.01

In this project we focus on one aspect of outcome selection for paediatric drug trials: outcomes that are shared between exploratory and confirmatory trials. This is a subset of the overall issue of selecting outcomes in exploratory and confirmatory trials. When an outcome is intended to be shared between stages of drug development, inappropriate
selection and measurement of the outcome can lead to waste in clinical research. This seems particularly important in paediatric clinical trials due to existing trends towards adopting outcomes from adult trials to be measured in young children, without evaluating their qualification, validity, feasibility, or responsiveness in the relevant age group [9].

We will focus on outcomes relevant to marketing authorizations. GRiP T3.01 aims to contribute to the greater GRIP goal by creating avenues towards standardizing outcomes selection and measurement, at the design stage of paediatric exploratory and confirmatory drug trials, and reporting of outcomes at the time of publication. As the focus of phase I pharmacology trials is to determine tolerance and drug toxicity with pre-defined pharmacological outcomes, our checklist will not focus on phase I trials. Recently, various research groups such as TORCH, OMERACT, COSMIN and COMET have been working on the development of guidelines for outcome selection in phase III and IV trials. These efforts will not be duplicated here.

4.4 Audience
This GRiP tool is intended for clinical investigators and others seeking to design, conduct or report Phase II or Phase III pediatric clinical trials. This includes, but is not limited to: pharmaceutical industry, and academic clinical investigators like pediatricians, and PhD students. Furthermore, it is intended for educators, and research and research infrastructure funders. Journal editors can use the tool as a checklist when making decisions about publication of paediatric clinical trials.

4.5 Methods
As in all other WP3 tasks, an environmental scan revealed that some of the originally planned activities had been taken up by other groups. To be most effective, a focus on outcome selection in early phase clinical drug development appeared to be the best use of our resources. In collaboration with the GRiP WP3 internal stakeholder group, the following five phase method was designed for GRIP Task 3.01. This agenda was approved in January 2015 by the GRiP Executive Board.

4.5.1 Scoping literature review
Objectives
I. to identify and critically appraise published guidance on outcome selection in early phase paediatric drug development
II. to identify and critically appraise published guidance on outcome selection in adult phase II-III drug trials and investigate its applicability in paediatric exploratory and confirmatory drug trials

A scoping review was performed to identify existing guidelines on outcome selection in early phase paediatric drug trials. A first step was to identify existing guidelines from regulators FDA and EMA and other sources (e.g. WHO, OMERACT) on the choice of outcomes in adult drug trials, and to explore their utility in paediatric exploratory and confirmatory drug trials. The purpose of searching for adult guidelines to select and measure outcomes was to develop a starting point or template to amend these guidelines in paediatric drug trials. Google search was conducted between October 7, 2014 to October 16, 2014. The following search terms were used to identify publications pertaining to outcome measures in clinical trials: ‘outcome measure in clinical trial’, ‘guideline for outcome measure’, ‘paediatric outcome measure’, and 'guideline for paediatric outcome measure'.

Concurrently, a brief targeted literature review was performed in MEDLINE using the Ovid interface (from 1946 to November Week 2 2014) for guidelines for outcome selection.
specific to phase II and III trials to ensure no key documents were missed. The keywords used were: outcome$.ti and phase II or phase III or clinical trial$.ti. The search was limited to English language. The identified evidence was synthesized.

4.5.2 Development of draft tool
The synthesis of the scoping review informed a draft of minimum standards and recommendations for the selection, measurement and reporting of outcomes in phase II and III paediatric drug trials: the preliminary checklist. Before this preliminary checklist is applied to an empirical evidence gathering in actual phase II and III drug trials, the applicability of this checklist was appraised by stakeholders from the pharmaceutical industry, and academic clinical trialists.

4.5.3 Qualitative feedback from industry partners
Following the development of the preliminary checklist in December 2014, we collaborated with three industry partners that have an established Pediatric Team (Novartis, Pfizer and Janssen) to review and road test this checklist to evaluate its relevance for designing early and pivotal drug studies. The respondents were requested to evaluate each item within the tool against existing phase II or phase III pediatric protocol within their respective pediatric working groups, and provide confidential feedback. Arrangements were made to meet with the industry partner through teleconference to further discuss their written feedback. The industry partners were asked to reflect on the following questions (Appendix III):
1. Is this a useful tool to apply in outcome selection in phase IIb and phase III drug trials?
2. Does it address all the relevant issues that go into outcome selection in phase IIb and phase III?
3. Would this tool prevent forgetting to address relevant questions when drafting a protocol?
4. Is it user friendly?
5. How much time does it take to complete one whole form per outcome?
6. Will it help or hinder the protocol-writing-process?

4.5.4 Survey with academic clinical trialists
The aim of this survey was to assess the usefulness and completeness of the tool and to identify knowledge gaps. An anonymous survey was conducted with the clinical trialists (n=11) who are involved in pediatric clinical trials and work with industry partners (Appendix IV). The clinical trialists were identified through an Innovative Clinical Trial Design course 2015, given at the Hospital for Sick Children. This course offers advanced level training in clinical trial design. The preliminary checklist was administered during the three day intensive course, following an extended interactive teaching session on outcome selection. The participants were asked to rate each item against four possible response options, namely:
- I would always address this item, so a checklist would make no difference
- I may sometimes address this item, but a checklist would help to remind me
- A checklist would make it much more likely for me to address this item
- I do not think this item is relevant, so a checklist would make no difference.

4.5.5 Final GRiP deliverable
The results from the scoping literature review and the feedback from both the industry partners and the clinical trialists were analysed and integrated into a final GRiP deliverable – an instrument referred from here onwards as InSPECT (Instrument for Selection of Pediatric Endpoints in Clinical Trials) – a checklist for selecting outcomes in paediatric phase II and III drug trial.
4.6 Results

4.6.1 Scoping literature review

A total of 17 relevant guidelines or recommendation documents were identified. The characteristics of the included reports / publications are provided in Table 1. Most of the identified guidelines were generic (i.e., not specifically developed for children) including the ones from the regulator, such as FDA and EMA. The ICH8 guideline was included to cover the generic aspects of clinical trials consideration. Paediatric specific guidelines on reporting of clinical trials protocols (CONSORT-C and SPIRIT-C) are currently under development.

No guidance regarding the choice of outcomes and outcome measurements in phase II and phase III of paediatric drug trials was found.

Table 1: Literature identified to support InSPECT development

<table>
<thead>
<tr>
<th>Guideline</th>
<th>Year of Publication</th>
<th>Publication Type</th>
<th>Consensus Achieved</th>
<th>Paediatric or Generic</th>
<th>Addressing Phase II and Phase III trials outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDA</td>
<td>2013 (check)</td>
<td>Regulator guideline</td>
<td>Method not reported</td>
<td>Generic</td>
<td>No</td>
</tr>
<tr>
<td>CONSORT-C</td>
<td>2015</td>
<td>Reporting guideline</td>
<td>Yes, through Delphi survey and Consensus meeting of stakeholders</td>
<td>Paediatric</td>
<td>No</td>
</tr>
<tr>
<td>SPIRIT-C</td>
<td>2015</td>
<td>Reporting guideline</td>
<td>Yes, through Delphi survey and Consensus meeting of stakeholders</td>
<td>Paediatric</td>
<td>No</td>
</tr>
<tr>
<td>SPIRIT</td>
<td>2013</td>
<td>Reporting guideline</td>
<td>Yes, through Delphi survey and Consensus meeting of stakeholders</td>
<td>Generic</td>
<td>No</td>
</tr>
<tr>
<td>Coster</td>
<td>2013</td>
<td>Review article</td>
<td>NA</td>
<td>Generic</td>
<td>No</td>
</tr>
<tr>
<td>OMERACT</td>
<td>2013</td>
<td>Reporting Guideline</td>
<td>Yes, through Consensus meeting of stakeholders</td>
<td>Generic</td>
<td>No</td>
</tr>
<tr>
<td>COSMIN</td>
<td>2012</td>
<td>Reporting guideline</td>
<td>Yes, through Delphi survey and Consensus meeting of stakeholders</td>
<td>Generic</td>
<td>No</td>
</tr>
<tr>
<td>PCORI</td>
<td>2012</td>
<td>Reporting guideline</td>
<td>Yes, through discussion with an expert panel</td>
<td>Generic</td>
<td>No</td>
</tr>
<tr>
<td>Williamson</td>
<td>2012</td>
<td>Evidence synthesis</td>
<td>NA</td>
<td>Generic</td>
<td>No</td>
</tr>
<tr>
<td>STAR</td>
<td>2012</td>
<td>Evidence synthesis</td>
<td>NA</td>
<td>Paediatric</td>
<td>No</td>
</tr>
<tr>
<td>GRADE</td>
<td>2011</td>
<td>Reporting guideline</td>
<td>Method not reported</td>
<td>Generic</td>
<td>No</td>
</tr>
<tr>
<td>McKenna</td>
<td>2011</td>
<td>Evidence synthesis</td>
<td>NA</td>
<td>Generic</td>
<td>No</td>
</tr>
<tr>
<td>CONSORT</td>
<td>2010</td>
<td>Reporting guideline</td>
<td>Yes, through Delphi survey and Consensus meeting of stakeholders</td>
<td>Generic</td>
<td>No</td>
</tr>
</tbody>
</table>
4.6.2 Themes identified for selection, measurement and reporting of outcomes

Several sources pointed out that, unfortunately, the quality of the description of outcomes in current paediatric clinical trial protocols and reports is remarkably poor; a systematic review of trials on neonatology demonstrated that these trials often fail to specify a primary outcomes [8]. Due to poor reporting, it is most often impossible to judge whether an in-depth analysis of possible outcomes was followed by a judicious selection process. Further, we found that, to date, very little work has been done on the guidance of appropriate outcome selection in paediatric clinical trials. Thus, the currently available literature does not provide a contemporary practical guide for outcome selection in phase II and III paediatric drug trials.

Given these findings, we proceeded to gather evidence from the scoping review and identified six key aspects that should be evaluated for selecting an appropriate outcome and reporting it in phase II and III paediatric drug trials. These key themes are reflected in a new 14 items InSPECT checklist (Appendix I), that includes the source of each recommendation.

Why: Rationale for selecting the outcome

Coster et al. recommend that creating an explanatory model of the intervention can help guide researchers select the appropriate outcomes [10]. This allows the researchers to visualize the intervention’s mechanism of change and how the study participants are impacted. The causal-explanatory model will also help illuminate the measurement construct of the outcome. This information allows stakeholders including participants, caregivers such as parents, and trial personnel, to assess the scientific and ethical basis for the trial, i.e. whether whether the trial rationale matches the trial methods. Moreover, it allows information users to determine whether the primary outcomes that are to be assessed meet the objectives of the study [11].

What: Qualification - Relevance of the outcome for the target population

The outcomes in a clinical trial should be biologically credible, clinically important, feasible, discriminative and reliable to the intervention [9]. Key aspects of qualification of an outcome include a justification of the clinical relevance of the outcome to the child, its growth and development, and to its family. Qualification speaks to how the outcome reflects real clinical benefit of the interventions, how the outcome maximizes translatability of the study’s findings into clinical practice, and whether decisions will be based on that outcome’s change (improvement) as ascribed to the intervention under study. Unlike in a phase III trial, in which the primary outcome often represent direct ‘clinical’ benefit to the patient (e.g. progression-free survival years), in phase II trial, the selection of primary outcome is often based on representing disease activity (e.g. tumour growth) [12].

Several frameworks that aim to enhance the selection of qualified outcomes in drug trials were identified. OMERACT identified four core areas that should be considered when developing the set of outcomes to be measured in a (phase III) trial, namely pathophysiological manifestation, life impact, health resource utilization and death [9].
However, further evidence is needed with regard to the applicability of these core areas in early phase exploratory pediatric drug trials.

**How: Determination of how the outcome will be measured**

This item refers to the measurement instrument that will be used for assessing the outcome of interest. An outcome instrument refers to a scale, scoring system, questionnaire, clinical test or other tool used for measuring an outcome. Extensive methodological work has been done assessing measurement qualities of a clinical outcome [13]. The four key properties of outcome measures that are an integral part of an investigator's evaluation of appropriate measures include **validity**, **reliability**, **responsiveness** and **feasibility** [9, 13].

**Reliability** is defined as a degree to which the measurement instrument is free from random measurement error. Different aspects of reliability include repeating measurements over time (test-retest), by different person on the same occasion (inter-rater); or by the same person on different occasion (intra-rater) [14].

**Validity** of a measure is how well it measures what it is supposed to measure. The outcome measurement instruments selected for a paediatric clinical trial should be relevant to, and have been validated for, the selected age group(s). Some measurement instruments that are valid for some paediatric and developmental stages may not be valid for others. Measures that have been validated for adults or specific age groups which have been newly modified to be used in another age group(s) need to be re-validated in those age group(s) as well.

The new SPIRIT-C reporting standard for trial protocols [15] recommends that the developers of paediatric clinical trial protocols use COSMIN (COConsensus-based Standards for the selection of health Measurement Instruments) to help evaluate the validity of measurement instruments for paediatric clinical trials [13].

Three key aspects of validity include **face validity** (the degree to which a measure is assessing what it is intended to measure), **content validity** (the extent to which a measure accurately and comprehensively measures what it is intended to measure), and **construct validity** (the degree to which an instrument is consistent with hypotheses, for instance with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups). Sinha et al. suggest that instruments that will measure the outcomes in a trial must be demonstrably valid or to be validated for the age specific subgroup in the study, the disease under investigation, and the setting of the trial [6]. Validation studies are strongly recommended if the validity of instruments has not been adequately tested. Moreover, references to support the validity of the tool, instrument, or method should be provided in the clinical trial protocol as well as trial reports.

**Responsiveness** is an ability of an instrument (methods, questionnaire etc.) to measure a significant change in disease activity over time. Instruments must be responsive and of sufficient specificity and sensitivity to allow detection of clinically significant treatment effects within the selected age groups.

**Feasibility** is a degree to which an assessment tool can be applied easily, given constraints of time, money, and interpretability. Some data collection methods that are relevant and feasible for certain age groups may not be for other age groups. One aspect of feasibility is acceptability: the burden imposed by the application of the outcome measure on the paediatric patient or parent should be appropriate. E.g., children with greater levels of illness or disability are less able to complete lengthy instruments. In some cases, the content of an instrument may be upsetting or otherwise unacceptable to respondents such as parents.
other cases, an instrument may be cognitively too challenging i.e., written at a reading level that is above that of the targeted paediatric population.

Another important aspect is to describe and justify the **minimally important difference** that an outcome measure is expected to identify. Clinically meaningful change in an outcome measure occurs when a child, clinician or caregiver perceives an improvement or worsening in his or her well-being. The *minimal* clinically important difference (MCID) of a measure is a numerical value indicating the smallest clinically meaningful change in an outcome measure that can be detected. It is important to describe a-priori the change in outcome measure expected from baseline and what this change would mean in terms of success or failure of a trial. Previous studies have identified MCID in Pediatric Quality of Life Inventory, separately for Children and Adolescent with different health conditions [16].

All these properties are specific to the population (i.e. the targeted paediatric age groups) and setting in which the measures are used; they affect the degree of measurement error or misclassification that an outcome measure is subject to. There may be an amalgamation of more than one outcome within one outcome domain, such as a score based on a variety of symptoms. More than one outcome measure may be possible to use to represent change in an outcome and is commonly referred to as a composite outcome and in such a case, the qualification (with regard to validity, responsiveness and feasibility) of each outcome within the composite endpoint should be determined.

**Who: Determination of the most appropriate source of information**

The fourth question that must be asked is “who”, referring to the source of the outcome information. There are usually four key sources for clinical outcomes assessment: *researcher, health professional, patient (child), or care-giver such as parents.* In case of patient (child) or care-giver, it is particularly important to ensure that the necessary sensory, literacy, cognitive, physical, and communication abilities of the outcome assessor have been considered [10]. Reporting the qualifications of outcome assessors in relation to the application of the instrument and training given to assessors would improve the transparency of the study, enabling bias assessment. Furthermore, training of assessors for use of a particular instrument enhances inter-rater reliability and, consequently, validity of results obtained from the said instrument [17]. Study validity has been shown to be higher when the same assessor took measures from all participants in a trial compared to when different assessors conducted the evaluations [18]. In case of multicenter studies, however, consideration should be given to assess outcome by the same assessor within a centre. Reporting of this item will allow readers to gauge the study’s validity.

Another consideration on the source of information is an appropriate masking of outcome assessors to the allocation of the interventions to prevent or reduce information bias. A recent study found that 19% of the pediatric clinical trial had a high or unclear risk of bias for masking [19].

**Where: Facilities at the location where the outcome is measured**

Outcomes in drug trials can be measured in the clinic, at home, at school, in a community facility, or any other place. The exact location may be another source of variability as the environment can affect all aspects of outcome measurement [20]. Therefore, description of where the outcome measure will be taken will allow for the evaluation of whether environmental confounders might impact the findings. Furthermore it will also allow a reviewer to assess whether necessary facilities required for a proper outcome assessment are available at the time and place of measurement.
**When: Specification of time points of measuring outcome**

Responsiveness to the intervention and variation within the target population (dynamic range) of the outcome measure is a vital consideration for outcome selection. Affirmation that the selected instrument operates during the time between implementation of intervention and measurement of outcome, determination of the optimal assessment point(s), and considerations given to the trajectory of change in the study population’s condition should be done [9, 10].

Temporal considerations are an important aspect of outcome selection and therefore require investigators to assess outcomes at appropriate time intervals. When selecting time point(s) of assessment it must be ensured that the time over which the chosen outcome instrument is responsive falls between time period of exposure of interventions to participants and outcome measurement. Description of these items allows readers to assess potential bias as well [10].

**Outcome reporting and Core Outcome Sets**

Outcomes should be described in great detail and strict definitions should be used. When designing a clinical trial examining efficacy and effectiveness of therapies, researchers should consider all aspects of the effects of interventions, population and intervention to identify primary and secondary outcomes that are important when making decisions about health care practice [6]. The “primary outcome” is the outcome of most interest and is the one used to calculate the sample size of the trial. Data on “secondary outcomes” are used to evaluate additional effects of the intervention. Selective reporting of outcomes, known as outcome reporting bias occurs when multiple outcomes are measured, but only selected ones are reported usually those with statistically significant results. The risk of reporting bias can be reduced by pre-specifying primary and secondary outcomes. Furthermore, follow-up period should be sufficient to observe hypothesized effects of treatment on primary and secondary outcomes.

In addition to the description of hierarchy of outcomes in a trial protocol, (e.g., primary or secondary endpoints), there is a recent emphasis on the development of Core Outcome Sets (COS) for a health condition. COS is an agreed standardized collection of outcomes for a specific health condition or intervention that can combat problems caused by inattention in selecting outcomes in clinical trials. A primary outcome of a trial should be one of the pre-defined / agreed COS. Williamson et al. argues that use of COS would reduce heterogeneity in reported outcomes between trials and reduce outcome reporting bias. Recommended methods for the development of COS have been proposed [5]. OMERACT identified four core areas that should be considered when developing the COS, namely pathophysiological manifestation, life impact, health resource utilization and death [9]. However further evidence is needed with regard to the applicability of the core areas in pediatric specific trials.

**4.6.3 Qualitative feedback from industry partners**

Three industry partners, the heads of the paediatric task force within their company with extensive experience in phase II and III pediatric drug trials, provided written feedback. Two industry partners provided further feedback through a telephone conference of one hour. The feedback on each individual item’s wording was incorporated in the updated checklist. Few excerpts from their feedback on the overall checklist are given below:

“The various items for consideration are well-thought out. The justification for these items followed a thorough review of the literature and FDA’s guidance. However, it’s not clear to me at what point in the protocol development this tool is intended to be utilized? For example, would this be utilized as the clinical trialist is developing the
study outline or would it be used after the protocol has been completed and is undergoing an “external” review process?” (Industry partner 1)

“I think it (the checklist) has the potential of being very valuable in guiding thinking as a study is being designed. It certainly seems to work well on traditional studies like an antihypertensive trial or antibiotic trial. But traditional studies are being done less and less as we recognize that we have had many trial failures for many reasons including working with the wrong endpoints. I think some modifications may need to be done to accommodate some new designs like how to handle adaptive designs with interim analyses, seamless phase 2/3 or 1/3 studies, extrapolation programs, or ranked analyses” (Industry Partner 2)

4.6.4 Survey of clinical trialists
The characteristics of the survey respondents are provided in Table 2. The respondents have varying level of research experience and role. The respondents were primarily involved in Phase I-III trials.

<table>
<thead>
<tr>
<th>Years of experience</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1 year</td>
<td>3</td>
</tr>
<tr>
<td>2-5 years</td>
<td>3</td>
</tr>
<tr>
<td>5-10 years</td>
<td>3</td>
</tr>
<tr>
<td>&gt;10 years</td>
<td>2</td>
</tr>
</tbody>
</table>

*Primary role in clinical trial*

<table>
<thead>
<tr>
<th>Role</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regulator</td>
<td>1</td>
</tr>
<tr>
<td>Academic Clinical trialist</td>
<td>7</td>
</tr>
<tr>
<td>Ethics committee member</td>
<td>1</td>
</tr>
<tr>
<td>Epidemiologist</td>
<td>1</td>
</tr>
<tr>
<td>DSMC</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experience with drug development phase*</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>3</td>
</tr>
<tr>
<td>II</td>
<td>2</td>
</tr>
<tr>
<td>III</td>
<td>2</td>
</tr>
<tr>
<td>IV</td>
<td>1</td>
</tr>
</tbody>
</table>

*Multiple responses possible

In Figure 1 – 14 the results of the survey are summarized. None of the items was identified as irrelevant. For most of the items, the draft InSPECT checklist was considered a helpful reminder and was thought to improve the likelihood of the item was being addressed.
Figure 1: Is there a well-specified explanatory model showing how the intervention links to the outcome?

Figure 2: Does the outcome match the objectives of the study?

Figure 3: Is the outcome qualified for the health condition and the pre-specified pediatric age group?
Figure 4: Have the most relevant dimensions or core areas of the outcome been specified

- I would always address this item, so a checklist would make no difference
- I may sometimes address this item, but a checklist would help to remind me
- A checklist would make it much more likely for me to address this item
- I do not think this item is relevant, so a checklist would make no difference

Figure 5: If this outcome classified as primary or secondary?
Provide justification for such classification

- I would always address this item, so a checklist would make no difference
- I may sometimes address this item, but a checklist would help to remind me
- A checklist would make it much more likely for me to address this item
- I do not think this item is relevant, so a checklist would make no difference

Figure 6: Has the tool, instrument or methods used to measure the outcome shown to be valid for children in the pre-specified age group, with the illness of interest and in the setting in which the trial is conducted

- I would always address this item, so a checklist would make no difference
- I may sometimes address this item, but a checklist would help to remind me
- A checklist would make it much more likely for me to address this item
- I do not think this item is relevant, so a checklist would make no difference
Figure 7: Would the burden imposed through the implementation of outcome measure on the pediatric patient be acceptable and feasible?

Figure 8: Who will be making the assessments?
Figure 9: Do the outcome assessors match the qualifications criteria of the instrument being considered?

- I would always address this item, so a checklist would make no difference
- I may sometimes address this item, but a checklist would help to remind me
- A checklist would make it much more likely for me to address this item
- I do not think this item is relevant, so a checklist would make no difference

Figure 10: Would training be required / provided to the outcome assessors for the purpose of data collection, administration procedures, completing process, handling of data?

- I would always address this item, so a checklist would make no difference
- I may sometimes address this item, but a checklist would help to remind me
- A checklist would make it much more likely for me to address this item
- I do not think this item is relevant, so a checklist would make no difference
I would always address this item, so a checklist would make no difference
I may sometimes address this item, but a checklist would help to remind me
A checklist would make it much more likely for me to address this item
I do not think this item is relevant, so a checklist would make no difference

Figure 11: Would the outcome assessors be available when the measurements are required (i.e., for all measurement points)

I would always address this item, so a checklist would make no difference
I may sometimes address this item, but a checklist would help to remind me
A checklist would make it much more likely for me to address this item
I do not think this item is relevant, so a checklist would make no difference

Figure 12: Would the location for outcome measurement be available for the time period of intervention?

I would always address this item, so a checklist would make no difference
I may sometimes address this item, but a checklist would help to remind me
A checklist would make it much more likely for me to address this item
I do not think this item is relevant, so a checklist would make no difference

Figure 13: Is the time point of outcome measurement justified?

I would always address this item, so a checklist would make no difference
I may sometimes address this item, but a checklist would help to remind me
A checklist would make it much more likely for me to address this item
I do not think this item is relevant, so a checklist would make no difference
4.6.5 Final GRiP deliverable
After incorporation of all relevant feedback, the adjusted tool underwent a final check in a consensus round among all GRiP partners. This approval process within GRiP was set in place to enhance the quality and relevance of the product, and to justify the GRiP branding of the final product. This also enhances the tool’s eventual implementation (benefitting from the reliable label of a EU research consortium) and at the same time contribute to GRiP visibility. The final GRiP tool: a checklist for outcome selection (InSPECT) is presented in Appendix 1. It is a “ready to use’ tool that aims to help trialists to address all six key aspects that should be evaluated for selecting, measuring reporting appropriate outcomes in phase II and III paediatric drug trials. Per outcome to be considered, InSPECT checks 14 items, providing a reference to the source of each consideration and recommendation.

4.7 Knowledge translation plan
A Knowledge Translation plan was developed upon consultation with Knowledge Translation specialists and GRiP WP1 leaders to bring the final GRiP deliverable to the attention of the end-users and stakeholders. This is addressed by the development of a training module in which paediatricians, medical sub-specialists, trialists, clinical research associates, ethics committee members and other relevant end-users can gain expertise on this topic. Moreover, it will be incorporated in the teaching material and lecture of module #4 for the GRiP Master Programme in Paediatric Medicines Development and Evaluation.’ For further knowledge transfer, the checklist will be published in a peer reviewed journal with pediatric trialists as readership. The final GRiP deliverable is available on the GRiP website, and downloadable for free use. Active participation of internal and external GRiP stakeholders is ongoing. This will ensure both internal and external validity, achieving a valid and sustainable Deliverable.

4.8 Limitations and indications for future research
4.8.1 Limitations
This tool was developed for a multidisciplinary, international audience, with the aim of presenting considerations of general interest, in order to encourage better research practice. However, though based on regulatory guidance an scientific literature, it does not have the status of a regulatory guideline. Producing such a guideline would have required formal consensus methods across institutions and jurisdictions, and the resources needed to engage key stakeholders in industry and academia, and decision makers at the regulatory agencies. To strengthen the external validity and up-to-dateness of the tool, the first draft
was enriched by suggestions representing the knowledge and experiences of experts within the GRiP consortium. Furthermore, industry partners and clinical trialists commented on whether the tool was feasible; this was largely the case and in a few cases the tool was further adjusted.

The trialists who provided feedback were partly identified through a convenience sampling of trialists enrolled in an advanced trial design course at the Hospital for Sick Children. Our small sample was therefore biased towards respondents with advanced experience in trial design. We did seek information on respondents’ basic characteristics such as a range of professional backgrounds and level of experience which indicate that their experience was diverse. Furthermore, our industry partners with extensive experience in pediatric phase II and III trials provided in-depth evaluation of the tool. External validity of the results could be enhanced by engaging trialists through (random) sampling of trial databases, e.g. Clinical trials.gov, to identify phase II and III pediatric clinical trials.

4.8.2 Future Research
Future steps could include the gathering of further empirical evidence for the utility of the tool. To this end, a set of recent, linked exploratory and confirmatory pediatric drug trial protocols could be examined using the checklist.

**Empirical Evidence Gathering**

**Hypothesis**
Outcomes that are used in both exploratory and confirmatory trials that support decisions about efficacy (for example marketing authorizations / licensing) are more likely to predict positive confirmatory trials if the exploratory trial contributes to the qualification and validation and of the outcomes used.

**Objectives**
1. to investigate the “reported rationale” for the selection of outcomes that are shared between exploratory and confirmatory phase of the study;
2. to explore which characteristics of outcome selection methods in exploratory studies might be predictive of the success or failure of subsequent confirmatory studies in children that use similar outcomes in exploratory and confirmatory trials.

**Methods**
One could develop case studies in therapeutic areas, e.g. oncology, neurology, rheumatology and neonatology, identify two pairs (one each of “success” and “failure”) of studies of exploratory and confirmatory paediatric drug trials. A study pair will be labelled as “success” if the results of both exploratory and confirmatory trials are positive. A study pair will be labelled as “failure” if the result of the exploratory trial are positive, however that of the confirmatory trials are negative.

These pairs of studies could be identified from the published literature, or as protocols in EMA Paediatric Investigation Plans (PIPs) or in EPAR, FDA, or by contacting industry or regulators. First, a confirmatory trial be identified in order to then identify the preceding exploratory study for inclusion.

**Inclusion / Exclusion Criteria:**
These studies should be
a. drug trials
b. focusing on paediatric population aged 0-18 years
c. in either exploratory or confirmatory phase

d. in one of the therapeutic areas: e.g. *oncology, neurology, rheumatology or neonatology*

e. describing the selection of primary outcomes (in enough detail)

f. only those pairs of studies using identical chemotherapy regimens will be included. Identical chemotherapy regimens are defined as those utilizing the same chemotherapeutic agents, the same administration schedule, and the same intended patient population. However, the study pair do not need to be by the same research team / same pharmaceutical industry.

Data on outcome selection and measurement should be extracted from both exploratory and confirmatory studies. Data on following aspects of the outcomes will be interesting: i) how an outcome was selected, ii) how an outcome qualified to be included in a trial, iii) what scales and tools were used to measure an outcome including identification of validated instruments, iv) how many of the InSPECT items were addressed.

The results regarding efficacy and safety of the exploratory and confirmatory phase could be compared to test the hypothesis. The results of the hypothesis testing will inform further steps.

**Checklist evaluation against AGREE II guideline**

**Objective**

To develop an evidence and consensus based checklist of minimal set of items to consider when selecting, measuring and reporting qualified and validated outcomes in exploratory and confirmatory phase of paediatric drug trials.

**Methods**

The results from the evaluation of existing guidance and empirical work on the critical analysis of study pairs could be used to develop a checklist of minimal set of items to consider when selecting, measuring and reporting validated outcomes that bridge between the design and conduct of exploratory and confirmatory paediatrics drug trial. Consensus on the checklist items will be achieved through a consensus meeting with an expert panel, who will appraise its quality and validity through the modified *Appraisal of Guidelines Research & Evaluation* (AGREE II) checklist [21]. The original AGREE II checklist is used to assess the quality of practice guidelines that evaluates the methodological rigor and transparency in which the guideline was developed. It is designed to assess new guidelines, existing guidelines and updates of existing guidelines, and can be applied in any disease area. AGREE II consists of 23 key items with six domains namely; i) Scope and purpose, ii) Stakeholder involvement, iii) Rigour of development, iv) Clarity and presentation, v) Applicability and, vi) Editorial independence. We may exclude 6 of the 23 items because they are only applicable to clinical practice guidelines and not to guidelines related to clinical research in children. The user will score each item with “yes”, “no” or “unknown”. A comments section will be provided for each item. Since there is no validated cut off score for the instrument, scores from the adapted AGREE II statement will be analyzed quantitatively, while comments will be qualitatively assessed using a thematic approach. As the quality and validity is appraised, then discussed, a first draft of an evidence and consensus based checklist of minimal set of items for selecting, measuring and reporting qualified and validated outcomes in exploratory and confirmatory phase of paediatric drug trials could be agreed on. An embedded and end of consensus meeting Knowledge to Action process will be used to further improve, disseminate and implement the checklist.
5 Document History

<table>
<thead>
<tr>
<th>Date</th>
<th>Author</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-Jun-2013</td>
<td>Winnie Chan, Martin Offringa</td>
<td>First draft</td>
</tr>
<tr>
<td>30-Nov-2013</td>
<td>Thivia Jegathesan, Martin Offringa</td>
<td>Second draft</td>
</tr>
<tr>
<td>31 July 2014</td>
<td>Mufiza Kapadia</td>
<td>3rd Draft</td>
</tr>
<tr>
<td>12 December 2014</td>
<td>Mufiza Kapadia, Praveen Thurairajah, Martin Offringa</td>
<td>4th Draft</td>
</tr>
<tr>
<td>15 January 2015</td>
<td>NA</td>
<td>Methods were approved by GRiP executive committee</td>
</tr>
<tr>
<td>31 March 2015</td>
<td>Mufiza Kapadia</td>
<td>Feedback incorporated from Industry partners</td>
</tr>
<tr>
<td>30 April 2015</td>
<td>Mufiza Kapadia</td>
<td>Feedback incorporated from Innovative Trial Design Course participants</td>
</tr>
<tr>
<td>31 August 2015</td>
<td>Mufiza Kapadia, Praveen Thurairajah, Martin Offringa</td>
<td>Final deliverable prepared</td>
</tr>
</tbody>
</table>

6 References


7 Appendices

7.1 Appendix I: InSPECT (Instrument for Selection of Pediatric Endpoints in Clinical Trials – Exploratory and Confirmatory Trials)

This checklist is to be completed, separately for each outcome. Instructions for completing this checklist are provided in green. For details on the method used to develop this guideline, refer to accompanying reference document.

**Q.1. What is the outcome domain (concept) under consideration?** Enter text here

**Q.2. What is the age group of the population?** *(Check all boxes that apply)*

- Neonate: Birth - 27 Days
- Infant: 28d - 12mo
- Toddler: 13mo - 2y
- Early Childhood: 2-5y
- Middle Childhood: 6-11y
- Early adolescence: 12-18y

**Q.3. Phase of Drug Trial:** ☐ Exploratory  ☐ Confirmatory  ☐ Other *(See Glossary(Appendix II) for the definition of terms)*
<table>
<thead>
<tr>
<th>Theme</th>
<th>Item No.</th>
<th>Item</th>
<th>Does the protocol describe this item?</th>
<th>Justification for judgment</th>
<th>Evaluation of an item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>This column is to be completed by the clinical trialist</td>
<td>This column is to be completed by the clinical trialist</td>
<td>This column is to be completed by Reviewer(s) only</td>
</tr>
<tr>
<td>Why: Determination of the rationale for selecting the outcome</td>
<td>1</td>
<td>Is there a wellspecified explanatory model or evidence from preclinical studies showing how the intervention links to the outcome in the pediatric population? [10, 22, 23]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Does the outcome match the objectives of the study? [23-25]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What: Determination of Outcome’s qualification for the health condition and population</td>
<td>3</td>
<td>Is the outcome qualified for the health condition in the pre-specified paediatric age group? [10, 26-28]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>If the outcome classified as primary or secondary endpoint? Provide justification for such classification [5, 6, 22, 29-31]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Which core area does the outcome under consideration represents? [10] (Use OMERACT Filter 2.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How: Determination of how the outcome will be measured</td>
<td>6</td>
<td>Has the tool, instrument, or method used to measure the outcome shown to be valid for children in the pre-specified age group, with the illness of interest, and in the setting in which the trial is conducted? If so, how? [5, 10, 29, 32] (Use COSMIN checklist)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Is the outcome responsive to meaningful change (e.g., minimally important difference has been established) in the pediatric population? [10] (Use COSMIN checklist)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.1</td>
<td>Is the expected magnitude of change in the value of outcome measure described (2, 14)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Would the burden imposed through the implementation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
of outcome measure on the paediatric patient or parent be acceptable and **feasible?** [33]

<table>
<thead>
<tr>
<th>Who: Determination of the most appropriate source of information</th>
<th>9</th>
<th>Who will be making the assessments (e.g., professional, caregiver)?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>Would training be required for the outcome assessors to implement the outcome instrument, and handle/storage of data? [10]</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>Would the outcome assessors be available when the measurements are required (i.e., for all measurement points, specification of which assessor be available at different time points of measurement)? [10]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Where: Determination of the location where the outcome measurement will occur</th>
<th>12</th>
<th>Is the location for outcome measurement suitable and all the necessary facilities available for the time period of intervention in the paediatric population? [32, 34]</th>
</tr>
</thead>
</table>

| When: Determination of timings of measuring outcome | 13 | Is the time point(s) of outcome measurement justified? |

**Source of recommendation**


### 7.2 Appendix II: Glossary of terms

<table>
<thead>
<tr>
<th>Terminology</th>
<th>Description</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome domain [5]</td>
<td>A relatively broad aspect of the effect of illness on a child, within which an improvement may occur in response to an intervention. In general these domains may not be directly measurable themselves, so outcomes are selected to assess change within them.</td>
<td>In clinical trials of children with asthma, outcome domains may include lung function, health care utilisation and symptom control</td>
</tr>
</tbody>
</table>
| Outcome [5]                          | A measurable variable within an outcome domain. The outcome can be measured at a variety of time points, which must be clearly stated by authors of clinical trials. The term ‘outcome’ and ‘endpoint’ are often used interchangeably in clinical trials, although endpoint is most frequently used in Phase II (efficacy) studies while outcome is most frequently used in Phase III / effectiveness studies. | (1) Absolute FEV1 expressed as change from baseline  
(2) Number of admissions to hospital within a six month period  
(3) Time to first seizure after starting an antiepileptic intervention |
| Outcome measurement [5]              | A scale, scoring system, questionnaire, clinical test or other tool used for measuring an outcome. Scales, scoring systems and questionnaires which rely on personal opinion / subjective interpretation pose a much higher level of difficulty to validate, compared to a test which relies on an objective laboratory analysis e.g. hemoglobin. Such scales, scoring systems and questionnaires are also often very difficult to be used for powering a study. There may be an amalgamation of more than one outcome within an outcome domain, such as a score based on a variety of symptoms. More than one outcome measure may be possible to use to represent change in an outcome and is commonly referred to as a composite outcome. | (1) Lung Function Test  
(2) The Paediatric Asthma Quality of Life Questionnaire  
(3) Bristol Stool Chart |
| Phase II / Exploratory Trial (EMA 1998) | Phase II is usually considered to start with the initiation of studies in which the primary objective is to explore therapeutic efficacy in patients. These trials are typically conducted on participants who are selected by relatively narrow criteria. Determination of dose(s) and regimen for Phase III trials is an important goal of Phase II trials. Additional objectives include evaluation of potential study endpoints, therapeutic regimens (including concomitant medications) and target populations (e.g. mild versus severe disease) for further study in Phase II or III. | A Phase II study of the safety and efficacy of teriflunomide in multiple sclerosis with relapses by O’Connor et al. [35] |
| Phase III / Confirmatory Trial (EMA 1998) | Phase III usually is considered to begin with the initiation of studies in which the primary objective is to demonstrate, or confirm therapeutic benefit. Further exploration of dose-response relationship; | Randomized phase III trial of pemetrexed versus docetaxel in patients with non-small-
and drug’s use in wider populations, in different stages of disease, or in combination with another drug are also done during these trials. Phase III trials complete the information needed to support adequate instructions for use of the drug, and thus provide adequate basis for marketing approval.

<table>
<thead>
<tr>
<th>Biomarker</th>
<th>A medical sign that can be consistently measured accurately to objectively assess the observed medical state from outside the patient [37]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surrogate outcome</td>
<td>A surrogate outcome of a clinical trial is a laboratory measurement or a physical sign used as a substitute for a clinically meaningful outcome that measures directly how a patient feels, functions or survives [5]</td>
</tr>
<tr>
<td>Composite outcome</td>
<td>A composite outcome in an outcome that consists of two or more different outcomes. A composite outcome is recognized when a person has experienced any one of the outcomes specified in the measure [39]. Construction of composite outcomes requires careful consideration and validation compared to when an outcome is used in isolation, and greatly increases the statistical challenges of powering a trial.</td>
</tr>
<tr>
<td>Core Outcome Set</td>
<td>It is an agreed standardized collection of outcomes for a specific health condition or intervention that can combat problems caused by inattention in selecting outcomes in clinical trials.</td>
</tr>
<tr>
<td>Pharmacokinetics</td>
<td>What the body does to a medicine. Description of absorption, distribution, metabolism, excretion of the medicine. [11]</td>
</tr>
<tr>
<td>Pharmacodynamics</td>
<td>What a medicine does to the body. Description of the pharmacological effects of a medicine on the various target organs. [11]</td>
</tr>
</tbody>
</table>
7.3 Appendix III: Invitation to industry partners

Dear XXX,

Toronto, 19 December 2014

I am pleased to invite you to contribute in developing an evidence based tool for the selection of outcomes in paediatric phase II and III drug trials.

What is GRiP? GRiP – Global Research in Paediatrics – Network of Excellence (GRiP) is an EU-funded project which aims to stimulate and facilitate the development and safe use of medicines in children. GRiP aims to create international consensus on standards, methods and “interoperability tools” for paediatric drug development. As one of 21 partners in GRiP, my research group works on methods and tools that make these studies more impactful.

What is InSPECT? GRiP asks us to develop a tool for outcome selection (qualification), consider measurement issues (validity, feasibility, responsiveness), and the utility of outcomes across phase II and III trials. InSPECT (Instrument for Selection of Pediatric Endpoints in Clinical Trials – Phase II and III) is a first draft of such a tool. Since there is no existing (regulatory) guidance on outcome selection in early pediatric drug trials, our preliminary tool (attached) is based on generic (regulatory) guidance.

How can you contribute?

IMPORTANT: We do not seek any confidential information about any particular protocol from your company.

We are currently collaborating with industry sponsors, regulators, and academics to review and subsequently road-test this tool. What is its relevance for designing early and pivotal drug studies? Your company’s paediatric drug development group can contribute to this critical methodological work by testing this tool against an efficacy’ outcome’in a phase Ib or Phase III pediatric drug protocol.

Please discuss and answer the following questions
1. Is this a useful tool to apply in outcome selection in phase Ib and phase II drug trials?
2. Does it address all the relevant issues that go into outcome selection in phase Ib and phase II?
3. Would this tool prevent forgetting to address relevant questions when drafting a protocol?
4. Is it user friendly?
5. How much time does it take to complete one whole form per outcome?
6. Will it help or hinder the protocol—writing process?

Please discuss, test, and prepare feedback by 20 February 2015. We shall arrange a teleconference with your team. Your valuable input would help us refine the tool, make it meaningful, before we proceed and work with regulators, and academics.

Sincerely,

Martin Offringa

---

Peter Gilgan Centre for Research and Learning
Room 11.9.706
666 Bay Street
Toronto, Ontario
Canada MSG 0A4
www.sickkids.ca

Andrea Verduge
Administrative Assistant
416–813–7654 ext. 306459
andrea.verduge@sickkids.ca

Lauren Kelly, PhD
Mufita Kapadia, MD PhD
Winnie Chan, MPH
Allison Needham, MSc
April O’Byrne – Shenin, MSc
Pravleen Tharaqah, MSc

Martin Offringa, MD PhD
Staff Neonatologist
Department of Neonatology
Professor of Pediatrics
Faculty of Medicine IHIPME
University of Toronto

Senior Scientist and
Program Head
Child Health Evaluative Sciences
CHRS
Hospital for Sick Children Research Institute
416–813–7654 ext. 328830
martin.offringa@sickkids.ca

EnBICH
Enhance Research in International Child Health

---

GRiP-D3.1-Evidence- and consensus-based guidance for the design, conduct and reporting of paediatric CTs (4)
7.4 Appendix IV: Survey – clinical trialists

InSPECT (Instrument for Selection of Pediatric Endpoints in Clinical Trials – Phase II and III)

Thank you for your willingness to complete this survey regarding InSPECT (Instrument for Selection of Pediatric Endpoints in Clinical Trials!)

Key points

- **InSPECT** is a reporting tool developed to **improve the selection and measurement of outcomes** in paediatric phase II and III drug trials. This preliminary tool is based on existing (regulatory) guidelines on outcome selection. It was developed for Global Research in Paediatrics – Network of Excellence (GRiP), an EU-funded project which aims to stimulate and facilitate the development and safe use of medicines in children.

- **Your contribution** - We are collaborating with clinical trialists, industry, regulators, sponsors and academics to road test this tool and evaluate its relevance for designing clinical trials. Please be aware that we do not intend to seek any confidential information from you about the protocols / RCT you are involved in.

The survey **takes approximately 10 minutes** to complete.

Should you experience any difficulty interpreting an item or have any questions, please do not hesitate to contact Mufiza Kapadia, via email Mufiza.kapadia@sickkids.ca

Thank you for completing the survey!

Dr. Martin Offringa and Dr. Mufiza Kapadia, on behalf of GRiP
Survey – End-User’s experience
InSPECT (Instrument for Selection of Pediatric Endpoints in Clinical Trials – Phase II and III)

Q.1. Please estimate how many years of experience you have in designing or conducting clinical trials in pediatrics?
☐ <1 Year ☐ 2-5 Years ☐ 5-10 Years ☐ >10 years

Q.2. Please describe your training related to clinical trials design or conduct

Q.3 Please describe your primary role in clinical trials
☐ Regulator ☐ Clinical Trialists ☐ Academic ☐ Ethic Committee Member
☐ Other, Specify………………………………………………

Q.4 Which Phase of Drug development / Trial are you involved in?
☐ Phase II ☐ Phase III ☐ Phase III ☐ Phase IV

Next page…
Q.4. Consider for each of the 12 items whether you would address that item (always / sometimes, regardless of its inclusion in a reporting checklist), or whether you think having it in a reporting checklist would be a helpful reminder. Use X to make your choice.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Item No.</th>
<th>Item</th>
<th>I would always address this item, so a checklist would make no difference</th>
<th>I may sometimes address this item, but a checklist would help to remind me</th>
<th>A checklist would make it much more likely for me to address this item</th>
<th>I do not think this item is relevant, so a checklist would make no difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why: Determination of the rationale for selecting the outcome</td>
<td>1</td>
<td>Is there a well-specified explanatory model showing how the intervention links to the outcome?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Does the outcome match the objectives of the study?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What: Determination of Outcome’s qualification for the health condition and population</td>
<td>3</td>
<td>Is the outcome qualified for the health condition and the pre-specified pediatric age group?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Have the most relevant dimension(s) or core area(s) of the outcome been specified clearly?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>If the outcome classified as primary or secondary? Provide justification for such classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How: Determination of how the outcome will be measured</td>
<td>6</td>
<td>Has the tool, instrument, or method used to measure the outcome shown to be valid for children in the pre-specified age group, with the illness of interest, and in the setting in which the trial is conducted? If so, how?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Would the burden imposed through the implementation of outcome measure on the pediatric patient be acceptable and feasible?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Who: Determination of the most appropriate source of information</td>
<td>8</td>
<td>Who will be making the assessments?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>Do the outcome assessors (e.g., professional, caregiver) match the qualifications criteria of the instrument being considered?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.1</td>
<td>Would training be required / provided to the outcome assessors for the purpose of data collection, administration procedures, completing process, handling/storage of data?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.2</td>
<td>Would the outcome assessors be available when the measurements are required (i.e., for all measurement points)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Where: Determination of the location where the outcome measurement will occur</td>
<td>10</td>
<td>Would the necessary facilities for outcome measurement be available for the time period of intervention, pediatric population and location of measurement?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When: Determination of</td>
<td>11</td>
<td>Is the time point(s) of outcome measurement justified? i.e., the outcome sensitive to</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Responsiveness of measuring outcome</td>
<td>meaningful change (e.g., minimally important difference has been established) in the pediatric population?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Does the expected change/time of outcome measure (e.g., change from baseline, final value, time to event, between time point X and Y) described?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>