University of Alberta

Methodological Issues in Randomized Trials of Pediatric Acute Diarrhea: Evaluating Probiotics and the Need for Standardized Definitions and Valid Outcome Measures

by

Bradley C. Johnston

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Experimental Medicine

Department of Medicine, Faculty of Medicine & Dentistry

Fall, 2009
Edmonton, Alberta

EXAMINING COMMITTEE

Dr. Sunita Vohra, Department of Pediatrics & School of Public Health, University of Alberta

Dr. Ross T. Tsuyuki, Department of Medicine & School of Public Health, University of Alberta

Dr. W. Todd Rogers, Department of Educational Psychology, University of Alberta

Dr. Richard N. Fedorak, Department of Medicine, University of Alberta

Dr. Hien Q. Huynh, Department of Pediatrics, University of Alberta

Dr. Lynne V. McFarland, Department of Health Services Research and Development, University of Washington

ABSTRACT

Background: In a 2006 WHO report, diarrheal diseases ranked second among conditions afflicting children. Pediatric acute diarrhea, although most often the result of a gastrointestinal infection, can also occur as a result of antibiotic exposure. This is often referred to as antibiotic-associated diarrhea (AAD). Previous research suggests that probiotics may be effective in the treatment or prevention of various types of PAD.

Methods: The first study involved a systematic review and meta-analysis of RCTs involving probiotics as an adjunct to antibiotics for preventing AAD in children. The second study was a systematic review of definitions and primary outcome measures employed in RCTs of PAD. The third study used a modified Delphi consensus procedure to develop a new instrument for evaluating the severity of PAD. The study involved steering committee discussions (phase 1) and two electronic surveys (phase 2 and 3) of leading experts in measurement and clinical gastroenterology.

Results: The per protocol meta-analysis of ten RCTs significantly favored probiotics to prevent the incidence of diarrhea (NNT = 10). However, this effect did not withstand ITT analysis and among included trials there was considerable inconsistency regarding definitions for the reviews primary outcome measure, the incidence of diarrhea. Study two identified 121 RCTs that reported 62 unique definitions of diarrhea, 64 unique definitions of diarrhea resolution and 62 unique primary outcome measures. Thirty-one trials used grading systems to support outcome evaluation. However, none of the trials (or their citations) reported evidence of their validation. In study three experts agreed on the inclusion of five attributes containing 13 items. Attributes proposed for the IPADDS

include: *Diarrhea Frequency* and *Duration, Vomiting Frequency* and *Duration, Fever*, *Restrictions in Normal Daily Activities* and *Dehydration*.

Conclusion: It is premature to draw a valid conclusion about the efficacy of probiotics for pediatric AAD. Definitions of diarrhea and primary outcome measures in RCTs of PAD are heterogeneous and lack evidence of validity. The third study represents content validity evidence for IPADDS. A numerical scoring system needs to be added and further empirical evidence of reliability and validity are required.

ACKNOWLEDGEMENTS

Committee Members

Sunita Vohra, thank you for supporting me, involving me in research projects outside of my dissertation and allowing me to be self-directed in my questions, writing and presentation of my thesis work. Your feedback was always timely, well-pointed and meticulous. *Ross Tsuyuki,* thank you for the all your support. Your advice as well as the learning opportunities you provided regarding the presentation of statistical concepts was excellent and very much appreciated. I always enjoyed working with your staff, especially Lily Yushko. Her generous and careful work in the formating and administration of the electronic survey for my final project was a testament to your research group. *Richard Fedorak,* thank you for participating in my committee. Although we did not spend a lot of time together you were always supportive and the discussions we did have were very helpful and important to my progress. It was always a pleasure to interact with you. *Todd Rogers,* you were a committee member and a friend. Although you were late addition to my committee, you were instrumental in my final project. Many thanks for your enthusiasm and endless support. Your door was always open.

Family

Mom & Dad, although you sometimes had a hard time understanding my seemingly endless pursuit of school - thank you for always encouraging me to follow my own path. You taught me the value of hard work and to "do onto others as you would wish them do onto you".

Friends

Adesiji, throughout my PhD journey your diplomactic advice was always influential. I was truly lucky to have had you as a roommate and as a friend. *Bruno,* your love for finding the comedy in every situation and passion for research always made for a productive and enjoyable friendship. *CARE students (Denise, Larissa and Sana)*, it was a pleasure to have each of you to work with (and "conference" with). Thanks for your support during my candidacy and defense preparation. I hope we can continue to be

colleagues in the future. *Ed,* you always supported my research endeavors through frank advice, opening important doors for me, acting as a reference and encouraging me to think big. I admire your ability to be real, to laugh, and to not let conceptual barriers hold you back. Your academic achievements are truly inspiring. *Edgar & Na,* what can I say, you two were there for many important moments including my "practice" candidacy and "practice" defense and, of course, to celebrate afterwards. We also shared many wonderful mountain adventures, barbeques and early morning lap swimming followed by Brasilian 'petro' coffee (to prevent cognitive decline)! *Luciana,* you are the sister I never had – thanks for being my roommate, making me a part of your wonderful family and for always being there. *Shafeek & Deepika,* you guys are awesome friends. Shaf – your persistence in getting me away from the computer for soccer or hockey was always much needed. Deepika, you are too kind – the world needs more of you.

Funding

SickKids Foundation: The funding you provided throughout my PhD studies was paramount to my success. It was a true enabler allowing me to not only successfully pursue my PhD studies, but to also gain experience in leading and collaborating on numerous other research projects.

Evidence-Based Practice Centre, University of Alberta: Your early financial support was a key reason why I chose to come to the University of Alberta. Not only was the funding you provided early in my studies much appreciated, but also the learning opportunity as a research assistant with your center in the summers of 2004 and 2005 was an excellent introduction to the importance of a team in producing high quality systematic reviews.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AAD	Antibiotic-Associated Diarrhea
ADL	Activities of Daily Living
CFU	Colony Forming Units
CI	Confidence Interval
Co	Consistency
DD	Diarrhea Duration
DF	Diarrhea Frequency
Fe	Fever
IPADDS	International Pediatric Acute Diarrheal Diseases Scale
IPADS	International Pediatric Acute Diarrhea Scale
IPAGS	International Pediatric Acute Gastroenteritis Scale
JDM	Judge Median Difference
PAD	Pediatric Acute Diarrhea
PADD	Pediatric Acute Diarrheal Diseases
PAG	Pediatric Acute Gastroenteritis
Md	Median
NNT	Number Needed to Treat
OR	Odds Ratio
ORS	Oral Rehydration Solution
Pa	Pain
R	Range
RCT	Randomzed Controlled Trial
RR	Relative Risk
RD	Risk Difference
VD	Vomiting Duration
VF	Vomiting Frequency
WHO	World Health Organization

CHAPTER 1: Overview

1. BACKGROUND

1.1.1 Pediatric Acute Diarrhea

Pediatric acute diarrhea presents as a change in normal bowel habit including a substantial increase in the frequency and/or a decrease in stool consistency. Acute diarrhea resolves in less than 7 days and no longer than 14 days (Guandalini 2004; Guarino 2008). The degree of severity can be related to the child's age and nutritional status, and the underlying cause of diarrhea. Though diarrhea acts as a defense mechanism in the body, quickly eliminating infective organisms, the most serious sequela is dehydration, particularly in malnourished or immuno-suppressed children (WHO 2006). Additional symptoms that accompany acute diarrhea often include abdominal pain, fever and vomiting (Guandalini 1998; Guandalini 2000; Vernacchio 2006).

1.1.2 Burden of Disease

Diarrheal diseases continue to be a major cause of morbidity and mortality in children worldwide. The WHO Global Burden of Disease initiative has adopted a health status measurement known as "disability-adjusted life-year", a unifying metric that combines mortality and morbidity. In 2002, diarrheal diseases ranked second among conditions afflicting children (Mathers 2002). An estimated 2 to 2.5 million diarrhea-associated deaths occurred in children less than 5 years of age, concentrated in the worlds most impoverished regions (Parashar 2003; Kosek 2003). Estimates from the 1990s suggest that approximately 1.4 billion diarrhea episodes occurred yearly among children less than 5 years of age in developing countries. In this population, research has demonstrated a median of 3.2 episodes of diarrhea per child-year (Kosek 2003). In areas where severe malnutrition exists, 6 to 8 episodes of diarrhea per child per year have been reported (Guandalini 2004). Pediatric acute diarrhea is a leading cause of child hospitalization and occurs most frequently during the first two years of life, with infants less than 6 months most at risk for severe diarrhea (Avendano 1993; Molbak 1997; Fagundes-Neto 1999). Of the 1.4 billion episodes, 123.6 million episodes necessitated outpatient medical care and 9 million cases resulted in hospitalization. The cost per diarrheal episode in the United States has been estimated at $289 USD in <36-month old ambulatory population

and the costs of hospitalization of 250,000 patients was $560 million or $2240 USD per case (Avendano 1993; Garthright 1988).

1.1.3 Etiology

Acute diarrhea is most frequently a result of a gastrointestinal infection, with viral infections being the most common cause. In developed countries *Rotavirus* results in 25 to 40% of cases (Guandalini 2004). However, the pathogenesis of acute diarrhea is multi-factorial and may be a result of other pathogens. In fact, more than 20 viral, bacterial and parasitic enteropathogens can trigger diarrhea (O'Ryan 2005). Other known causes of acute diarrhea include drug or radiation induced enteritis, food allergies, digestive/absorptive disorders, vitamin deficiencies or ingestion of heavy metals (Guandalini 2004). See Table 1-1 for a list of key pathogenic and non-pathogenic causes of acute diarrhea.

Table 1-1: *Known Causes of Acute Diarrhea*

Infections
Enteric infections (including food poisoning)
Extraintestinal infections
Drug/Medical Intervention Induced
Antibiotic-associated
Chemotherapy or radiation-associated
Other drugs
Enteral tube feeding
Food Allergies
Cow's milk protein allergy
Soy protein allergy
Multiple food allergies
Disorders of Digestive/Absorptive Processes
Sucrase-isomaltase deficiency
Late-onset (or "adult type") hypolactasia
Vitamin Deficiencies
Niacin deficiency
Zinc deficiency
Ingestion of Heavy Metals
Copper, Tin, Zinc

Adapted from Guandalini S (2004)

1.2 Antibiotic-Associated Diarrhea

Between 10 and 100 trillion bacteria inhabit the human gut, and a balance of these micro-organisms is important in normal gastrointestinal function (Gill 2006). Antibiotic treatment may disturb the colonization resistance of gastrointestinal flora, resulting in a range of clinical symptoms, most notably, diarrhea. This is referred to as antibiotic-associated diarrhea (AAD). The first case report of AAD appeared in the *Bulletin of the John Hopkins Hospital* in 1893, where John Finney and Sir William Osler described the case of a young woman who died of a severe case of "diphteric colitis" shortly after gastric surgery (Finny 1893). It was not until the mid-1900s, after the discovery and wide spread use of antibiotics, that AAD became a medical problem (Kingston 2008). Subsequently, the increased use of antibiotics was accompanied by an increased prevalence of AAD. The symptoms of AAD are typically benign, self-limiting and include frequent watery bowel movements, urgency and crampy abdominal pain. In severe cases, often associated with *Clostridium difficile*, AAD may lead to electrolyte disturbances, dehydration, pseudomembranous, hemorrhagic colitis, toxic megacolon, and possibly death (Berrington 2004).

1.2.1 Prevalence and Risk Factors

The incidence of AAD varies greatly. While reports in the general population indicate that AAD occurs in approximately 5-62% of patients between initiation of therapy and up to two months after the end of treatment (Wistrom 2001; McFarland 1998; Elstner 1983; Turck 2003; LaRosa 2004), the weighted baseline risk of AAD for different age groups indicates a pooled incidence rate of 25% in adults, 28.2% in the elderly, and 42.5% in children (Scheike 2006). Based on an estimated 23.4 million antibiotic prescriptions dispensed in Canada in 2004 (Marra 2006), the number of AAD events can be estimated to range from 1.2 to 14.5 million AAD per year. Risk factors that predispose an individual to AAD include compromised immunity, abdominal surgery, comorbidity, types and prolonged use of antibiotics, length of hospitalization and age (elderly or infant) (McFarland 1998; Turck 2003). But the pre-eminent risk factor is the use of antibiotics. Although almost all antibiotics are associated with AAD, those most

commonly associated with diarrhea are ampicillin, amoxicillin-clavulanate, cephalosporins and clindamycin (McFarland 1998; Louie 2004; Turck 2003).

While AAD can be associated with numerous enteropathogens (*Clostridium perfringens, Klebsiella oxytoca, Staphylococcus aureas, Candida species*), *Clostridium difficile* is most often associated with AAD, including its most serious complications. *C. difficile* is a spore-forming, anaerobic, gram positive bacillus, which may colonize the gut after its ingestion, when normal intestinal microflora is disrupted by antibiotics (McFarland 1998; McFarland 2007). Although the asymptomatic carriage of *C. difficile* is not thought to be a major health concern to carriers themselves, the emergence of fluoroquinolone resistant strain BI/NAP1/027, which produces the characteristic toxins A and B as well as binary toxins has been associated with community acquired cases both adults and children, many sufficiently serious to require hospitalization (CDC 2005; McFarland 2007; Morinville 2005). However, *C. difficile* diarrhea occurs most often in older, immuno-compromised, hospitalized adults exposed to broad-spectrum antibiotics (approximately one-third of AAD can be attributed to *C. difficile*) (McFarland 1998). A difference in rates of colonization has been described between inpatients (20% to 30%) and outpatients (3%) (Bartlett 2002). The evidence for the relationship between *C. difficile* and pediatric AAD is less extensive in pediatric inpatients (Starr 2005; Gogate 2005). For instance, in a case-control study of 500 hospitalized children aged 5-12 years having developed diarrhea, 250 whom were prescribed 5 or more days of antibiotics, and 250 age and sex matched controls, 18% of cases tested positive for *C. difficile* versus 0% of controls (Gogate 2005). There are few data on *C. difficile*-associated diarrhea in the pediatric outpatient population (Elstner 1983; MacFarland 2007; Surawicz 2003).

1.3 Probiotics

In the early 19th century, microbiologists observed the indigenous microflora of healthy versus diseased individuals, noting distinct differences. They proposed that the re-inoculation and normalization of unbalanced indigenous microflora might prevent disease and restore health. *Lactobacillus bulgaricus*, discovered by Metchinkoff in 1905, was the first microbe used for this purpose and was subsequently promoted as a life sustaining

cultured dairy product (Metchikoff 1907). The term *probiotic,* is derived from the Greek language, meaning "for life", and was first used in 1965 to describe "substances secreted by one microorganism which stimulates the growth of another" as opposed to the term antibiotic (Lilly 1965). Since this time, many definitions for the term probiotic have been suggested. Currently, the term probiotic is commonly defined as "*a live microbial culture or cultured diary product which beneficially influences the health and nutrition of the host*" (Schreznmeir 2001). The rationale behind probiotic administration may be mediated by a number of functions (e.g., re-inoculation of disturbed indigenous microflora secondary to antibiotic use, inhibition of pathogens by competition for nutrients and intestinal adhesion sites) (Fedorak 2004; Gueimonde 2006) using specific probiotic species such as *Lactobacillus*, *Bifidobacterium* and *Saccharomyces* (See Table 1-2 for details).

Table 1-2: *Probiotic Species*

Common probiotic species often found in commercial products	**Common probiotic species used in antibiotic-associated diarrhea studies**	**Probiotic species used in studies relevant to antibiotic-associated diarrhea in children**
Bidobacterium *B. animalis* subsp. *lactis (BB12)* *B. bifidum* *B. infantis* *B. longum* *B. thermophilum* *B. adolescenti* **Clostridium** *C. butyicum* **Lactobacillus** *L. acidophilus* *L. brevis* *L. casei* *L. cellobiosus* *L. curvatus* *L. delbruecki* subsp. *bulgaricus* *L. fermentum* *L. plantarum* *L. reuteri* *L. rhamnosus GG (LGG)* *L. salivarius* *L. sporogenes* **Streptococcus** *S. salivarius* subsp. *thermophilus* **Yeasts** *Saccharomyces boulardii* *Saccharomyces cerevisiae*	**Bifidobacterium** *B. animalis* subsp. *lactis (BB12)* *B. bifidum* *B. longum* **Clostridium** *C. butyicum* **Lactobacillus** *L. acidophilus* *L. casei* *L. delbruecki* subsp. *bulgaricus* *L. rhamnosus GG (LGG)* **Streptococcus** *S. salivarius* subsp. *thermophilus* **Yeasts** *Sarrharomyces boulardii*	**Bifdobacterium** *B. bifidum* *B. infantis* *B. lactis** **Clostridium** *C. butyicum* **Lactobacillus** *L. acidophilus* *L. delbruecki* subsp. *bulgaricus* *L. rhamnosus GG (LGG)* *L. sporogenes* **Streptococcus** *S. salivarius* subsp. *thermophilus* **Yeasts** *Sarrharomyces boulardii*

Adapted from Madsen K (2001); *Based on Correa et al. 2005 unclear whether BB12

1.3.1 Evidence from Clinical Trials

Five per-protocol meta-analyses representing 25 trials of probiotics have pointed to the potential efficacy of co-administrating probiotics with antibiotics to prevent AAD in the general population, each reporting significantly reduced AAD (Cremonini 2002; D'Souza 2002; McFarland 2006; Sazawal 2006; Szajewska 2005). The first of these, a meta-analysis of *Saccharomyces boulardii* and *Lactobacillus GG* (*LGG)* co-administered with

antibiotics for the prevention of AAD in a diverse population (881 patients ranging in age from two weeks to elderly, including inpatients, outpatients and populations from developing countries), provided promising evidence to suggest that probiotic agents prevent AAD (RR 0.40; 95% CI 0.27, 0.57) (Cremonini 2002). Szajewska reviewed five trials involving children and adults (n=1076) that administered S. boulardii compared with placebo reporting a significantly reduced risk of AAD (RR 0.43; 95% CI: 0.23, 0.78) (Szajewska 2005). More recent meta-analyses include substantially more patients (Sazawal 2006; McFarland 2006). Sazawal reviewed 19 trials involving 2008 patients in mostly developed countries, demonstrating that probiotics (involving 14 different strains) prevent AAD in the general population (RR 0.48; 95% CI: 0.35, 0.65, $I^2 = 53\%$) (Sazawal 2006). McFarland reviewed 25 trials, 13 reporting a significant reduction of AAD with a pooled relative risk of 0.43, (95% CI 0.31, 0.58). Although promising, each of these meta-analyses are limited, especially with regards to their relevance to pediatric AAD. They did not apply an ITT analysis, nor did any of these reviews meta-analyze for adverse events. While other reviews did conduct *a priori* subgroup analyses on children versus adults (McFarland 2006), and *a priori* subgroups by probiotic strain (McFarland 2006) and dose (McFarland 2006; Szajewska 2005), unlike our review, the latter subgroups were not specific to children. Our review was the most comprehensive, and in addition to the analyses mentioned above, also conducted *a priori* subgroups on the definition of diarrhea, the risk of bias (Jadad score <3 vs. ≥3), and compared the random versus the fixed effects model.

.

1.3.2 Safety

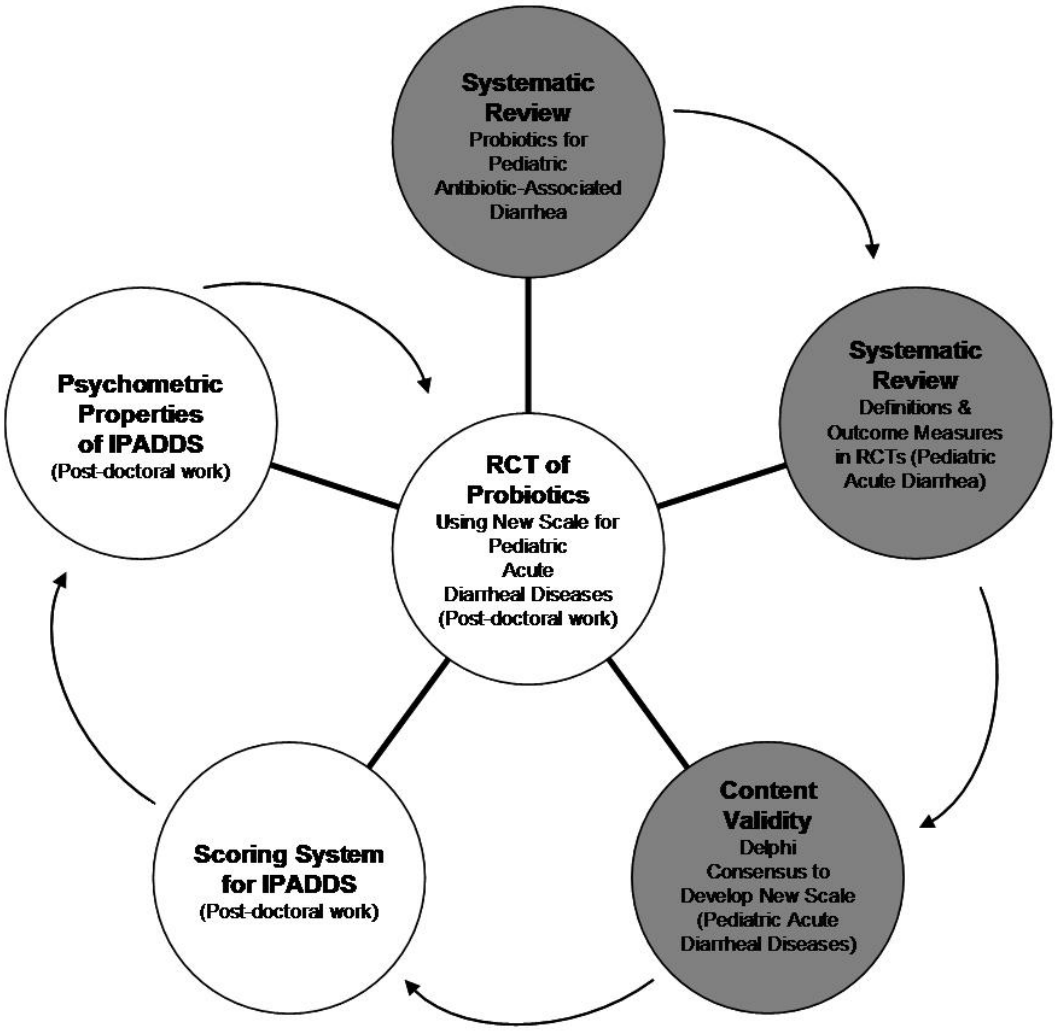
According to the best available evidence, safety does not appear to be a concern in healthy individuals; however, safety is relative, not absolute (Hammerman 2006). Although infections (e.g. bacteremia, endocarditis, septicemia, pneumonia, deep abdominal abscesses) resulting from probiotic use have been reported in neonates, severely debilitated and immuno-compromised individuals (Salminen 1998; McFarland 1998), there is still debate on probiotics safety in these patients. Prospective studies have demonstrated the safety of probiotics in immuno-compromised adults and children with HIV and preterm neonates, with no infections secondary to probiotics reported

(Saliminen 2004; Cunningham 2000; Bin-Nun 2005; Lin 2005). A workshop involving international clinical experts was convened in 2001 to review the overall safety of probiotic consumption over a wide range of probiotic doses, strains, and intervention intervals and concluded "*Current evidence suggests that the risk of infection with probiotic lactobacilli or bifidobacterium is similar to that of infection with commensal strains, and that the consumption of such products presents a negligible risk to consumers, including immunocompromised hosts*" (Borriello 2003).

1.4 PURPOSE

The initial purpose of this doctoral dissertation was to systematically evaluate the randomized trial literature on the use of probiotics to prevent pediatric AAD and to address the methodological limitations of these trials. Contingent upon what was identified, we planned to develop and conduct a RCT to address these limitations. The initial systematic review identified the absence of a standardized definition or a valid and reliable primary outcome measure for PAD. Based on this, a second systematic review of definitions and primary outcome measures in RCTs of PAD was conducted confirming what was uncovered in our first systematic review. This review revealed three disease activity indexes (Diarrheal Disease Index, Ruuska & Vesikari Scale, and Clark Scale), each having limited evidence of validity and no published evidence of reliability for evaluating PAD morbidity. Recognizing the importance of a valid and reliable primary outcome measure to avoid misleading results regarding the efficacy of interventions evaluated in RCTs, and the need for standardized definitions and outcome measures when pooling trials, we then aimed to address this deficiency. Consequently, for the third study we did not conduct an RCT. Rather, the purpose of this study was to develop and begin to collect validity evidence to support the IPADDS, developed for use in RCTs to assess PAD morbidity in children ≤5 years of age. The IPADDS represents the first pediatric acute diarrhea scale developed using sound systematic and scientific methods. Development of a validated outcome measure for diarrhea will be an important methodological advance, which may improve the internal and external validity of therapeutic trials in this area. See Figure 1-1 below for an overview of work.

Figure 1-1: Overview of PhD Objectives



Note: Dissertation objectives in grey

1.5 SPECIFIC OBJECTIVES

1) In randomized controlled trials of children (0 -18 years) involving the outcome measure AAD, to systematically assess: a) whether probiotics (any specified strain or dose) co-administered with antibiotics (any agent) reduce the incidence of antibiotic-associated diarrhea in children; b) adverse events of probiotics when co-administered with antibiotics in children; and c) which probiotic strain(s) and dose(s) yield the most beneficial results in reducing the incidence of diarrhea.

2) In randomized controlled trials of children (0 -18 years) involving acute diarrhea as their primary outcome measure, to systematically document: a) how acute diarrhea and its resolution is defined; b) how acute diarrhea is assessed; and c) the reporting of clinimetric properties (i.e., the validity and reliability) of the outcome measures.

3) In children (0-5 years of age): a) How should acute diarrhea and its resolution be defined? b) What items should be included in a scale to evaluate the severity of PAD? and; c) Given these items, what response scale should be used to obtain reliable data that can be validly interpreted?

1.6 ORGANIZATION OF DISSERTATION

This dissertation is organized as follows: Chapter 2 presents a systematic review and meta-analysis of probiotics for the prevention of pediatric AAD. The chapter is broken to the standard background, methods, results and discussion. The results are ordered as follows: main results (incidence of diarrhea, adverse events), secondary results (mean duration of diarrhea, mean stool frequency), *a priori* subgroups (probiotic strain, definition of diarrhea, dose-response, and antibiotic agent), *a priori* sensitivity analyses (random vs fixed effects, quality using the Jadad and Shultz critera, ITT analysis and publication bias). The chapter concludes with implications for practice and research.

Chapter 3 is a systematic review of the uniformity, reliability and validity of definitions of diarrhea and primary outcome measures in RCTs of PAD. The chapter is also divided into the standard background, methods, results and discussion. The results are presented

in four sections: definitions of diarrhea and diarrhea resolution, primary outcome measures, reliability of grading systems and validity of grading systems to support outcome evaluation. The chapter concludes with implications for research for RCTs and systematic reviews of interventions for PAD, including further implications relevant to chapter 2. *Chapter 4* reports the development of a scale to measure the severity of pediatric acute diarrheal disease involving discussions with a steering committee and a modified Delphi consensus study. The study is presented in three sequential phases: (i) steering committee discussions to develop the initial definitions of PAD and the accompanying items; (ii) collection of content validity evidence by an external panel of local, national and international experts and revision of the initial definitions and items in the instrument; and (iii) assessment of the utility of the final instrument by the same panel of experts. *Chapter 5* provides a summary of research methods employed, the major study findings and study limitations, followed by implications for practice and future research.

REFERENCES

Avendano P, Matson DO, Long J, Whitney S, Matson CC, Pickering LK. Costs associated with office visits for diarrhea in infants and toddlers. Pediatr Infect Dis J. 1993;12:897-902.

Bartlett JG. Clostridium difficile-associated enteric disease. Curr Infect Dis Rep. 2002;4:477-483.

Berrington A, Borriello SP, Brazier J, Duckworth G, Foster K, Freeman R et al. National Clostridium difficile standards group: Report to the department of health. J Hospit Infect. 2004;56:1-38.

Bin-Nun A, Bromiker R, Wilschanski M, Kaplan M, Rudensky B, Caplan M, et al. Oral probiotics prevent necrotizing enterocolitis in very low birth weight neonates. J Pediatr. 2005;147:192-196.

Borriello SP, Hammes WP, Holzapfel W, Marteau P, Schrezenmeir J, Vaara M, Valtonen V. Safety of probiotics that contain lactobacilli or bifidobacteria. Clin Infect Dis. 2003;36:775-780.

Centers for Disease Control and Prevention. Severe Clostridium difficile-associated diesease in populations previosly at low risk – four states, 2005. MMWR Morb Mortal Wkly Rep. 2005;54:1201-1205.

Correa NB, Peret Filho LA, Penna FJ, Lima FM, Nicoli JR. A randomized formula controlled trial of Bifidobacterium lactis and Streptococcus thermophilus for prevention of antibiotic-associated diarrhea in infants. J Clin Gastroenterol. 2005;39:385-389.

Cremonini F, DiCaro S, Nista EC, Bartolozzi F, et al. Meta-analysis: the effect of probiotic administration on antibiotic associated diarrhea. Aliment Pharmacol Ther. 2002;16:1461-1467.

Cunningham-Rundles S, Ahrne S, Bengmark S, Johann-Liang R, Marshall F, Metakis L, et al. Probiotics and immune response. Am J Gastroenterol. 2000;95(1 Suppl):S22-S25.

D'Souza A, Rajkumar C, Cooke J, Bulpitt CJ. Probiotics in prevention of antibiotic associated diarrhea: meta-analysis. BMJ 2002;324:1361-1364.

Elstner CL, Lindsay AN, Book LS, Matsen JM. Lack of relationship of Clostridium difficile to antibiotic-associated diarrhea in children. Pediatr Infect Dis J. 1983;2:364-6.

Fagundes-Neto U, de Andrade JA. Acute diarrhea and malnutrition: lethality risk in hospitalized infants. J Am Coll Nutr. 1999;18:303-308.

Fedorak RN, Madsen KL. Probiotics and prebiotics in gastrointestinal disorders. Curr Opin Gastroenterol. 2004;20:146-155.

Finney J. Gastro-enterostomy for cicatrizing ulcer of the pylorus. Bull John Hopkins Hosp. 1893;4:53.

Garthright WE, Archer DL, Kvenberg JE. Estimates of incidence and costs of intestinal infectious diseases in the United States. Public Health Rep. 1988;103:107-115.

Gueimonde M, Salminen S. New methods for selecting and evaluating probiotics. Dig Liv Dis 2006;38:S242-S247.

Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Buck SS, et al. Metagenomic analysis of the human distal gut microbiome. Science 2006;312:1355-1359.

Gogate A, De A, Nanivadeka R, Mathur M, Saraswathi K, Jog A, et al. Diagnostic role of stool cultured & toxin in detection of antibiotic-associated diarrhea. Indian J Med Res. 2005;122:518-524.

Guandalini S (2004) Acute Diarrhea. In: Walker AW, Goulet O, Kleinman RE, Sherman PM, Shneider BL, Sanderson IR, editors. Pediatric Gastrointestinal Disease: Pathophysiology, Diagnosis, Management. Toronto: BC Decker Inc. pp. 166-179.

Guandalini S, Dincer AP. Nutritional management in diarrhoeal disease. Baillieres Clin Gastroenterol. 1998;12(4):697-717.

Guandalini S, Pensabene L, Zikri MA, Dias JA, Casali LG, Hoekstra H, et al. Lactobacillus GG administered in oral rehydration solution to children with acute diarrhea: a multicenter European trial. J Pediatr Gastroenterol Nutr. 2000;30:54-60.

Guarino A, Albano F, Ashkenazi S, Gendrel D, Hoekstra JH, Shamir R, Szajewska H. Paediatric Infectious Diseases evidence-based guidelines for the management of acute gastroenteritis in children in Europe: executive summary. Pediatr Gastroenterol Nutr. 2008; 46(Supple. 2):S81-S122.

Hammerman C, Bin-Nun A, Kaplan M. Safety of probiotics: comparison of two popular strains. BMJ 2006;333:1006-1008.

Johnston BC, Thorlund K. Probiotics for the prevention of Clostridium difficile associated diarrhea in adults and children (Protocol). Cochrane Database Syst Rev: Cochrane Library 2009, Issue 1. CD006095.

Kingston W. "Irish contributions to the origins of antibiotics". Ir J Med Sci. 2008;177:87–92.

Kosek M, Bern C, Guerrant RL. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. Bull World Health Organ. 2003;81:197-204.

LaRosa M, Bottaro G, Gulino N, Gambuzza F, Di Forti F, Ini G, et al. Prevention of antibiotic-associated diarrhea with Lactobacillus sporogens and fructo-oligosaccharides in children: A multicentric double-blind vs placebo study. Minerva Pediatr. 2003;55:447-452.

Lilley DM, Stillwell RH. Probiotics: growth promoting factors produced by microorganisms. Science 1965;147:747-748.

Lin HC, Su BH, Chen AC, Lin TW, Tsai CH, Yeh TF, et al. Oral probiotics reduce the incidence and severity of necrotizing enterocolitis in very low birth weight infants. Pediatrics 2005;115:1-4.

Louie TJ, Meddings J. Clostridium difficile infection in hospitals: risk factors and responses. CMAJ 2004;171:45-46.

Marra F, Patrick DM, Chong M, Bowie WR. Antibiotic use among children in British Columbia, Canada. J Antimicrob Chemother. 2006;58:830-839.

Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3:e442.

Madsen KL. The use of probiotics in gastrointestinal disease. Can J Gastroenterol. 2001; 15(12):817-22.

McFarland LV. Epidemiology, risk factors and treatments for antibiotic-associated diarrhea. Digest Dis. 1998;16:292-307.

McFarlandLV. Meta-Analysis of probiotics for the prevention of antibiotic-associated diarrhea and the treatment of Clostridium difficle disease. Am J Gastroenterol. 2006;101:812-822.

McFarland LV, Beneda HW, Clarridge JE, Raugi GJ. Implications of the changing face of Clostridium difficle disease for health care practitioners. Am J Infect Control 2007;35:237-253.

Metchnikoff E. The prolongation of life. Optimistic studies. London: William-Heinemann, 1907.

Molbak K, Jensen H, Ingholt L, Aaby P. Risk factors for diarrheal disease incidence in early childhood: a community cohort study from Guinea-Bissau. Am J Epidemiol. 1997;146:273-282.

Morinville V, McDonald J. Clostridium difficile-associated diarrhea in 200 Canadian children. Can J Gastroenterol. 2005;19:497-501.

O'Ryan M, Prado V, Pickering LK. A millennium update on pediatric diarrheal illness in the developing world. Semin Pediatr Infect Dis. 2005;16:125-36.

Parashar UD, Bresee JS, Glass RI. The global burden of diarrhoeal disease in children. Bull World Health Organ. 2003;81:236.

Parashar UD, Hummelman EG, Bresee JS, Miller MA, Glass RI. Global illness and deaths caused by rotavirus disease in children. Emerg Infect Dis. 2003;9:565-572.

Ruuska T, Vesikari T. Rotavirus disease in Finnish children: use of numerical scores for clinical severity of diarrhoeal episodes. Scand J Infect Dis. 1990;22:259-67.

Salminen MK, Tynkkynen S, Rautelin H, Poussa T, Saxelin M, Ristola M, et al. The efficacy and safety of probiotic Lactobacillus rhamnosus GG on prolonged, noninfectious diarrhea in HIV Patients on antiretroviral therapy: a randomized, placebo-controlled, crossover study. HIV Clin Trials 2004;5:183-191.

Salminen S, von Wright A, Morelli L, Marteau P, Brassart D, de Vos WM, et al. Demonstration of safety of probiotics - a review. Intern J Food Microbiol. 1998; 44:93-106.

Sazawal S, Hiremath G, Dhingra U, Malik P, Deb S, Black RE. Efficacy of probiotics in acute diarrhea: a meta-analysis of masked, randomized, placebo-controlled trials. Lancet Infect Dis. 2006;6:374-382.

Scheike I, Connock M, Taylor R, Fry-Smith A, Ward D. Probiotics for the prevention of antibiotic-associated diarrhea. West Midlands Health Technology Assessment Collaboration, Department of Public Health and Epidemiology, University of Birmingham. Report Number 56, 2006.

Schrezenmeir J, de Vrese M. Probiotics, prebiotics, and synbiotics: approaching a definition. Am J Clin Nutr. 2001;73:S361–S364.

Starr J. Clostridium difficile-associated diarrhoea: diagnosis and treatment. BMJ. 2005;331:498-501.

Surawicz CM. Probiotics, antibiotic-associated diarrhoea and Clostridium difficile diarrhoea in humans. Best Pract Res 2003; 17:775-783.

Turck D, Bernet JP, Marx J, Kempf H, Giard P, Walbaum O, et al. Incidence and risk factors of oral antibiotic-associated diarrhea in an outpatient pediatric population. J Pediat Gastroenterol Nutr. 2003;37:22-26.

Vernacchio L, Vezina RM, Mitchell AA, Lesko SM, Plaut AG, Acheson DWK. Diarrhea in American infants and young children in the community setting: Incidence, clinical presentation and microbiology. Pediatr Infect Dis J. 2006;25:2-7.

Wistrom J, Norrby SR, Myhre EB, Eriksson S, Granstrom G, Lagergren L, et al. Frequency of antibiotic-associated diarrhoea in 2462 antibiotic-treated hospitalized patients: a prospective study. J Antimicrob Chemother. 2001;47:43-50.

World Health Organization (2005) The treatment of diarrhea: A manual for physicians and other senior health workers. Accessed May 06, 2009: http://whqlibdoc.who.int/publications/2005/9241593180.pdf

CHAPTER 2: Probiotics for the Prevention of Pediatric Antibiotic-Associated Diarrhea: A Systematic Review & Meta-analysis

In August 2006, a limited version of this systematic review of placebo-controlled trials only was published in the *Canadian Medical Association Journal* (Johnston et al. 2006). In April 2007, the complete systematic review of placebo, active control and no treatment controlled trials was published in the *Cochrane Database of Systematic Reviews* (Johnston et al. 2007). The latter is presented here.

2.1 ABSTRACT

Background

Antibiotics alter the microbial balance within the gastrointestinal tract. Probiotics may prevent AAD via restoration of the gut microflora. Antibiotics are prescribed frequently in children and AAD is common in this population.

Objectives

To assess the efficacy and adverse effects of probiotics (any specified strain or dose) for the prevention of antibiotic-associated diarrhea in children.

To assess adverse events associated with the use of probiotics when co-administered with antibiotics in children.

Search strategy

MEDLINE, EMBASE, CENTRAL, CINAHL, AMED, and the Web of Science (inception to August 2006) were searched along with specialized registers including the Cochrane IBD/FBD Review Group, CISCOM, Chalmers PedCAM Research Register and trial registries from inception to 2005. Letters were sent to authors of included trials, nutra/pharmaceutical companies, and experts in the field requesting additional information on ongoing or unpublished trials. Conference proceedings, dissertation abstracts, and reference lists from included and relevant articles were hand searched.

Selection criteria

Randomized, parallel, controlled (placebo, active, or no treatment) trials comparing co-administered probiotics with antibiotics for the prevention of diarrhea secondary to antibiotic use in children (0 to 18 years).

Data collection & analysis

Methodological quality assessment and data extraction were conducted independently by two authors (BCJ, AS). Dichotomous data (incidence of diarrhea, adverse events) were combined using pooled relative risks, and continuous data (mean duration of diarrhea, mean daily stool frequency) as weighted mean differences, along with their corresponding 95% confidence intervals. Adverse events were summarized using risk difference. For overall pooled results on the incidence of diarrhea, *a priori* sensitivity analyses included per protocol versus ITT, random versus fixed effects, and methodological quality criterion. Subgroup analysis were conducted on probiotic strain, dose, definition of antibiotic-associated diarrhea, and antibiotic agent.

Main results

Ten studies met the inclusion criteria. Trials included treatment with either *Lactobacilli spp., Bifidobacterium spp., Streptococcus spp.,* or *Saccharomyces boulardii* alone or in combination. Six studies used a single strain probiotic agent and four combined two

probiotic strains. The per protocol analysis for 9/10 trials reporting on the incidence of diarrhea show statistically significant results favouring probiotics over active/non active controls (RR 0.49; 95% CI 0.32 to 0.74). However, ITT analysis showed non-significant results overall (RR 0.90; 95% CI 0.50 to 1.63). Five of ten trials monitored for adverse events (n = 647); none reported a serious adverse event.

Reviewers' conclusions

Probiotics show promise for the prevention of pediatric AAD. While per protocol analysis yields treatment effect estimates that are both statistically and clinically significant, as does analysis of high quality studies, the estimate from the ITT analysis was not statistically significant. Future studies should involve probiotic strains and doses with the most promising evidence (*Lactobacillus GG, Lactobacillus sporogenes, Saccharomyces boulardii* at 5 to 40 billion colony forming units/day). Research done to date does not permit determination of the effect of age (e.g., infant versus older children) or antibiotic duration (5 days versus 10 days). Future trials would benefit from a validated primary outcome measure for antibiotic-associated diarrhea that is sensitive to change and reflects what treatment effect clinicians, parents, and children consider important. The current data are promising, but it is premature to routinely recommend probiotics for the prevention of pediatric AAD.

2.2 BACKGROUND

2.2.1 Antibiotic-Associated Diarrhea

More than 400 species of bacteria inhabit the human gut, and a balance of these micro-organisms is important for normal gastrointestinal function (Madsen 2001). Antibiotic treatment may disturb the colonization resistance of gastrointestinal flora, resulting in a range of clinical symptoms, most notably, diarrhea. In particular, antibiotics that act on anaerobes, such as aminopenicillins, cephalosporins and clindamycins, are most associated with diarrhea (Wistrom 2001; McFarland 1998). Antibiotic-associated diarrhea is associated with altered intestinal microflora, mucosal integrity, vitamin, mineral metabolism and crampy abdominal pain (Saavedra 1999). If severe, AAD may lead to electrolyte disturbances, dehydration, premature discontinuation of antibiotic therapy, pseudomembranous colitis, toxic megacolon and possibly death (Arvola 1999). Reports in the general population indicate that the incidence of AAD ranges from 5 to 62%, occurring at any point from the initiation of therapy to two months after the end of treatment (Wistrom 2001; McFarland 1998; LaRosa 2003). The incidence of diarrhea in children receiving broad spectrum antibiotics has been reported in the range of 11 to 40% (Turck 2003; Elstner 1983). The overgrowth of many enteropathogens has been associated with antibiotic-induced diarrhea. *Clostridium difficile* overgrowth is the bacterial agent most associated with AAD (Bartlett 1978; McFarland 1998). *C. difficile* diarrhea is associated with the most serious adverse events associated with AAD, and occurs most often in older, immunocompromised, hospitalized adults, but also occurs in children (Gogate 2005).

The definition of AAD varies across trials. Although the WHO defines pediatric acute diarrhea as three or more abnormally loose bowel movements per 24 hours, the definition in adult and pediatric trials ranges from one to three abnormally loose stools per 24 to 48 hours (Kotowska 2005; Tankanow 1990). Additionally, stool frequency is more difficult to quantify in diaper-aged children with diarrhea and may vary substantially between infants and older children.

2.2.2 Probiotics

Probiotics refer to so-called "friendly" non-pathogenic bacterial or yeast microbiota intended to benefit the host via altering the microflora by implantation or colonization (Schrezenmeir 2001; Havenaar 1992). The rationale behind probiotic administration is based on re-inoculation and normalization of unbalanced indigenous microflora using specific probiotic strains. Probiotics have been administered both prophylactically and therapeutically in an attempt to modify the mucosal, epithelial, intestinal and systemic immune activity in ways that may benefit human health. Probiotics are reported to improve microbial balance in the intestinal tract and display both antibacterial and immune regulatory effects in humans (Gibson 1998; Goldin 1998). Five meta-analyses have been published on the use of probiotics for the prevention of AAD in the general population. The results favoured probiotic co-administration with antibiotics (RR 0.40, 95% CI 0.27 to 0.57; OR 0.37, 95% CI 0.26 to 0.53; RR 0.48, 95% CI 0.35 to 0.65; and RR 0.43, 95% CI 0.31 to 0.58 and RR 0.43; 95% CI: 0.23, 0.78) (Cremonini 2002; D'Souza 2002; Sazawal 2006; McFarland 2006; Szajewska 2006). In addition, we recently published a meta-analysis of placebo-controlled trials to evaluate the efficacy of probiotics for preventing antibiotic-induced diarrhea in children (Johnston 2006). Probiotics commonly administered in randomized controlled trials of AAD are: *Lactobacillus acidophilus, Lactobacillus bulgaricus, Lactobacillus casei, Lactobacillus rhamnosus, Bifidobacteria bifidum, Bifidobacteria longum, Streptococcus thermophilus, Saccharomyces boulardii* and *Clostridium butyicum.* Safety does not appear to be a concern in healthy individuals, although serious infections (e.g., pneumonia, bacteremia, endocarditis, deep abdominal abscesses, meningitis) have been reported in neonates, severely debilitated and or immunocompromised individuals (Land 2005; Salminen 2004; Mackay 1999; Rautio 1999;Piarroux 1999; Salminen 1998; Saxelin 1996; Hata 1988; Sussman 1986).

Aims of Treatment

Prevent or ameliorate diarrhea (i.e., shorten duration and/or severity of diarrhea).

Treatment Options

Surrogate antibiotic with a decreased risk of inciting diarrhea;

Anti-diarrhea drugs (e.g., diosmectite, loperamide);

Concurrent probiotics (vs. probiotics administered after diarrhea occurs); and

Discontinuation of antibiotic therapy.

2.3 OBJECTIVES

PRIMARY

1) To systematically assess whether probiotics (any specified strain or dose) co-administered with antibiotics (any agent) reduce the incidence of antibiotic-associated diarrhea in children.

2) To systematically assess adverse events of probiotics when co-administered with antibiotics in children.

SECONDARY

1) To systematically assess which probiotic strain(s) and dose(s) yield the most beneficial results in reducing the incidence of diarrhea.

2) To systematically assess whether probiotics (any specified strain or dose) co-administered with antibiotics (any agent) reduce the mean duration of diarrhea.

3) To systematically assess whether probiotics (any specified strain or dose) co-administered with antibiotics (any agent) reduce the mean daily stool frequency.

Criteria for considering studies for this review

Types of studies

All randomized controlled trials irrespective of language or publication status, where a specified probiotic agent has been compared to placebo, active, or no treatment control.

Types of participants

Children (0 to 18 years of age), male or female of any ethnic group being administered antibiotic therapy for any reason.

Types of interventions

Intervention group: specific, identified probiotic. Trials investigating non-specific probiotic agents (e.g., products that do not label the probiotic strain and dose), yogurt or other fermented foods were not considered. Trials combining probiotics with prebiotics were included if the prebiotic dose was less than 275 mg, as this was judged to be of limited impact to alter the gut milieu (Roberfroid 1998). Control group: placebo, active, or no treatment control. All studies comparing probiotics to conventional care (i.e., diosmectite, loperamide) or probiotics plus conventional care vs. conventional care plus placebo or no treatment were considered for the review.

Types of outcome measures

PRIMARY OUTCOMES

1. Incidence of diarrhea using the primary investigators' definition (i.e., frequency, consistency of bowel movements)
2. Number and type of adverse events (e.g., bacteremia, meningitis)

SECONDARY OUTCOMES

1. Mean duration of diarrhea
2. Mean stool frequency

Search strategy for identification of studies

ELECTRONIC SEARCHES

A comprehensive search of the following relevant databases irrespective of publication status or language was conducted: The Cochrane Central Register of Controlled Trials (CENTRAL) on the Cochrane Library (2006 Issue 3), The trial registers of the Cochrane IBD/FBD Review Group, the Cochrane Complementary Medicine Field's Register of Controlled Trials, MEDLINE (1966-2006), EMBASE (1980-2006), CINAHL (1982-2006), AMED (1985-2006), Web of Science (1945-2006), the Chalmers Research Group PedCAM Database, and a request for the staff of the Research Council for Complementary Medicine, UK, to search CISCOM (Centralised Information Service for Complementary Medicine), going as far back as individual databases will go.

Conference proceedings and dissertation abstracts were searched through Ovid's OCLC (Online Computer Library Center Inc. 1992-2005), Conference Papers Index (1982-2005), and Dissertation Abstracts (1980-2005) to identify additional studies.

HANDSEARCHES

Bibliographies of randomised controlled trials and review articles were checked for additional studies not identified by the electronic searches. Handsearches of the American Gastroenterological Association meeting abstracts were conducted to identify trials that may not have been published in full.

ADDITIONAL SEARCHES

Primary authors of identified pediatric AAD trials, nutraceutical companies that manufacture probiotic agents and individuals working in the field were contacted to further identify any additional unpublished, ongoing, or planned trials. Ongoing trials were searched through Current Controlled Trial Register, which houses the NHS Controlled Trials Register, the National Institute of Health Register, the National Research Register, and the International Standard Randomized Controlled Trial Number Register. Organisations and individuals working in the field were asked to review the completeness of the search and to further identify additional unpublished, ongoing, planned or relevant trials.

The MEDLINE search strategy is as follows. (See Appendix 2-1 for additional search strategies).

1. exp PROBIOTICS/tu or probiotic$.tw.
2. exp LACTOBACILLUS/ or lactobacill$.tw. or "l acidophilus".tw. or "l casei".tw. or bifidobacter$.mp. or "b infantis".tw. or "b bifidum".tw. or "b longum".tw. or saccharomyce$.mp. or "s boulardii".tw. or clostridium butyricum.tw. or clostridium difficile.mp. or "streptococcus thermophilus".tw. or enterococcus faecium.mp.
3. exp antibiosis/ or biotherapeutic agent$.tw.
4. or/1-3
5. exp Anti-Bacterial Agents/ or antimicrobial$.tw. or antibiotic$.tw.

6. ((antimicrobial or anti microbial or antimycrobial or antimycobacteri$ or antibacteri$ or bacteriocid$) adj3 agent$).tw.

7. 5 or 6

8. exp DIARRHEA/ or diarrhea.tw. or diarrhoe$.tw. or diarhe$.tw. or diahoe$.tw. or dysenter$.tw. or gastro enteritis$.tw. or gastroenteriti$.tw.

9. and/4,7-8

10. child/ or infant/ or adolescence/ or exp infant, new born/ or exp child, preschool/

11. (child$ or newborn$ or adolescen$ or infan$).tw.

12. (preschool$ or pre-school$).tw.

13. teen$.tw.

14. (kindergarten$ or kindergarden$).tw.

15. elementary school$.tw.

16. secondary school$.tw.

17. nursery school$.tw.

18. high school$.tw.

19. highschool$.tw.

20. youth$.tw.

21. (baby$ or babies$ or preemie$ or premature$).tw.

22. (schoolchild$ or "school child$").tw.

23. (schoolage$ or school age$).tw.

24. toddler$.tw.

25. pubert$.tw.

26. (pre-pubescen$ or prepubescen$ or post-pubescen$ or postpubescen$).tw.

27. (kid or kids or boy$ or girl$).tw.

28. juvenile.tw.

29. or/10-28

30. 9 and 29

31. Cochrane RCT filter (Dickersin 1994).

32. 30 and 31

[/=MeSH term, exp=explode, tw=textword, mp=multipurpose word, $=truncation]

2.4 METHODS

STUDY SELECTION

Search results were screened independently by two authors (BCJ, AS) using titles of papers, and when available, abstracts. The full text of the selected articles was retrieved and two authors (BCJ, AS) independently assessed each article for inclusion according to pre-specified selection criteria. Inter- rater reliability was measured using kappa statistics and disagreement was resolved by discussion.

QUALITY ASSESSMENTS

Methodological quality was assessed independently by two authors (BCJ, AS). Quality components were assessed for selection, detection, performance and loss to follow-up bias. Each of the included studies was evaluated using the validated 5-point Jadad Scale to assess randomization, double blinding, withdrawals and dropouts (Jadad 1996). Concealment of allocation was assessed as adequate, inadequate or unclear using trial design methodology described by Schulz (Schulz 1995). Inter-rater reliability was assessed for both quality scales by using kappa statistics and disagreement was resolved by consensus.

DATA EXTRACTION

Using a standardized data extraction form, two authors (BCJ, AS) independently extracted the following data items: author, year of publication, language, study setting, methodological design, funding source, definition and diagnostic criteria for diarrhea, inclusion and exclusion criteria for participants, patient characteristics (age, gender, diagnosis, socioeconomic status), number of patients allocated to each group, presence/absence of an ITT analysis, reasons for withdrawal, measures of compliance, specified antibiotic, specified probiotic, duration, dosage and schedule of antibiotic, duration, dosage and schedule of probiotic, outcome measures(incidence of diarrhea, number of adverse events, mean duration of diarrhea, mean stool consistency, and mean stool frequency).

STATISTICAL ANALYSIS

Results were combined unless diversity (clinical and/or methodological heterogeneity) or statistical heterogeneity (non-overlapping confidence intervals) suggested combination was unreasonable. Dichotomous data are presented as relative risks, and continuous data as weighted mean differences, along with their corresponding 95% confidence intervals. The number needed to treat (NNT) or the number needed to harm (NNH) was considered for statistically significant dichotomous outcomes. Adverse events were summarized using risk difference since these events were rare. Random effects models were used and fixed effects models were considered in sensitivity analyses. Heterogeneity was investigated using the I^2 statistic (Higgins 2003).

Meta-regression or the Chi-square test for heterogeneity were used in subgroup analyses depending on the number of trials included. Subgroup analyses were subdivided by: probiotic strain(s) (when two or more trials administered the same strains), antibiotics that are specific to anaerobes (most associated with diarrhea side effects), diagnostic criteria for diarrhea (e.g. the WHO definition: fewer than 3 abnormally loose bowel movements per 24 hours vs. 3 or more abnormally loose bowel movements per 24 hours), dosage of probiotic (≥5 billion colony forming units of live bacteria, <5 billion colony forming units of live bacteria). Sensitivity analyses were done on quality criterion (e.g., Jadad Scale). To fully appreciate the potential influence of missing responses (e.g. children lost to follow up) sensitivity analyses based on ITT principles were applied for the primary outcome, incidence of diarrhea, using the best and worse case scenarios (extreme-case). In addition to funnel plots, the rank correlation test (Begg 1994) and weighted regression (Egger 1997) were considered to test for publication bias. If publication bias was apparent, adjustment of the pooled estimates was considered using the trim and fill method (Duval 2000). More than one method to evaluate publication bias was considered since the relative merits of the methods are not well established.

2.5 DESCRIPTION OF STUDIES

A total of 652 studies were identified from the primary electronic databases (EMBASE 351, MEDLINE 211, Web of Science 48, CENTRAL 25, CINAHL 20, AMED 3). Of

these, 132 were identified as duplicates, leaving 526 titles and/or abstracts identified as original publications. Independent review (BCJ, ALS) of the titles and/or abstracts identified 25 potentially relevant studies for full-text review (24 full-text and 1 abstract). Additional citations identified through the grey literature included: contact with authors (n = 4), nutraceutical companies (n = 11), experts in the field (n = 2), the trial registers of the Cochrane Inflammatory Bowel Disease and Functional Bowel Disorders Review Group (n = 10), and The Chalmers Research Group PedCAM database (n = 2), and reference lists (n = 1). These additional citations provided two further relevant articles (Chapoy 1985; Erdeve 2005) resulting in 27 articles for full review. Two authors (BCJ, ALS) independently assessed these studies and identified 10 studies that met the inclusion criteria (7 English, 2 Italian, 1 French). Reasons for exclusion were as follows: three studies did not randomize (Chapoy 1985; Czerwionka 2006; Seki 2003); three were not associated with antibiotic use (Lei 2006; Michielutti 1996; Weizman 2005); four involved a non-pediatric population (Duman 2005; Siitonen 1990; Thomas 2001; Witsell 1995); four did not report outcomes particular to AAD (Schrezenmeir 2004; Srinivasan 2006; Sykora 2005; Zoppi 2001), two letters to the editor on *S. boulardii* for pediatric AAD (not randomized controlled trials) (Erdeve 2005; McFarland 2005), and one did not include probiotics as an intervention (Brunser 2006). The inter-rater proportion of overall agreement on inclusion and exclusion was 93% with a kappa coefficient of 0.83.

Design

All included studies were prospective, randomized, controlled trials (placebo, active or no treatment control arm).

Patient population

After withdrawals, the ten studies included a total of 1986 patients (1015 treatment, 971 controls). Patients were diagnosed with upper and lower respiratory tract infections or dermatological infections (Arvola 1999; Benhamou 1999; Contardi 1991; Kotowska 2005; LaRosa 2003; Tankanow 1990; Vanderhoof 1999), and meningitis or septicemia (Jirapinyo 2002). The health care setting was provided in eight studies and consisted of: hospitalized inpatients (Correa 2005; Jirapinyo 2002), private primary care practices

(Benhamou 1999; Contardi 1991; Tankanow 1990; Vanderhoof 1999) and an outpatient university teaching hospital (Arvola 1999; Kotowska 2005). Children enrolled were from families of diverse socioeconomic status, including developed and developing countries such as Brazil (Correa 2005), Thailand (Jirapinyo 2002) and Turkey (Erdeve 2004). Children ranged from 1 month to 15 years of age. Seven studies provided information regarding the participants' mean age: 4.5 years (Arvola 1999), 2.4 years (Benhamou 1999), 1.8 years (Correa 2005), 4.8 years (Kotowska 2005), 6.6 years (LaRosa 2003), 2.5 years (Tankanow 1990) and 4 years (Vanderhoof 1999). Two studies provided only the age range of enrolled participants: 8 months - 3 years (Contardi 1991) and 1 month - 3 years (Jirapinyo 2002). Six studies included both males and females (371 males and 439 females), and four studies did not state sufficient information regarding gender (Arvola 1999; Benhamou 1999; Erdeve 2004; Jirapinyo 2002).

Interventions

All trials reported exposing children to 5 to 15 days of oral antibiotics. One trial administered intravenous antibiotics (cefuroxime) to patients 60/246 (12%) requiring hospitalization (Kotowska 2005). Two studies reported exposure to oral amoxicillin alone (Contardi 1991; Tankanow 1990) using a standard pediatric dosage range (20 to 50 mg/kg/day), whereas the remaining trials reported a mixture of oral antibiotic agents including: bactericidal cephalosporins (e.g., cefotaxime, cefprozil), bacteriostatic macrolides (e.g.,clarithromycin, erythromycin), and the bactericidal beta-lactams/penicillins. In particular, five studies described the antibiotic classes administered. Three studies administered a host of cephalosporins (n = 229) and beta-lactams/penicillins (n = 756) (Benhamou 1999; Correa 2005; Kotowska 2005), one study reported cephalosporins (n = 49), beta-lactams/penicillins in the form of amoxicillin-clavulanate (n = 36) and macrolides in the form of erthromycin (n = 34) (LaRosa 2003), and one study reported exposing children to beta-lactams/penicillins in the form of sulbactam-ampicillin (n = 234) and macrolides in the form of azithromycin (n = 232) (Erdeve 2004).

Trials included treatment with either *Lactobacilli spp., Bifidobacterium spp., Streptococcus spp.,* or *Saccharomyces boulardii*. The strain(s) and daily dosage of the probiotic interventions included: Lactobacillus GG, 20 to 40 billion colony forming units (CFU) bacteria per day (Arvola 1999*); Saccharomyces boulardii*, 4.5 billion yeast/day (Benhamou 1999); *Saccharomyces boulardii*, 5 billion CFU yeast/day (Erdeve 2004); *Lactobacillus acidophilus* and *Bifidobacterium bifidus*, 3 billion CFU bacteria/day (Contardi 1991); *Bifidobacterium lactis* and *Streptococcus thermophilus*, 825 million CFU bacteria/day (Correa 2005); *Lactobacillus acidophilus* and *Bifidobacterium infantis*, dose not reported (Jirapinyo 2002); *Saccharomyces boulardii*, 10 billion CFU of yeast/day (Kotowska 2005); *Lactobacillus sporogenes* and fructo-oligosaccaride (a prebiotic); 5.5 billion CFU bacteria/day and 250 mg prebiotic/day (LaRosa 2003); *Lactobacillus acidophilus* and *Lactobacillus bulgaricus*, 2 billion CFU bacteria/day (Tankanow 1990); and *Lactobacillus GG* and inulin (a prebiotic), 10 to 20 billion CFU bacteria/day equalling 100 mg and 225 mg of the prebiotic inulin/day (Vanderhoof 1999). The latter trial was the only study to use a weight-based approach.

Comparison

In seven studies, the probiotic(s) intervention was compared to a placebo control group, two trials compared probiotics to conventional care (Benhamou 1999; Correa 2005) and one trial compared probiotics to no treatment (Erdeve 2004). In the placebo-controlled trials, contact with authors revealed that the placebo contained an inert amount of inulin (325 mg) - a prebiotic used as capsule filler (Vanderhoof 1999). Three trials provided information on the choice of comparison stating that the placebos contained 'sugar', Saccharum lactis, and 'lactose' respectively (Jirapinyo 2002; Kotowska 2005; Tankanow 1990). For the two trials involving active controls with conventional care, one trial administered diosmectite (an anti-spasmolytic drug) (Benhamou 1999) and the second administered formula containing vitamins, minerals and protein (Correa 2005).

Outcomes

Nine studies (n = 1946) provided data on the incidence of diarrhea, five studies (n = 647) reported on adverse events, four studies (n = 574) reported mean duration of diarrhea, and

three studies (n = 347) reported mean stool frequency. The criteria for defining the incidence of diarrhea varied among the studies and ranged from "one or more abnormally loose bowel movements per day" (Tankanow 1990); two or more liquid stools per day on at least two occasions during the course of the study (Vanderhoof 1999); three or more liquid/watery stools per day (Benhamou 1999; Erdeve 2004), to three or more watery/loose/liquid stools per day for two consecutive days (Arvola 1999; Correa 2005; Kotowska 2005).

Two studies reported on viral and bacterial analysis of fecal samples (Arvola 1999; Kotowska 2005). Along with viral and bacterial fecal analysis, one trial reported on the metabolic activity of gut microflora: fecal urease, ß-glucosidase and ß-glucuronidase activity (Arvola 1999). The second trial reported on frequencies of retroviral diarrhea, salmonella diarrhea, shigella diarrhea and *C. difficile* diarrhea as 'secondary' outcome measures (Kotowska 2005). Other outcomes of potential interest included mean diarrhea incubation and percentage suffering from dehydration reported in one study (Correa 2005). No studies reported on cost-effectiveness related to absenteeism from the workplace, daycare or school between treatment and control groups.

2.6 METHODOLOGICAL QUALITY OF INCLUDED STUDIES

Only two studies reported information concerning *a priori* sample size calculations (Kotowska 2005; Vanderhoof 1999). Loss to follow-up was substantial (i.e. >20%) in 4/9 trials (Arvola 1999; Benhamou 1999; Erdeve 2004; Tankanow 1990). In particular, loss to follow-up was 37% in Tankanow 1990 and 29% in Arvola 1999. Only one trial provided a 'CONSORT' like flow diagram providing details regarding drop-outs (Kotowska 2005). All studies were randomized parallel group designs. Eight studies adopted a "double-blind" procedure (n = 1480); whereas, two trials were unblinded (n = 506) (Contardi 1991; Erdeve 2004). The latter two trials were included because they met *a priori* inclusion criteria - 'prospective, randomized, controlled trial'. Jadad quality scores were as follows: four studies scored two (Arvola 1999; Contardi 1991; Erdeve 2004; Jirapinyo 2002), two studies scored three (Benhamou 1999; Tankanow 1990), three studies scored four (Correa 2005; LaRosa 2003; Vanderhoof 1999), and one study scored

five (Kotowska 2005). The mean Jadad score was 3.2, indicating good quality overall. Allocation concealment was mixed. Five studies demonstrated adequate allocation concealment (Correa 2005; Kotowska 2005; LaRosa 2003; Tankanow 1990; Vanderhoof 1999). Three studies were unclear as to whether allocation concealment was performed properly (Arvola 1999; Benhamou 1999; Jirapinyo 2002) and two studies were inadequate (were not blinded) (Contardi 1991; Erdeve 2004). Kappa statistics were not calculated since there was full agreement between reviewers for both the Jadad and Shultz quality scales.

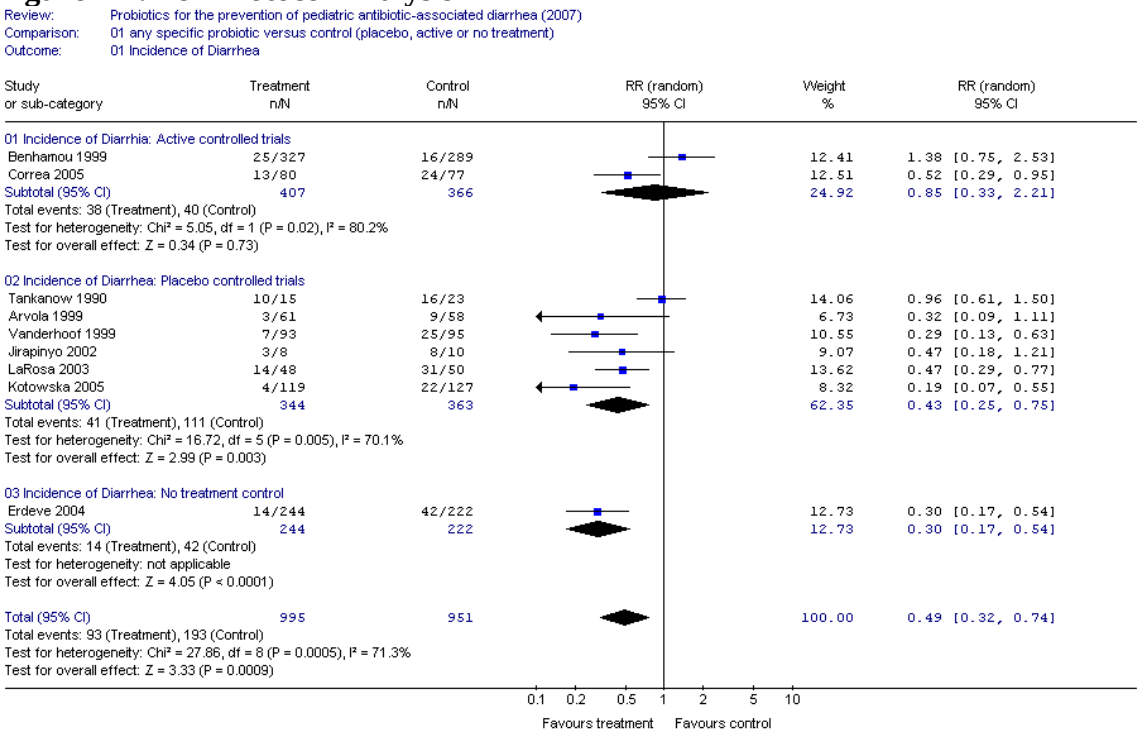
2.7 RESULTS

2.7.1 Incidence of Diarrhea

To allow for a heterogeneous definition of diarrhea, data (as a binary outcome) were included based on the primary authors' definition of the presence/absence of diarrhea. Nine studies (n = 1946) reported incidence of diarrhea. Using a per protocol approach, three placebo-controlled studies showed a statistically significant reduction in the incidence of AAD ($P < 0.05$) (Kotowska 2005; LaRosa 2003; Vanderhoof 1999); one active-controlled study (formula) was statistically significant (Correa 2005), and one no treatment-control study demonstrated statistical significance (Erdeve 2004). Three placebo-controlled studies (Arvola 1999; Jirapinyo 2002; Tankanow 1990) and one active-control (diosmectite) study (Benhamou 1999) did not show statistical significance. One trial showed a non-significant trend toward a positive effect for the control intervention (diosmectite) (Benhamou 1999). The overall pooled results using a per protocol analysis showed that the use of probiotics produced a statistically significant reduction in the incidence of AAD (RR 0.49; 95% CI 0.32 to 0.74; random effects). However, statistical heterogeneity was moderate with respect to percent variability due to between (or inter -) study variability ($I^2 = 71.3\%$) (Higgins 2003). I-squared was interpreted with caution with special attention given to clinical heterogeneity via subgroup testing (below) due to the small number of trials. A sensitivity analysis was conducted using random (RR 0.49; 95% CI 0.32 to 0.74) versus fixed effects models (RR 0.49; 95% CI 0.39 to 0.61) for the incidence of diarrhea, indicating limited differences between the relative risk and corresponding 95% confidence intervals. Nonetheless,

because the I-squared statistic demonstrated moderate heterogeneity within and between studies, a random effects model was used for all remaining statistical analyses. An ITT sensitivity analysis (see ITT explanation below) was not statistically significant (RR 0.90; 95% CI 0.50 to 1.63, $I^2 = 92.9\%$).

Figure 2-1: Per-Protocol Analysis

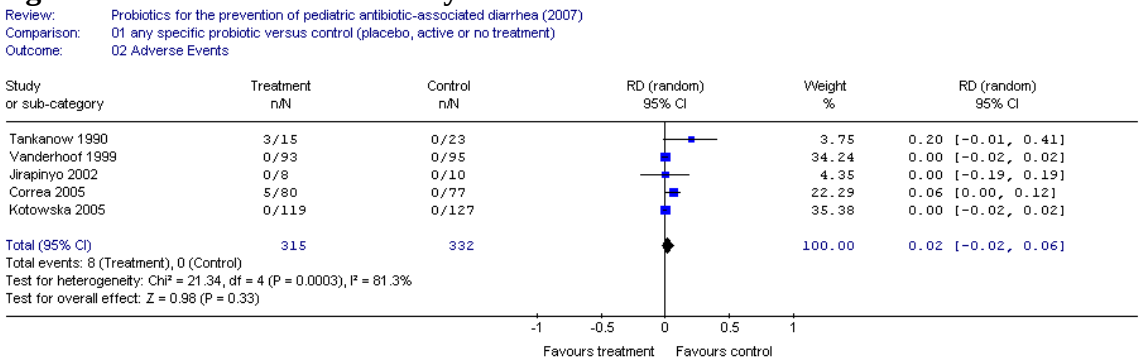


2.7.2 Adverse Events

None of the studies (10/10) defined adverse events *a priori*. Three trials reported no adverse events (Jirapinyo 2002; Kotowska 2005; Vanderhoof 1999). Two trials reported adverse events (Correa 2005; Tankanow 1990). Tankanow 1990 reported 14 adverse events including rash, gas, vomiting, increased phlegm and chest pain. However, for each of the 14 events it was not clear in which group (treatment or control) the adverse events occurred. It appears that the 14 adverse events occurred in three subjects (Tankanow 1990). It was assumed for the meta-analysis that the adverse events were in the treatment group. Correa 2005 reported five subjects with adverse events in the treatment group. These adverse events were related to the tolerability of the formula

supplemented with probiotics. Meta-analysis demonstrated no statistically significant differences in the incidence of adverse events (RD 0.02; 95% CI -0.02 to 0.06).

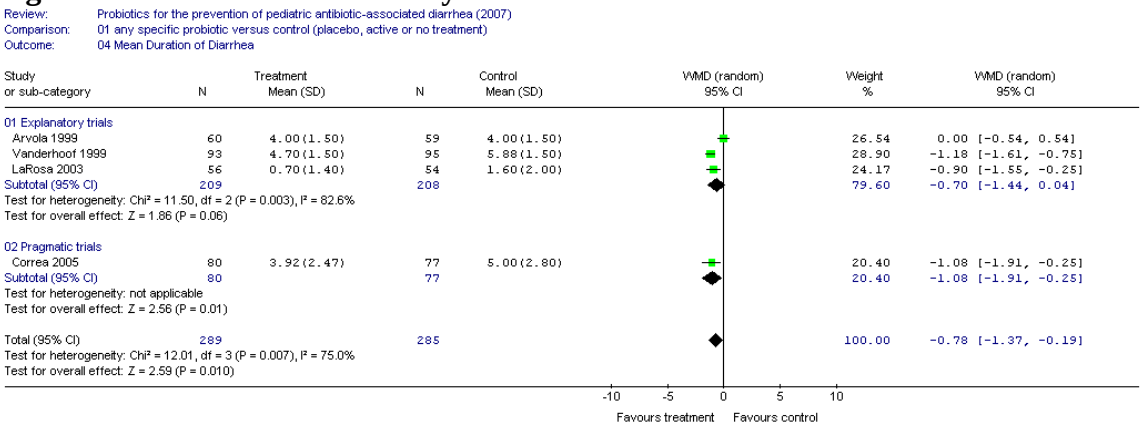
Figure 2-2: Adverse Events Analysis



2.7.3 Mean Duration of Diarrhea

Four studies recorded the mean duration of diarrhea (Arvola 1999; Correa 2005; LaRosa 2003; Vanderhoof 1999). The standard deviation (SD) for one of the trials was not reported and this information was requested from authors with no response (Vanderhoof 1999). The SD was imputed for Vanderhoof 1999 from a study reporting a similar mean duration of diarrhea for treatment and control (Arvola 1999). A post hoc sensitivity analysis was conducted to test the robustness of the mean duration results both before and after imputing data. The WMD was statistically significant before including Vanderhoof 1999 (WMD -0.68; 95% CI -1.31 to -0.08) and after imputing the SD data (WMD -0.78; 95% CI -1.37 to -0.19).

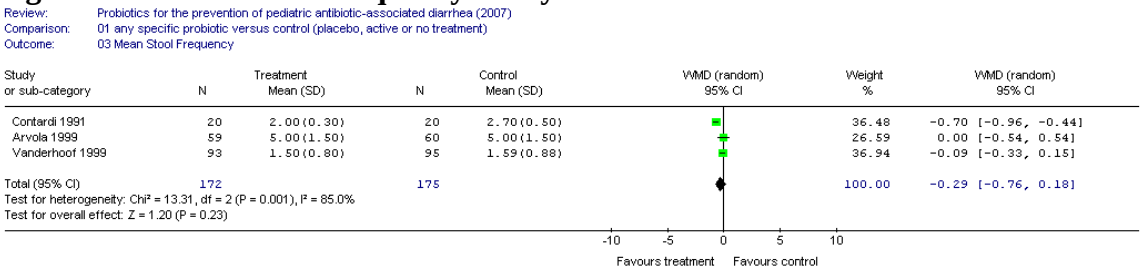
Figure 2-3: Mean Duration Analysis



2.7.4 Mean Stool Frequency

Three RCTs recorded mean stool frequency (Arvola 1999; Contardi 1991; Vanderhoof 1999). SD data were imputed for one study (Arvola 1999). This study reported a range for the mean stool frequency for both treatment and control which was used to impute a SD for each study arm. A post hoc sensitivity analysis was conducted to test the robustness of the mean stool frequency results both before and after imputing data. The WMD excluding Arvola 1999 was -0.39 (95% CI -0.99 to 0.20). After imputing SD data the WMD was -0.29 (95% CI -0.76 to 0.18). Both results were not statistically significant.

Figure 2-4: Mean Stool Frequency Analysis

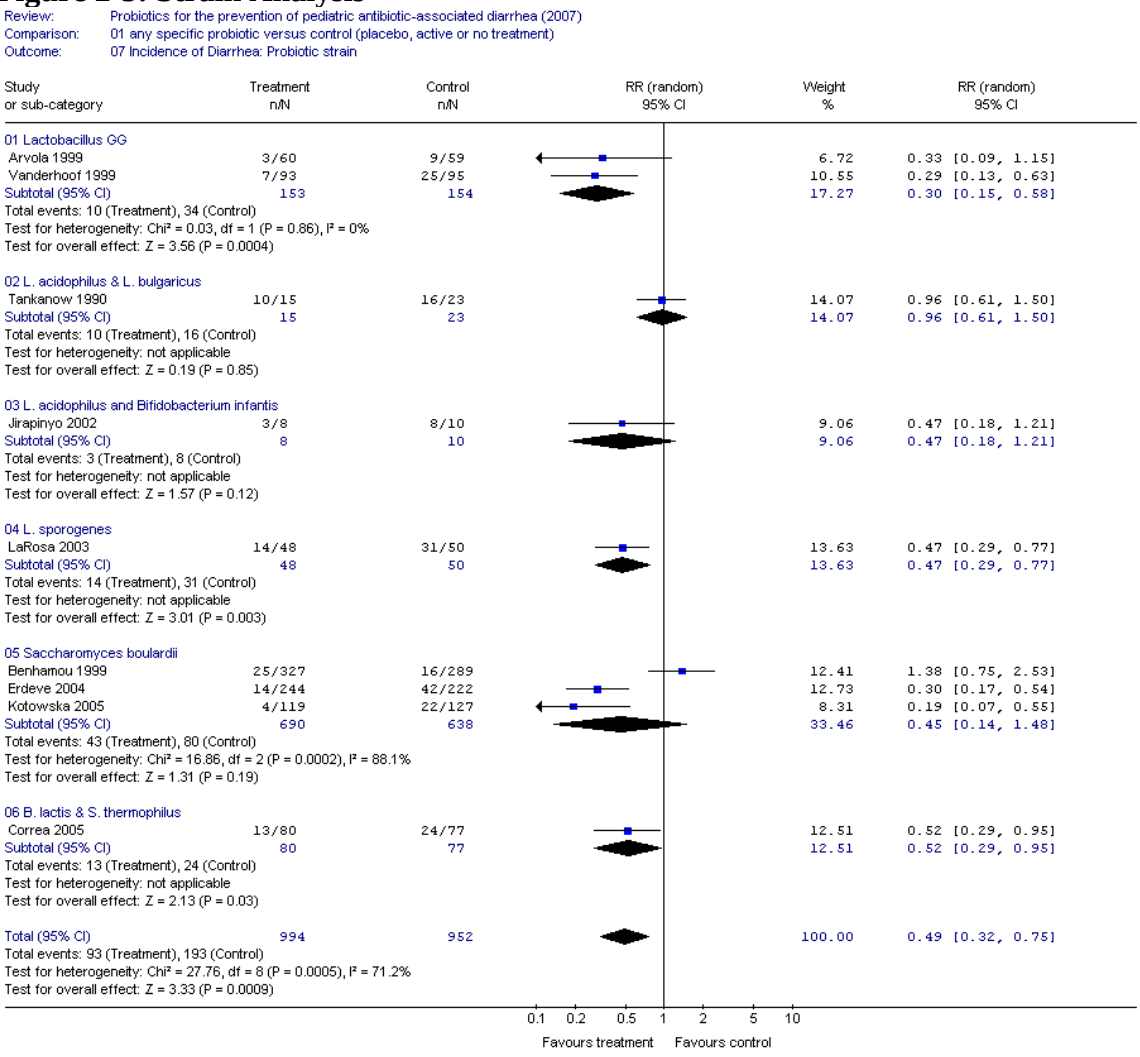


2.8 a priori Subgroups

2.8.1 Probiotic Strain

Two of nine trials administered *Lactobacillus GG* (*Lactobacillus casei spp rhamnosus*) (Arvola 1999; Vanderhoof 1999), while three studied the yeast *Saccharomyces boulardii* (Benhamou 1999; Erdeve 2004; Kotowska 2005). The summary statistic from the *Saccharomyces boulardii* trials (n = 1328) was not statistically significant (RR 0.45; 95% CI 0.14 to 1.48, $I^2 = 88.1\%$). Combined results from two studies (n = 307) were statistically significant indicating a protective effect for *Lactobacillus GG* (RR 0.30; 95% CI 0.15 to 0.58, $I^2 = 0\%$). *Lactobacillus sporogenes* also provided evidence of efficacy (RR 0.47; 95% CI 0.29 to 0.77) for the prevention of childhood AAD (LaRosa 2003).

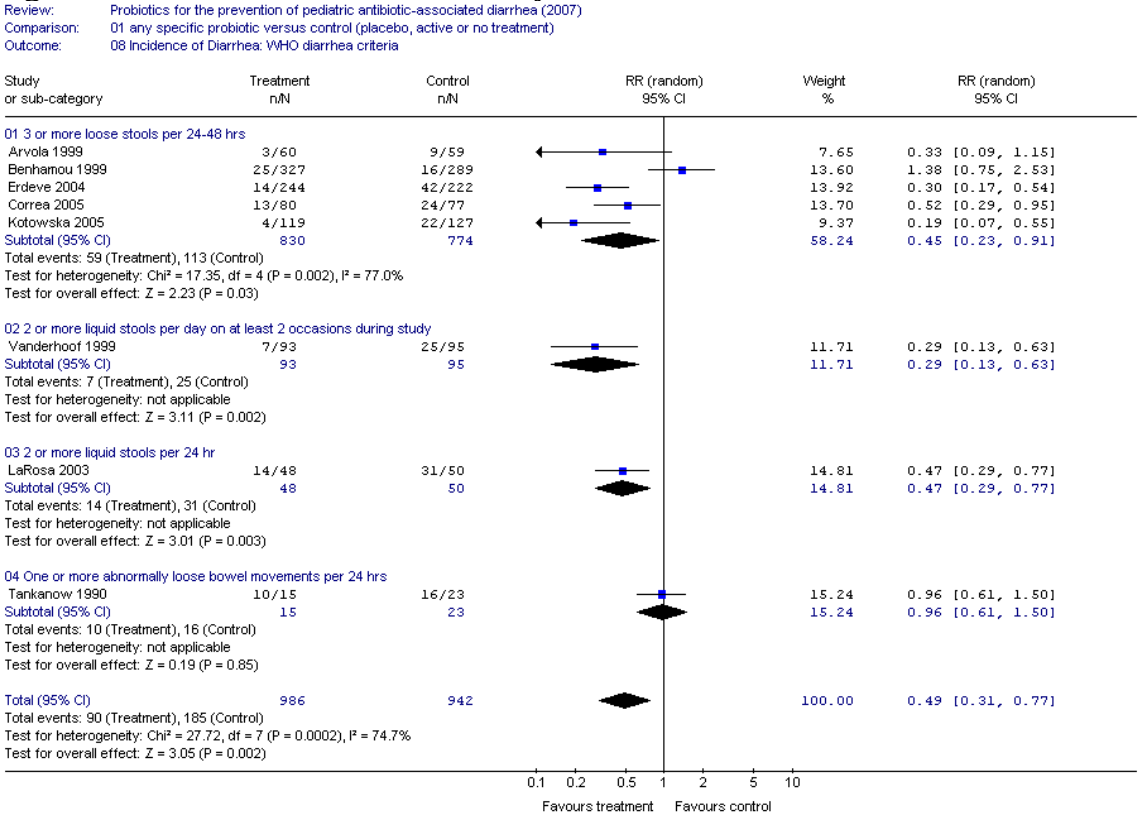
Figure 2-5: Strain Analysis



2.8.2 Definition of Diarrhea

The criteria for diarrhea varied amongst the studies and only eight studies defined diarrhea. Five studies (Arvola 1999; Benhamou 1999; Correa 2005; Erdeve 2004; Kotowska 2005) defined diarrhea synonymous to World Health Organization criteria: three or more abnormally loose bowel movements per 24 to 48 hours (RR 0.45; 95% CI 0.23, 0.91). Tankanow 1990 defined diarrhea as one or more abnormally loose bowel movements per 24 hours (RR 0.96; 95% CI 0.61, 1.50). LaRosa 2003 defined diarrhea as at least two liquid bowel movements per 24 hour period (RR 0.47; 95% CI 0.29, 0.77). Vanderhoof 1999 defined diarrhea as two or more loose stools on two or more occasions throughout the study period (RR 0.29; 95% CI 0.13, 0.63).

Figure 2-6: Definition of Diarrhea Analysis

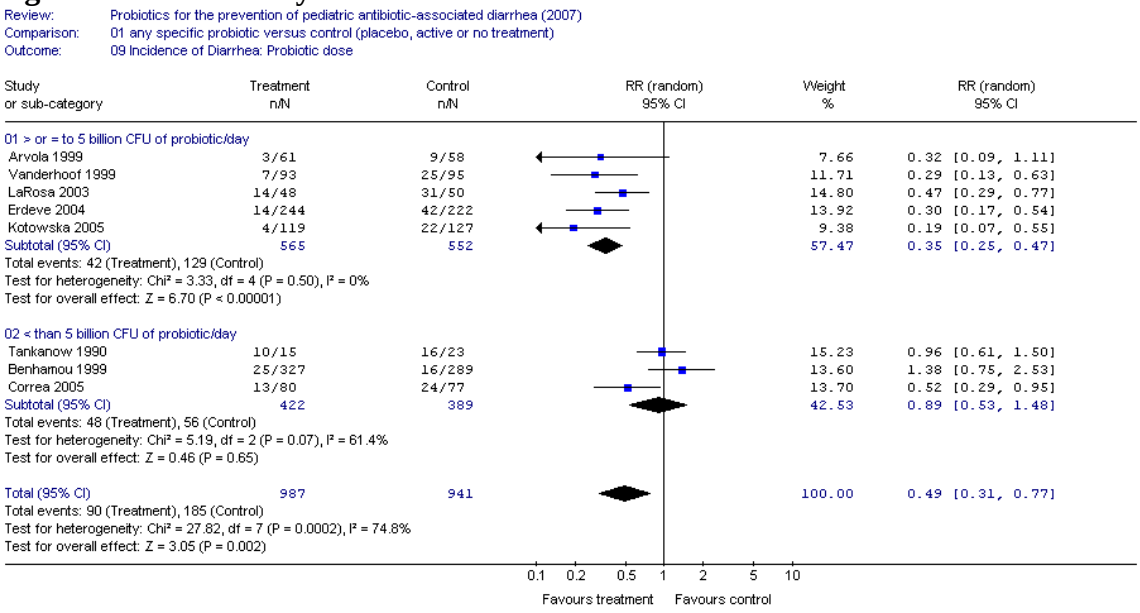


2.8.3 Dose

The daily dosage of probiotic(s) varied greatly (2 to 40 billion CFU/day). Eight of nine studies that reported incidence of diarrhea data provided dosage information (Arvola 1999; Vanderhoof 1999; Erdeve 2004; Kotowska 2005; LaRosa 2003; Tankanow 1990; Benhamou 1999; Correa 2005). The *a priori* subgroup analyses on dose compared <5 billion CFU/day versus ≥ 5 billion CFU/day. Five studies providing children with 5 - 40 billion bacteria/yeast cells per day showed evidence for the preventative effects of probiotics (RR 0.35; 95% CI 0.25 to 0.47) (Arvola 1999; Erdeve 2004; Kotowska 2005; LaRosa 2003; Vanderhoof 1999); whereas, three studies providing <5 billion CFU bacteria/yeast per day: 825 million CFU/day (Correa 2005), 2 billion CFU/day (Tankanow 1990), and 4.5 billion CFU/day (Benhamou 1999) demonstrated non-significant results when combined (RR 0.89; 95% CI 0.53 to 1.48, $I^2 = 61.4\%$). Trials providing greater than five billion CFU of probiotic per day yielded significant results

and no between (or within) study heterogeneity (I^2 = 0%). A chi square test revealed statistically significant dose related heterogeneity (P = 0.0004).

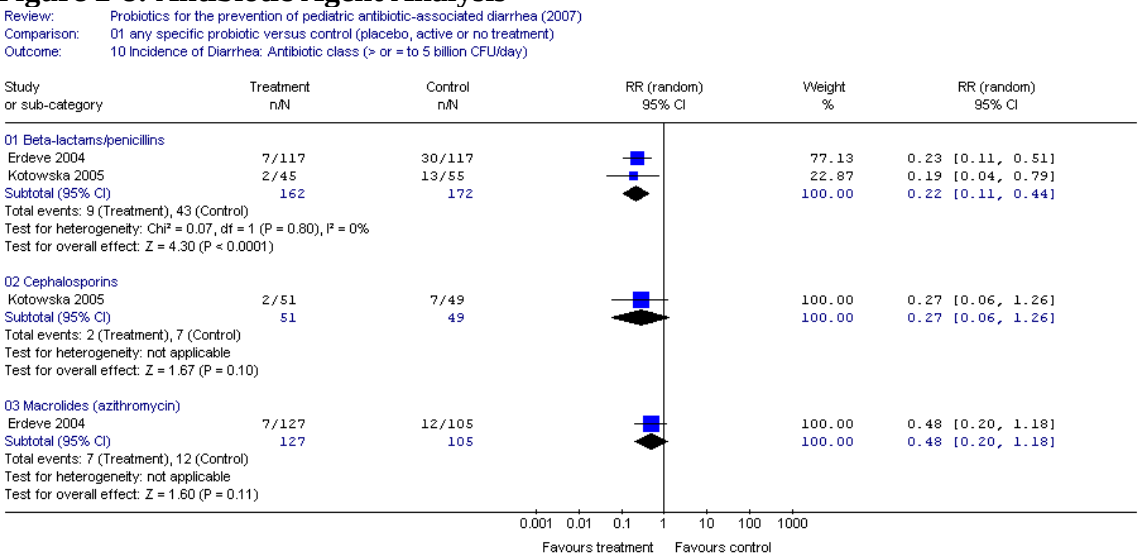
Figure 2-7: Dose Analysis



2.8.4 Antibiotics

Four of nine trials reported adequate details regarding the exposure to antibiotic agents (co-administered with probiotics) (Benhamou 1999; Correa 2005; Erdeve 2004; Kotowska 2005). Antibiotic agents were categorized into antibiotic class (e.g., beta-lactams/penicillins, cephalosporins, macrolides) and a post hoc subgroup analysis was performed on 2/4 trials that administered ≥5 billion CFU/day (Erdeve 2004; Kotowska 2005). In a subgroup of patients given beta-lactams/pencillins only, a statistically significant difference between probiotic treatment and control groups (RR 0.22; 95% CI 0.11 to 0.44) was observed. Only one trial reported AAD incidence rates in patients administered cephalosporins and macrolides both demonstrating non-significant difference (RR 0.27; 95% CI 0.06 to 1.26; RR 0.48; 95% CI 0.20 to 1.18) (Kotowska 2005; Erdeve 2004) respectively.

Figure 2-8: Antibiotic Agent Analysis



2.9 a priori Sensitivity Analyses

2.9.1 Random vs Fixed Effects

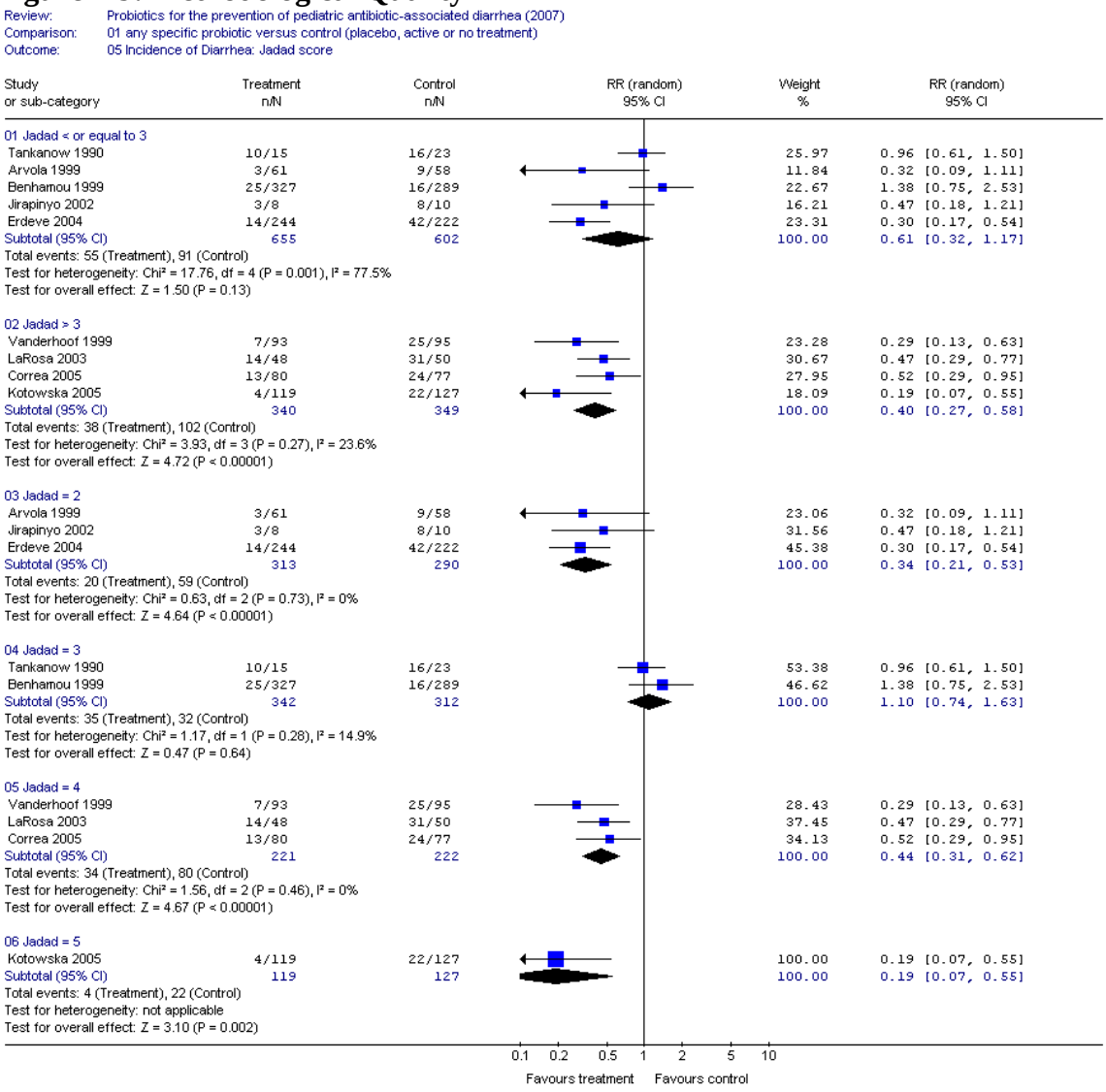
A sensitivity analysis using random (RR 0.49; 95% CI 0.32 to 0.74) versus fixed effects models (RR 0.49; 95% CI 0.39 to 0.61) for the incidence of diarrhea, indicated limited differences between the relative risk and corresponding 95% confidence intervals. Nonetheless, because the I-squared statistic demonstrated moderate heterogeneity within and between studies, a random effects model was used for all remaining statistical analyses.

2.9.2 Quality

A sensitivity analysis was conducted on the methodological quality in studies reporting incidence of diarrhea using the Jadad Scale 0-5 (a score of less than three indicates a poor quality study and a score of five indicates maximum quality) (Jadad 1996). The overall quality of the trials reporting on the primary outcome (incidence of diarrhea) was good with a mean Jadad score of 3.2 out of 5. A subgroup analysis of trials with a Jadad score of ≤3 versus >3 demonstrated a relationship between quality and efficacy. The subgroup of five trials (n = 1257) with a Jadad score >3 were shown to be more efficacious than the four trials (n = 689) with a Jadad score ≤3 (RR 0.40; 95% CI 0.27 to 0.58, I^2 = 23.6% versus RR 0.61; 95% CI 0.32 to 1.17, I^2 = 77.5%) respectively. The chi square test for

heterogeneity (five studies with a Jadad score ≤3 versus four studies with a Jadad score >3 was significant (P = 0.006). However, summary statistics from the sub-categories for Jadad = 2 (Arvola 1999; Erdeve 2004; Jirapinyo 2002), Jadad = 3 (Benhamou 1999; Tankanow 1990), Jadad = 4 (Correa 2005; LaRosa 2003; Vanderhoof 1999), Jadad = 5 (Kotowska 2005) showed no trend between quality and sub-category point estimates. In other words, as quality (e.g., randomization, blinding) increased, the therapeutic effect of probiotics did not decrease.

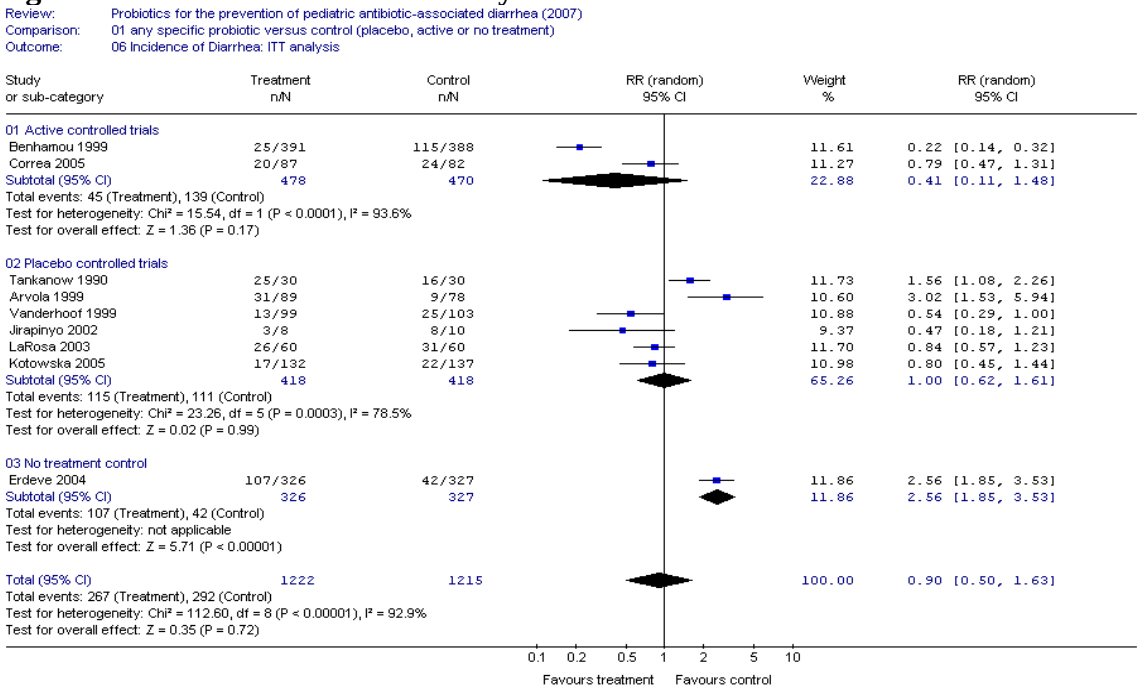
Figure 2-9: Methodological Quality



2.9.3 Intention-to-Treat Analysis

There were 2437 pediatric subjects originally randomized in the nine trials reporting on the primary outcome (incidence of diarrhea). Eight of nine trials reported loss to follow-up of which four reported substantial attrition concerns. Loss to follow-up was 21%, 29%, 29% and 37% in the Arvola 1999, Benhamou 1999, Erdeve 2004, and Tankanow 1990 studies respectively. A highly conservative "extreme-case" ITT analysis was calculated. If all the data on the number of patients randomized to each group and the number of dropouts from each group were available all patients lost to the group that fared better were assigned a poor outcome (diarrhea), and all patients lost to the group that fared worse were assigned a good outcome (no diarrhea). If no information on the number of patients randomized to each group, or the number of dropouts from each group (e.g., not reported in the published trial or unsuccessful contact with authors) was available, it was assumed that the denominators were as even as possible (block randomization), and all patients lost to the group that faired better were assigned a poor outcome (diarrhea), and all lost to the group that faired worse (i.e., Erdeve 2004) were assigned a good outcome (no diarrhea). The ITT analysis was not statistically significant (RR 0.90; 95%CI 0.50, 1.63, $I^2 = 92.9\%$).

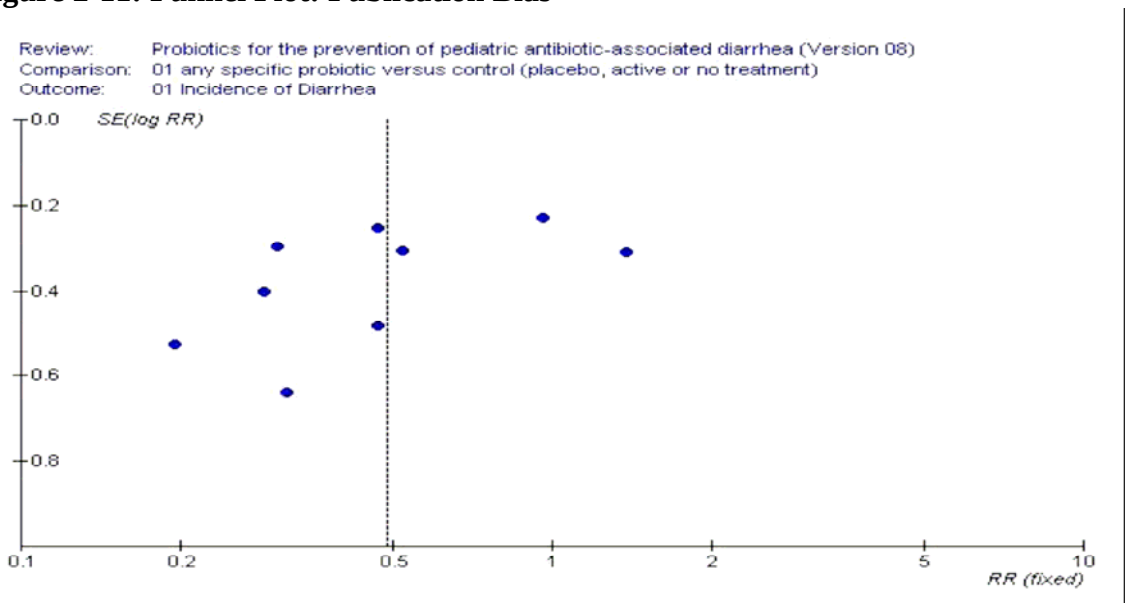
Figure 2-10: Intention-to-Treat Analysis



2.9.4 Publication Bias

There were too few trials reporting on the incidence of diarrhea (n = 9) to properly analyze for publication bias. However, a funnel plot analysis shows asymmetry of the funnel for the relationship between risk ratio and standard error. Six studies lie on the left of the funnel and three on the right, suggesting publication bias. RCTs favouring probiotics seem more likely to be published (e.g., trials to the left of the line of no effect indicating that probiotics reduce the incidence of diarrhea) than those with inconclusive results. Small studies with negative or inconclusive results seem hard to find despite extensive searching in the "grey literature".

Figure 2-11: Funnel Plot: Publication Bias



2.10 DISCUSSION

The objective of this review was to determine if the co-administration of probiotics with antibiotics prevents or ameliorates AAD in children. This review included two trials (Benhamou 1999; Erdeve 2004) not included in previous meta-analyses of probiotics for AAD (Johnston 2006; McFarland 2006; Sazawal 2006). Like the previous meta-analyses, this study used a per protocol analysis for primary and sub-group analysis. Per protocol pooled results of nine studies reporting on the incidence of diarrhea suggest that

probiotics are effective for preventing AAD (RR 0.49; 95% CI 0.32 to 0.74, $I^2 = 71.3\%$). The number needed to treat to prevent one case of diarrhea is ten (NNT 10; 95% CI 7 to 18). Regarding safety, no trials reported a serious adverse event although only 5/10 trials included reported on adverse events. Meta-analysis demonstrated no significant differences in the incidence of any adverse events between treatment and control. Of note, no trial defined *a priori* what was considered to be an adverse event. The results of this review are consistent with the results of four earlier meta-analyses in the general population: D'Souza et al. (9 RCTs included) used a per-protocol analysis (OR 0.37; 95% CI 0.26, 0.53) (D'Souza 2002); whereas, Cremonini et al. (7 RCTs) reported that trials with a loss to follow-up of 15% or greater would be excluded (RR 0.40; 95% CI 0.27, 0.57) (Cremonini 2002). Sazawal 2006 (19 RCTs) and McFarland 2006 (25 RCTs) also reported similar results (RR 0.48; 95% CI 0.35, 0.65, $I^2 = 53\%$ and RR 0.43; 95% CI 0.31, 0.58, $p < 0.001$) respectively. The per-protocol results similarly reflect meta-analyses conducted in the general population further validating that probiotics hold promise not only in the adult population but also in children. However, further trials need to better define what probiotic(s), at what dose, in what age group, co-administered with what antibiotic, at what duration demonstrate efficacy.

Concerning the secondary outcomes of mean duration of diarrhea (four trials, n = 574) and mean stool frequency, using a per protocol analysis probiotics decreased the mean duration of diarrhea by approximately three quarters of a day (WMD -0.78; 95% CI -1.37, -0.19), a statistically significant difference. However, the differences in mean stool frequency were not statistically significant (WMD -0.29 95% CI -0.76, 0.18) between the treatment and control groups. Of the ten pediatric trials included here, only two conducted *C. difficile* stool assays (Arvola 1999; Kotowska 2005). Considering that *C. difficile* diarrhea is the most serious adverse event associated with AAD, and that 26% to 50% of AAD can be attributed to *C. difficile* (McFarland 1998), future trials should better address the efficacy of probiotics in preventing AAD in children caused by enteropathogens; and, in particular, *C. difficile*-associated diarrhea.

The ITT sensitivity analysis did not reveal any statistically significant differences (RR 0.90; 95% CI 0.50, 1.63, $I^2 = 92.9\%$). Four studies with very large loss to follow-up (range 21% to 37%) may explain why the ITT results were not statistically significant (Benhamou 1999; Arvola 1999; Erdeve 2005; Tankanow 1990). In one study these losses to follow-up may have been related to feasibility issues, e.g. 3 month follow-up with families was hard to accomplish (Personal communication, Arvola 2005). Although no consensus exists about how to account for missing data in an ITT analysis (Hollis 1999), ITT has a number of advantages over per protocol analyses. Intention-to-treat is intended to minimize selection bias and provide an unbiased estimate of treatment effect. ITT also minimizes bias due to non-compliance, protocol deviations, and loss to follow-up (Hollis 1999). By accepting that non-compliance and protocol deviations are likely to occur in real clinical situations, ITT analysis preserves randomization integrity, enhances the external validity of the results by controlling for the removal of non-compliers from the analysis and avoids over optimistic statements regarding the efficacy of an intervention (Heritier 2003). However an ITT analysis may not be robust if loss to follow-up rates are high. Since this review involved meta-analysis of nine trials where 4 of 9 trials had high losses to follow-up, it was felt the validity of ITT analysis was in question. Thus a per-protocol analysis was performed as other authors in this field have done (Cremonini 2002; D'Souza 2002; McFarland 2006; Sazawal 2006).

Statistical heterogeneity was moderate in the per-protocol analysis (71.3%) and large in the ITT analysis (92.9%). The quantification of heterogeneity is only one component of a wider investigation of variability across studies, perhaps the most important being diversity in clinical aspects which was explored via subgroup analysis (Figures 2-2 and 2-3) (Higgins 2003). The per-protocol subgroup analyses explored potential reasons for statistical heterogeneity including probiotic dose and strain. For instance, probiotics may be compared to antibiotics in that both have many different agents or classes that perform similar biological actions albeit via different mechanisms of action. In addition, these actions may or may not be dose-dependent. For example, different antibiotic classes have different mechanisms of action (beta-lactams inhibit cell wall synthesis while macrolides inhibit bacterial protein synthesis), and are usually dose dependent. With

regard to probiotics, their mechanism of action, and the most appropriate dose for possible prevention of AAD are not known. Some probiotics are well studied such as *S. boulardii* and their mechanism of action for preventing AAD may involve pathogenic microbial antagonism or their ability to help regulate water and electrolyte exchanges in the intestines (Girard 2003). However, potential mechanisms of action of other probiotics included in this review are not so well studied. In the absence of any reliable data relative to which mechanism of action(s) are involved in preventing AAD, the possibility of strain dependent or dose dependent effects must be considered. Randomized controlled trials on the strain(s) and doses themselves are required before efficacy claims can be made. Although data were pooled for analyses, statistical heterogeneity was apparent. This meta-analysis is the first to conduct *a priori* strain and dose subgroup analyses to explore statistical heterogeneity.

No dose ranging studies have been reported to determine the minimal effective dose of a probiotic for preventing pediatric AAD. In an effort to separate trials of sub-therapeutic doses from others, a subgroup analysis of ≥5 billion CFU/day versus <5 billion CFU/day was chosen. LaRosa 2003 demonstrated efficacy for probiotics for preventing pediatric AAD using 5.5 billion CFU/day of *Lactobacillus sporgenes*. While dosage recommendations on the labels of probiotic products available in health food stores varies from 1-40 billion CFU/day, doses in the lower range may not colonize the intestine (Raza 1995). The subgroup analyses provided preliminary evidence that probiotic dose may be responsible for the observed clinical and statistical heterogeneity (I^2 = 71.3%). A per protocol subgroup analysis on five studies providing ≥5 billion CFU/day (5 - 40 billion single strain probiotics per day (*Lactobacillus GG*, *Lactobacillus sporogenes, Saccharomyces boulardii*) showed strong evidence for the preventative effects of probiotics (0.35; 95% CI 0.25 to 0.47, I^2 = 0%). The chi square test for potential dose related heterogeneity was statistically significant (P = 0.0004) and may explain the moderate statistical heterogeneity (I^2 = 71.3%) observed regarding the primary outcome, the incidence of diarrhea. Considering these analyses, single strain interventions *Lactobacillus GG, Lactobacillus sporogenes* or *Saccharomyces boulardii* (5 - 40 billion CFU/day) co-administered with antibiotics are worthy of further investigation for the

prevention of pediatric AAD. The organism originally termed *L. sporogenes* by Horowitz-Wlassowa in 1933 has been reclassified as *Bacillus sporogenes* and moved into the B. coagulans group (De Vecchi 2006). Therefore, assumptions about safety and benefits derived from the extensive literature on the genus *Lactobacillus* may not apply to this species (De Vecchi 2006). Although a few small studies involving *L. sporogenes/B coagulans* have demonstrated preliminary efficacy and safety there is very limited scientific evidence on *B. coagulans*. For this reason the use of the wrong nomenclature of L. sporogenes becomes questionable, since it seems to try to benefit from association with the extensive literature on the safety and health benefits of the genus *Lactobacillus* (De Vecchi 2006).

The relationship of the effect of probiotics to the class of antibiotics used is of interest. Although almost all antibiotics cause gastrointestinal side effects; antibiotics directed against anerobes (e.g., aminopencillins, cephalosporins and clindamycin) have been cited as having the highest risk of diarrheal side effects (McFarland 1998). Adequate data were not available to do subgroup analysis on each of these anaerobic agents separately. Sufficient data were not available to analyze the effects of these antibiotics together versus other antibiotics. To avoid multiple testing and the potential for false positives only trials that provided ≥5 billion CFU/day and reported data on the incidence of diarrhea by antibiotic class were used for this analysis. Only two trials fulfilled these criteria for beta-lactams/penicillins (Erdeve 2004; Kotowska 2005). For cephalosporins and macrolides, only one trial reported data on each antibiotic class. Although probiotics significantly reduced the incidence of AAD when administered with beta-lactams/penicillins (vs. non-significant results when probiotics were co-administered with cephalosporins or macrolides); drawing conclusions based on such limited data is inappropriate. Future trials should consider post-hoc analysis of the effects of the aminopenicillins, cephalosporins and clindamycin or together versus other antibiotics.

A sensitivity analysis was conducted on the methodological quality in studies reporting incidence of diarrhea (average Jadad score of 3.2 out of 5) demonstrating that quality did not influence the magnitude of effect. Of interest, a relationship between quality and

efficacy was noted: the subgroup of five trials (n = 1257) with a Jadad score ≥3 were shown to be more efficacious than the four trials (n = 689) with a Jadad score <3 (RR 0.40; 95% CI 0.27 to 0.58, I^2 = 23.6% versus RR 0.61; 95% CI 0.32 to 1.17, I^2 = 77.5%) respectively.

There are two additional issues worthy of further consideration when reviewing pediatric trials of AAD: the effect of age and the definition of AAD. Whereas four previous meta-analyses included both children and adults, with the majority of included trials comprised of adults (Cremonini 2002; D'Souza 2002; McFarland 2006; Sazawal 2006), this review was restricted to the pediatric population (aged two weeks to 15 years). The difficulty in accurately measuring frequency or consistency of diarrhea in diapered infants versus that of children and adolescents may make it difficult to detect differences between treatment and control groups. For example, infants not only have bowel movements more frequently (Fontana 1989), there stools may also be looser in consistency than those of older children, which may make infant stools more likely to meet the investigators definition of diarrhea. Furthermore, it may be difficult to accurately detect the frequency and consistency of stools in diapered infants. Although randomized trials may evenly distribute infants between treatment and control groups so as to eliminate this potential confounder, the considerable differences in bowel habits in children at varying ages may obscure the effect of probiotics for this population, and should be taken into consideration when planning pediatric trials. Trials did not report adequate data to analyze the effect of age on probiotics ability to prevent AAD.

The definition of AAD in the included studies was heterogeneous. Of the eight trials that defined diarrhea, the incidence of AAD in the control groups ranged from 6 - 80%. Definitions of AAD included: "one or more abnormally loose bowel movements per day," (Tankanow 1990) as opposed to "at least three watery or loose stools per day for a minimum of two consecutive days" (Arvola 1999). One trial did not provide a definition (Jirapinyo 2002). Although a trend between the definition of diarrhea used in the included trials and efficacy was not found (increased probiotic efficacy in trials with more conservative definitions of AAD (e.g., one or more loose stools per day), the

development of an outcome measure for stool frequency and consistency that is valid and sensitive to change in children pre and post antibiotic administration would be ideal. A survey of clinicians and parents may be helpful in defining what is considered a clinically meaningful reduction in AAD and at what point would parents and/or clinicians would consider co-administering a probiotic.

This systematic review has several strengths. The search strategy for this review was more comprehensive than that used in previous reviews of probiotics for AAD in the general population. Although some evidence of publication bias was found, several strategies were implemented to control for the effect of this bias within the review, including explicit searches in multiple databases, and inclusion of any relevant trial irrespective of language or format of publication (including abstracts). *A priori* subgroup (i.e., probiotic strain, dose) and sensitivity analyses (per-protocol versus ITT) were conducted to further explore the data. Finally, trials were included that assessed probiotics versus no treatment, placebo or standard conventional care for AAD (i.e., diosmectite, infant formula) (Benhamou 1999; Correa 2005) and demonstrated that probiotic treatment had higher efficacy versus no treatment control (RR 0.30; 95% CI 0.17, 0.54) than when compared to placebo (RR 0.43; 95% CI 0.25, 0.75). In addition, probiotic treatment was demonstrated to be beneficial versus placebo control (RR 0.43; 95%CI 0.25, 0.75), but not effective versus active controls (diosmectite, infant formula) (RR 0.85; 95%CI 0.33, 2.21). This finding is in keeping with the literature: therapies when compared to inactive controls have a higher probability of demonstrating benefit than when a therapy is compared to active controls (Djulbegovic 2000). The results relative to the control group(s) further validate the findings of this review. In addition, the inclusion of active control trials provided effectiveness data increasing the external validity of the results of this review (McMahon 2002).

This review also has a number of limitations. The United European Gastroenterology Week, North American Society for Pediatric Gastroenterology, and Hepatology and Nutrition conference proceedings were not searched. Some readers may question the pooling of different probiotic strains. In keeping with the justification for the combining

of probiotic strains used in two trials included in this review (Tankanow 1990 administered both *L. acidophilus* with *L. bulgaricus*; Jirapinyo 2002 administered both *L. acidophilus* with *B. infantis*), data were pooled because the probiotics used in each trial share the recommended characteristics of a viable probiotic: non-pathogenic properties (noting that further study is needed on *L. sporogenes*), the ability to survive transit through the gastrointestinal tract, adherence to intestinal epithelium, colonization in the intestinal tract, production of antimicrobial substances, and a good shelf life in food or powdered form (Goldin 1998).

The ITT analysis did not reveal a statistically significant benefit of probiotics for the prevention of pediatric AAD. This may have been due to substantial losses to follow-up in the included studies. In contrast, evidence limited to high quality trials (particularly for *Lactobacillus GG* and *Saccharomyces boulardii* at 5-40 billion CFU/day) suggests that probiotics may be effective. Future studies should involve these probiotic strains and doses. The effects of age (i.e., infant vs older children), and antibiotic duration (e.g., 5 days versus 10 days), should be determined, losses to follow-up should be limited and adverse events reported. In addition, trials would benefit from a validated primary outcome measure for AAD that is sensitive to change and reflects what clinicians, parents, and children consider important with regards to stool frequency and consistency. The current data are promising, but inconclusive. It is premature to draw a conclusion about the efficacy and safety of probiotics for pediatric AAD until such trials are completed.

Reviewers' Conclusions

Implications for Practice

1. The current data are promising, but inconclusive. The use of *Lactobacillus GG,* or *Saccharomyces boulardii* at a dose of 5 to 40 billion CFU/day appear to hold promise as an option for co-administration with antibiotics, but there is insufficient evidence to recommend their use at this time. The evidence for safety and efficacy of *Lactobacillus sporogenes* comes from a single trial.

2. No serious adverse events were reported in the included studies. However safety is better assessed in population-based samples.

Implications for Research

Future pediatric trials should use:

1. probiotic strains and doses with the most promising evidence (e.g., *Lactobacillus GG*, or *Saccharomyces boulardii* at 5 - 40 billion CFU/day);
2. determine the effect of age (i.e., infant, child, adolescent), and antibiotic duration (e.g., 5 days vs. 10 days) on probiotic efficacy;
3. define potential adverse events *a priori* and monitor for these adverse reactions accordingly;
4. limit losses to follow-up and analyze results using ITT analysis;
5. a validated primary outcome measure for antibiotic-associated diarrhea that is sensitive to change and reflects what clinicians, parents, and children consider important with regards to stool frequency and consistency; and

REFERENCES

Arvola T, Laiho K, Torkkeli S, Mykkanen H, Salminen S, Maunula L et al. Prophylactic Lactobacillus GG reduces antibiotic-associated diarrhea in children with respiratory infections: a randomized study. Pediatrics 1999;104:e64.

Bartlett JG, Chang TW, Gurwith M, Gorbach SL, Onderdonk AB. Antibiotic-associated pseudomembranous colitis due to toxin producing clostridia. N Engl J Med. 1978;298:531-534.

Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50:1088-1101.

Benhamou PH, Berlier P, Danjou G, Plique O, Jessueld D, DuPont C. Antibiotic-associated diarrhoea in children: A computer monitored double-blind outpatients trial comparing a protective and a probiotic agent. Med Chir Dig. 1999;28:163-168.

Bezkorovainy A. Probiotics: determinants of survival and growth in the gut. Am J Clin Nutr. 2001;73(2 Suppl):399S-405S.

Brunser O, Gotteland M, Cruchet S, Figueroa G, Garrido D, Steenhout P. Effect of a milk formula with prebiotics on the intestinal microbiota of infants after an antibiotic treatment. Pediatr Res. 2006;59:451-456.

Chapoy P. Traitement des diarrhees aigues infantiles: essai controle de Saccharomces boulardii [Treatment of infantile diarrhea: controlled trial of Saccharomyces boulardii]. Ann Pediatr (Paris) 1985;32:561-563.

Contardi I. Batterioterapia orale quale prevenzione della diarrea da antibiotici in eta pediatrica [Oral bacterial therapy in prevention of antibiotic-induced diarrhea in childhood]. Clin Ter. 1991;136:409-413.

Correa NB, Peret Filho LA, Penna FJ, Lima FM, Nicoli JR. A randomized formula controlled trial of Bifidobacterium lactis and Streptococcus thermophilus for prevention of antibiotic-associated diarrhea in infants. J Clin Gastroenterol. 2005;39:385-389.

Cremonini F, Di Caro S, Nista EC, Bartolozzi F, Capelli G, Gasbarrini G, et al. Meta-analysis: the effect of probiotic administration on antibiotic associated diarrhea. Aliment Pharmacol Ther. 2002;16:1461-1467.

Czerwionka-Szaflarska M, Kuczynska R, Mierzwa G, Bala G, Murawska S. Effect of probiotic bacteria supplementation on the tolerance of helicobacter pylori eradication therapy in children and youth. Pediatria Polska. 2006;81:334-341.

D'Souza AL, Rajkumar C, Cooke J, Bulpitt CJ. Probiotics in the prevention of antibiotic associated diarrhoea: meta-analysis. BMJ 2002;324:1361-6.

De Vecchi E, Drago L. Lactobacillus sporogenes or bacillus coagulans: Misidentification or mislabelling? Int J Probiot Prebiot. 2006;1:3-10.

Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M; Davey Smith G, and Altman DG, editor(s). Systematic reviews in health care. 2nd edition. London, UK: BMJ Books, 2001:300.

Dickersin K, Scherer R, Lefebvre C. Indentifying relevant studies for systematic reviews. BMJ 1994;309:1286-1291.

Djulbegovic B, Lacevic M, Cantor A, Fields KK, Bennett CL, Adams JR, et al. The uncertainty principle and industry-sponsored research. Lancet 2000;356:635-8.

Duggan C, Penny ME, Hibberd P, Gil A, Huapaya A, Cooper A, et al. Oligofructose-supplemented infant cereal: 2 randomized, blinded, community-based trials in Peruvian infants. Am J Clin Nutr. 2003;77:937-42.

Duman DG, Bor S, Ozutemiz O, Sahin T, Oguz D, Istan F, et al. Efficacy and safety of Saccharomyces boulardii in prevention of antibiotic-associated diarrhoea due to Helicobacterpylori eradication. Eur J Gastroenterol Hepatol. 2005;17:1357-1361.

Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000;56(2):455-463.

Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315(7109):629-34.

Elstner CL, Lindsay AN, Book LS, Matsen JM. Lack of relationship of Clostridium difficile to antibiotic-associated diarrhea in children. Pediatr Infect Dis. 1983;2:364-366.

Erdeve O, Tiras U, Dallar Y. The probiotic effect of Saccharomyces boulardii in a pediatric age group. J Trop Pediatr. 2004;50:234-236.

Erdeve O, Tiras U, Dallar Y, Savas S. Saccaromyces boulardii and antibiotic-assoiciated diarrhoea in children. Aliment Pharmacol Ther. 2005;21:1507-1508.

Fahey T, Stocks N, Thomas T. Quantitative systematic review of randomised controlled trials comparing antibiotic with placebo for acute cough in adults. BMJ 1998;316:906-910.

Fedorak RN, Madsen KL. Probiotics and prebiotics in gastrointestinal disorders. Curr Opin Gastroenterol. 2004;20:146-155.

Fontana M, Bianchi C, Cataldo F, Conti Nibali S, Cucchiara L, Gobio Casali L, et al. Bowel frequency in healthy children. Acta Paediatr Scand. 1989;78:682-684.

Gibson GR. Dietary modulation of the human gut microflora using prebiotics. Br J Nutr. 1998;80:S209-212.

Girard P, Pansart Y, Lorette I, Gillardin JM. Dose-response relationship and mechanism of action of Saccharomyces boulardii in castor oil-induced diarrhea in rats. Dig Dis Sci. 2003;48:770-774.

Gismondo MR, Drago L, Lombardi A. Review of probiotics available to modify gastrointestinal flora. Int J Antimicrob Agents 1999;12:287-292.

Gogate A, De A, Nanivadekar R, Mathur M, Saraswathi K, Jog A, et al. Diagnostic role of stool culture and toxin detection in antibiotic associated diarrhoea due to Clostridium difficile in children. Indian J Med Res. 2005;122:518-524.

Goldin BR. Health benefits of probiotics. Br J Nutr. 1998;80:S203-S207.

Hamilton-Miller JM, Shah S, Winkler JT. Public health issues arising from microbiological and labelling quality of foods and supplements containing probiotic microorganisms. Public Health Nutr. 1999;2:223-229.

Hata D, Yoshida A, Ohkubo H, Mochizuki Y, Hosoki Y, Tanaka R, et al. Meningitis caused by bifidobacterium in an infant. Pediatr Infect Dis J. 1988;7:669-671.

Havenaar R, Huis in't Veld JHT. Probiotics: a general view. In: Wood BJB, editor. Lactic acid bacteria in health and disease. Vol 1 of series, The lactic acid bacteria. Amsterdam: Elsevier Applied Science Publishers 1992.

Heritier SR, Gebski VJ, Keech AC. Inclusion of patients in clinical trial analysis: the intention-to-treat principle. Med J Aust. 2003;179:438-440.

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-560.

Hollis S, Campbell F. What is meant by intention-to-treat analysis? Survey of published randomised controlled trials. BMJ 1999;319:670-674.

Jadad AR, Moore RA, Carrol D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996;17:1-12.

Jirapinyo P, Densupsoontorn N, Thamonsiri N, Wongarn R. Prevention of antibiotic-associated diarrhea in infants by probiotics. J Med Assoc Thai. 2002;85:S739-S742.

Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. Ann Intern Med. 2001;135:982-989.

Kotowska M, Albrecht P, Szajewska H. Saccharomyces boulardii in the prevention of antibiotic-associated diarrhoea in children: a randomized double-blind placebo-controlled trial. Aliment Pharmacol Ther. 2005;21:583-590.

Land MH, Rouster-Stevens K, Woods CR, Cannon ML, Cnota J, Shetty AK. Lactobacillus sepsis associated with probiotic therapy. Pediatrics 2005;115:178-181.

LaRosa M, Bottaro G, Gulino N, Gambuzza F, Di Forti F, Ini G, et al. Prevention of antibiotic-associated diarrhea with Lactobacillus sporogens and fructo-oligosaccharides in children: A multicentric double-blind vs placebo study. Minerva Pediatr. 2003;55:447-552.

Lei V, Friis H, Michaelsen KF. Spontaneously fermented millet product as a natural probiotic treatment for diarrhoea in young children: an intervention study in northern Ghana. Int J Food Microbiol. 2006;110:246-253.

Lilley DM, Stillwell RH. Probiotics: growth promoting factors produced by microorganisms. Science 1965;147:747-748.

Mackay AD, Taylor MB, Kibbler CC, Hamilton-Miller JM. Lactobacillus endocarditis caused by a probiotic organism. Clin Microbiol Infect. 1999;5:290-292.

Madsen KL. The use of probiotics in gastrointestinal disease. Can J Gastroentrol. 2001;15:817-822.

McFarland LV. Epidemiology, risk factors and treatments for antibiotic-associated diarrhea. Dig Dis. 1998;16:292-307.

McFarland LV. Can Saccharomyces boulardii prevent antibiotic-associated diarrhea in children? Nat Clin Pract Gastroenterol Hepatol. 2005;2:262-263.

McFarland LV. Meta-analysis of probiotics for the prevention of antibiotic associated diarrhea and the treatment of Clostridium difficile disease. Am J Gastroenterol. 2006;101:812-822.

Michielutti F, Bertini B, Presciuttini B, Andreotti G. Valutazione clinica di un nuovo batterioterapico orale in pazienti di eta pediatrica con dirrea acuta [Clinical assessment of a new oral bacterial treatment for children with acute diarrhea]. Minerva Med. 1996;87:545-550.

Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomized trials effect estimates of intervention efficacy reported in meta-analyses? Lancet 1998;352:609-613.

Piarroux R, Millon L, Bardonnet K, Vagner O, Koenig H. Are live saccharomyces yeasts harmful to patients? Lancet 1999;353:1851-1852.

Rautio M, Jousimies-Somer H, Kauma H, Pietarinen I, Saxelin M, Tynkkynen S, et al. Liver abscess due to a Lactobacillus rhamnosus strain indistinguishable from L. rhamnosus strain GG. Clin Infect Dis. 1999;28:1159-1160.

Raza S, Graham SM, Allen SJ, Sultana S, Cuevas L, Hart CA. Lactobacillus GG promotes recovery from acute nonbloody diarrhea in Pakistan. Pediatr Infect Dis J. 1995;14:107-111.

Roberfroid MB. Prebiotics and synbiotics: concepts and nutritional properties. Br J Nutr. 1998;80:S197-S202.

Saavedra JM. Probiotics plus antibiotics: regulating our bacterial environment. J Pediatr. 1999;135:535-537.

Salminen S, von Wright A, Morelli L, Marteau P, Brassart D, de Vos WM, et al. Demonstration of safety of probiotics - a review. Int J Food Microbiol. 1998;44:93-106.

Salminen MK, Rautelin H, Tynkkynen S, Poussa T, Saxelin M, Valtonen V, et al. Lactobacillus bacteremia, clinical significance, and patient outcome, with special focus on probiotic L. rhamnosus GG. Clin Infect Dis. 2004;38:62-69.

Sandler RS, Drossman DA. Bowel habits in young adults not seeking health care. Dig Dis Sci. 1987;32:841-845.

Saxelin M, Chuang NH, Chassy B, Rautelin H, Makela PH, Salminen S, et al. Lactobacilli and bacteremia in southern Finland, 1989-1992. Clin Infect Dis. 1996;22:564-566.

Sazawal S, Hiremath G, Dhingra U, Malik P, Deb S, Black RE. Efficacy of probiotics in prevention of acute diarrhoea: a meta-analysis of masked, randomised, placebo-controlled trials. Lancet Infect Dis. 2006;6:374-382.

Schrezenmeir J, de Vrese M. Probiotics, prebiotics, and synbiotics - approaching a definition. Am J Clin Nutr. 2001;73:S361-S364.

Schrezenmeir J, Heller K, McCue M, Llamas C, Lam W, Burow H et al. Benefits of oral supplementation with and without synbiotics in young children with acute bacterial infections. Clin Pediatr. (Phila) 2004;43:239-249.

Schulz KF, Chalmers I, Hayes RJ, Altman DG. Emperical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408-412.

Seki H, Shiohara M, Matsumura T, Miyagawa N, Tanaka M, Komiyama A, et al. Prevention of antibiotic-associated diarrhea in children by Clostridium butyricum MIYAIRI. Pediatr Int. 2003;45:86-90.

Siitonen S, Vapaatalo H, Salminen S, Gordin A, Saxelin M, Wikberg R et al. Effect of Lactobacillus GG yoghurt in prevention of antibiotic associated diarrhoea. Ann Med. 1990;22:57-59.

Srinivasan R, Meyer R, Padmanabhan R, Britto J. Clinical safety of Lactobacillus casei shirota as a probiotic in critically ill children. J Pediatr Gastroenterol Nutr. 2006;42:171-173.

Sussman JI, Baron EJ, Goldberg SM, Kaplan MH, Pizzarello RA. Clinical manifestations and therapy of Lactobacillus endocarditis: report of a case and review of the literature. Rev Infect Dis. 1986;8:771-776.

Sykora J, Valeckova K, Amlerova J, Siala K, Dedek P, Watkins S, et al. Effects of a specially designed fermented milk product containing probiotic Lactobacillus casei DN-114 001 and the eradication of H. pylori in children: a prospective randomized double-blind study. J Clin Gastroenterol. 2005;39(8):692-698.

Takata GS, Chan LS, Shekelle P, Morton SC, Mason W, Marcy SM. Evidence assessment of management of acute otitis media: I. The role of antibiotics in the treatment of uncomplicated acute otitis media. Pediatrics 2001;108:239-247.

Tankanow RM, Ross MB, Ertel IJ, Dickinson DG, McCormick LS, Garfinkel JF. A double-blind, placebo-controlled study of the efficacy of Lactinex in the prophylaxis of amoxicillin-induced diarrhea. DICP 1990;24:382-384.

Thomas MR, Litin SC, Osmon DR, Corr AP, Weaver AL, Lohse CM. Lack of effect of Lactobacillus GG on antibiotic-associated diarrhea: a randomized, placebo-controlled trial. Mayo Clin Proc. 2001;76:883-889.

Turck D, Bernet JP, Marx J, Kempf H, Giard P, Walbaum O, et al. Incidence and risk factors for of oral antibiotic-associated diarrhea in an outpatient pediatric population. J Pediatr Gastroenterol Nutr. 2003;37:22-26.

Vanderhoof JA, Whitney DB, Antonson DL, Hanner TL, Lupo JV, Young RJ. Lactobacillus GG in the prevention of antibiotic-associated diarrhea in children. J Pediatr. 1999;135:564-568.

Weizman Z, Asli G, Alsheikh A. Effect of a probiotic infant formula on infections in child care centers: comparison of two probiotic agents. Pediatrics 2005;115:5-9.

Wistrom J, Norrby SR, Myhre EB, Eriksson S, Granstrom G, Lagergren L, et al. Frequency of antibiotic-associated diarrhoea in 2462 antibiotic-treated hospitalized patients: a prospective study. J Antimicrob Chemother. 2001;47:43-50.

Witsell DL, Garrett CG, Yarbrough WG, Dorrestein SP, Drake AF, Weissler MC. Effect of Lactobacillus acidophilus on antibiotic-associated gastrointestinal morbidity: a prospective randomized trial. J Otolaryngol 1995;24:230-233.

Zoppi G, Cinquetti M, Benini A, Bonamini E, Minelli EB. Modulation of the intestinal ecosystem by probiotics and lactulose in children during treatment with ceftriaxone. Curr Ther Res. 2001;62:418-435.

CHAPTER 3: Definitions and Primary Outcome Measures in Randomized Trials of Pediatric Acute Diarrhea: A Systematic Review

3.1 ABSTRACT

Background
In a 2006 WHO report, diarrheal diseases ranked second among conditions afflicting children. Despite the disease burden, and the many existing randomized controlled trials (RCTs) of interventions to prevent and treat pediatric diarrheal diseases, there is an absence of guidelines, designed using consensus techniques, related to the definition or measurement of PAD in RCTs.

Objectives
In pediatric RCTs involving acute diarrhea as their primary outcome measure, objectives include: 1) document how acute diarrhea and its resolution is defined; 2) document how acute diarrhea is assessed; and 3) document reporting of the psychometric properties of the outcome measure(s).

Methods
We searched four major databases (CENTRAL, EMBASE, Global Health, MEDLINE) from inception to February 2007 for English-language RCTs in children <19 years of age measuring acute diarrhea as a primary outcome measure. We evaluated each of the included studies for methodological quality (e.g., concealment, blinding, loss to follow-up) and recorded details of outcome measurement.

Results
We identified 121 RCTs reporting one or more primary outcome measures related to PAD. The overall methodological quality of included trials was good. Authors used 62 different definitions of acute diarrhea and 64 different definitions of what constituted resolution of diarrhea. Trials used 62 different primary outcome measures related to diarrhea, the most common of which was duration of diarrhea (n = 65 trials). Thirty-one trials used grading systems (e.g., single and multi-faceted ordinal and continuous scales) to support outcome measure assessment. Of these, three trials stated that their grading system was valid, however, none of the trials (or their citations) reported evidence of this.

Conclusion
Even in what would be considered methodologically sound clinical trials, definitions of diarrhea, outcome measure selection, and grading systems employed in PAD RCTs are heterogeneous and lack evidence of validity. The external validity of such trials is limited and the opportunity to conduct meaningful knowledge synthesis is greatly impeded. Standardized definitions of PAD and its resolution, as well as validated, reliable primary outcome measures would facilitate valid inferences regarding the efficacy of different interventions for PAD.

3.2 INTRODUCTION

Pediatric acute diarrhea presents as a change in normal bowel habit including a substantial increase in the frequency of bowel movements and/or a decrease in stool consistency. The degree of change can be related to the child's age and nutritional status and the underlying cause of diarrhea. The most common cause of acute diarrhea involves gastrointestinal infection (Guandalini 2004). Though diarrhea acts as a defense mechanism in the body, quickly eliminating infective organisms, the most serious sequela is dehydration, particularly in malnourished or immuno-suppressed children (WHO 2006).

The WHO Global Burden of Disease initiative has adopted a health status measurement known as "disability-adjusted life-year", a unifying metric that combines mortality and morbidity. In 2002, diarrheal diseases ranked second among conditions afflicting children (Mathers 2006). The cost per diarrheal episode in the United States has been estimated at $289 USD in <36-month old ambulatory population and the costs of hospitalization of 250,000 patients was $560 million or $2240 USD per case (Avendano 1993; Garthright 1988). With respect to rotavirus diarrhea, there are an estimated 25 million clinic visits, 2 million hospitalizations, and over 600,000 deaths worldwide in children less than 5 years of age (Parashar 2003).

We previously conducted a study of probiotics for the prevention of pediatric antibiotic-associated diarrhea (chapter 2), uncovering the need for the development of standard definitions and validated primary outcome measure (Johnston 2007). Although the pathophysiology, etiologies and clinical sequelae of acute diarrhea have been well described in recent years (Guandalini 2004), there appears to be limited consensus on how to best measure PAD in clinical trials. The definitions of acute diarrhea vary in surveillance studies and clinical trials, making the incidence of diarrhea difficult to estimate (Johnston 2007; Wright 2007). Indeed, few studies have been done to address the choice of outcome measures for clinical trials, and in most pediatric subspecialties, no research has been conducted (Sinha 2008). In recognizing the importance of standard definitions (trial entry criteria, diarrhea resolution criteria) and valid and reliable outcome

measures when planning a PAD study, our aim was to systematically assess how diarrhea is defined and measured in RCTs of PAD.

3.3 METHODS

3.3.1 Search Strategy

In pediatric RCTs involving acute diarrhea as their primary outcome measure, our objectives were to: 1) document how acute diarrhea and its resolution is defined; 2) document how acute diarrhea is assessed; and 3) document reporting of clinimetric properties of the primary outcome measures. In consultation with an expert health services librarian, a systematic review was performed of four major databases published in English from inception to February 2007. We searched Ovid's EMBASE (1980-2007), MEDLINE (1966-2007), CENTRAL (2007 Issue 1) and Global Health (1973-2007). Search terms included extensive controlled vocabulary and keyword searches for (randomized controlled trials) AND (anti-diarrheal treatment or prophylaxis e.g., oral rehydration, vaccine, zinc) AND (pediatric) AND (diarrhea). There were variations in search terms depending on the specific indexing of each database. To identify additional articles that utilized or supported the development of measurement instruments, we searched the bibliographies of included RCTs having employed an instrument related to the measurement of PAD. The search strategy for each database can be found in Appendix 3-1.

3.3.2 Study Selection

We included studies that were: i) randomized controlled trials (treatment or prophylaxis) (ii) that were published in English; iii) covering a pediatric population (0-18 years); iv) of any intervention (e.g., oral rehydration, probiotic, vaccine); and v) with an explicit statement that primary outcome measure was acute diarrhea or an *a priori* sample size calculation based on acute diarrhea. For example, we accepted statements such as "the sample size was estimated to give a power of 90% (alpha = 5%) to detect a 15% reduction in the incidence of diarrhea" as indicative of the study's primary outcome measure (Barreto 1994). Searches were screened using titles of papers and, when available, abstracts. The full texts of the selected articles were retrieved and two

reviewers (BCJ, BRC) independently assessed each article for inclusion according to pre-specified inclusion criteria. Inter-rater reliability was assessed for inclusion criteria by using kappa statistics and disagreement was resolved by consensus.

3.3.3 Data Extraction

Using a standardized data extraction form, two reviewers (BCJ, LS) independently extracted data items (Buscemi 2006) including: study setting, type of trial (prevention or treatment), number and age of children enrolled, intervention, control, the definition of diarrhea and diarrhea resolution, the primary outcome measure(s), the validity and reliability of the outcome measures used, and an assessment of the methodological quality for each trial included. The definitions of diarrhea and its resolution were classified into categories based on three key dimensions: frequency, consistency, and duration (Lebak 2003). Measurement properties of studies that employed a grading system to support outcome measure assessment were also extracted, regardless of whether it was the primary endpoint. Bibliographies of RCTs employing grading systems were searched for previous reports of the instrument in an attempt to uncover evidence of validity and trace the lineage of their development. Any discrepancies were noted and resolved by joint review of the items in question.

3.3.4 Quality Assessment

Each of the included studies was evaluated using the validated 5-point Jadad Scale to assess sequence generation, double-blinding, withdrawals and dropouts (Jadad 1996). Concealment of allocation was assessed as adequate, inadequate or unclear using trial design methodology described by Schulz (Schulz 1995). The full data extraction form can be found in Appendix 3-1.

3.3.5 Data Analysis

For the purposes of this systematic review, combining of data was not appropriate or necessary. Descriptive statistics are employed to illustrate the characteristics of trials measuring acute diarrhea as their primary outcome measure(s).

3.4 RESULTS

Our electronic search yielded 2,738 references (Figure 3-1). The title and abstract screening identified 657 potentially eligible citations. The chance-adjusted, between-reviewer agreement on the application of study inclusion criteria was very good (kappa = 0.89; 95% CI 0.84, 0.93), resulting in the inclusion of 121 RCTs. The 121 RCTs enrolled 69,376 children. Of the included trials, 80 were conducted in infants and toddlers up to 3 years of age, 22 in children up to 5 years of age and 19 trials involved children up to 18 years. Thirty-six (30%) were prophylaxis trials and 85 (70%) were treatment trials. Thirty-eight (31%) trials were community-based studies, 79 (65%) were conducted in a healthcare setting, one was conducted in both settings and in three studies the study setting was unclear. Interventions were diverse, involving 30 therapies (alone or in combination, see Table 3-1). Using the Jadad 0 to 5 scale, based on the reports of the included RCTs, the median methodological quality was 3.0 (range = 1 to 5). Concealment of allocation was adequate in 51 (42%) of trials, inadequate in one, and unclear in the remaining 69 (57%) trials.

3.4.1 Definitions of Diarrhea and its Resolution

Authors reported a definition of diarrhea (i.e., onset) in 119 (98%) trials. Sixty-two (51%) trials reported a unique and independently distinguishable definition of diarrhea (i.e., inclusion criteria for treatment trials, diagnostic criteria for prophylaxis trials). The definitions of diarrhea were classified into 9 categories based on frequency, consistency, and duration. Thirty-seven trials used a variation of the WHO definition of diarrhea, although most did not refer explicitly to the WHO definition (WHO 2008) (see Table 3-2). A number of studies operationalized diarrhea according to the mother's perception (Long 2006; Mitra 1997: Sarker 2005; Savarino 2002; Sur 2003).

Ninety trials (74%) provided a definition of diarrhea resolution (i.e., offset) with 64 (53%) providing a unique description of resolution. Of the 64 unique definitions, 24 provided ≥3 criteria (frequency, consistency, duration); 20 provided 2 criteria and 20 provided 1 criterion. Ten of the 20 studies providing just one criterion employed a previously validated definition of diarrhea resolution (i.e., three intervening diarrhea-free

days) (Barretto 1994; Baqui 2003; Bhandari 1997; Bhandari 2002; Faruque 1999; Sazawal 1995; Sazawal 1997; Sazawal 2007; Sempertegui 1999; Sur 2003). Barretto (1994) was the lone study citing the previous validity report of the definition of diarrhea resolution (Barretto 1994; Baqui 1991).

3.4.2 Primary Outcome Measures

Seventy-six trials (63%) explicitly stated one or more primary outcome measures, 97 (80%) provided a statement regarding a sample size calculation and 53 (44%) trials clearly stated both their primary outcome measure(s) and a statement regarding sample size. Trials reported 62 different primary outcome measures; none reported the use of a valid and reliable primary measure. The most common primary outcome measure was duration of diarrhea in 65 trials (see Table 3-3).

3.4.3 Reliability of Grading Systems to Support Outcome Measure Evaluation

Thirty-one trials (26%) employed grading systems (Allen 1994; Arvola 1999; Barreto 1994; Becker 2006; Bernstein 1999; Block 2007; Faruque 1999; CSG 2001; Fawzi 2000; Fayad 1999; Hoekstra 2004; Jacobs 1994; Kaplan 1999; Khan 2005; Linhares 1996; Long 2006; Nakamura 2006; Narkeviciute 2002; Raghupathy 2006; Rosenfeldt 2002; Ruiz-Palacios 2006; Saha 2005; Salazar-Lindo 2004; Salinas 2005; Sarker 2005; Savarino 2002; Sharieff 2006; Thibault 2004; Vesikari 2006; Weizman 2005; Zaman 2001) of which 18 trials (15%) used uni-faceted scales involving stool consistency categories (some with accompanying pictures), which were applied to determine primary outcome measures such as the incidence of diarrhea or duration of diarrhea (Allen 1994; Arvola 1999; Becker 2006; Faruque 1999; Fawzi 2000; Hoekstra 2004; Jacobs 1994; Khan 2005; Nakamura 2006; Narkeviciute 2002; Raghupathy 2006; Rosenfeldt 2002; Saha 2005; Sarker 2005; Sharieff 2006; Thibault 2004; Weizman 2005; Zaman 2001). One study stated the reliability properties of a stool consistency grading scheme used to support the measurement of their chosen primary outcome measure – duration of diarrhea, and provided a citation for a reliability study supporting their grading system choice (Arvola 1999).

3.4.4 Validity of Grading Systems

Of the 31 trials employing grading systems, 7 used one of 3 multi-faceted scales: the 15-point modified Diarrheal Index Score, the 20-point Ruuska & Vesikari Scale, and the 24-point scale proposed by Clark (1988). The authors did not report the measurement properties of these instruments (Bernstein 1999; Block 2007; Jacobs 1994; Linhares 1996; Ruiz-Palacios 2006; Salinas 2005; Vesikari 2006). However, Block et al. 2007 stated that the 24-point clinical scoring system they employed was validated, citing Clark et al. 1988 (Block 2007; Clark 1988). In retrieving the full-text of this article and any related articles cited by Clark (1988), no papers contained a report of the measurement properties of the 24-point scale (Clark 1988; Duffy 1986; Riepenhoff-Talty 1981).

Two additional trials stated that their grading system was valid (Long 2006; Faraque 1999). The grading systems included a chart with picture symbols to record the frequency of diarrhea stools (Faraque 1999) and a questionnaire (Long 2006). The former cited an unpublished article in support of their claims of validation (Faraque 1999); whereas the latter provided no citation to support claims of validity for the questionnaire employed (Long 2006).

With respect to the additionally identified multi-faceted grading systems (i.e., Ruuska & Vesikari 20-point scale, 15-point Diarrheal Index Score), we also searched for previous reports of these scales and collected the full-text papers in an attempt to trace the lineage of their development (Bernstein 1995; Ericsson 1983; Ericsson 1987; Flores 1987; Hjelt 1987; Hjelt 1987; Joensuu 1997; Ruuska 1990; Ruuska 1991). We were unable to locate any reports describing the clinimetric properties of these instruments. Additional uni/multi-faceted scales and questionnaires employed as outcome measures in the included trials are described in Table 3-4.

The uniformity of outcome measure selection in trials examining vaccines for the prevention of acute infectious diarrhea (i.e., rotavirus gastroenteritis) appear to be superior to that of other included trials. Specifically, six out of eight vaccine trials (Bernstein 1999; Block 2007; Linhares 2005; Ruiz-Palacios 2006; Salinas 2005; Vesikari

2006) included for review employed either the 20 or 24 point grading systems cited above. The ability of these instruments to distinguish the severity of diarrhea and to be responsive to change in multiple RCTs provides some reassurance of their validity.

3.5 DISCUSSION

Worldwide, acute diarrheal diseases rank second among conditions afflicting children, yet there is surprisingly little agreement on how to define and measure this illness. We identified 121 RCTs that reported 62 unique definitions of diarrhea, 64 unique definitions of diarrhea resolution and 62 unique primary outcome measures. Thirty-one trials used grading systems to support outcome measure evaluation and three of these stated that their grading system was valid (Block 2007; Faraque 1999; Long 2006). However, none of the trials (or their citations) reported evidence of their validation. Given our results, the external validity of these trials is limited and the opportunity to conduct meaningful knowledge synthesis is impeded.

Aside from a single stool consistency grading system, previously shown to be reliable, (Allen 1994), the majority of scales employed are unpublished or provided no citation regarding the development or measurement properties of the scale. The use of grading systems that are unpublished, when employed in clinical trials, may impart misleading results. An investigation involving 300 RCTs of published versus unpublished instruments demonstrated that trials involving therapies for schizophrenia were more likely to report claims of treatment superiority based on instruments that were not published (Marshall 2000). Similar results have been found in RCTs of attention-deficit hyperactivity-disorder (Jadad 1999). Our findings lend creadance to the need for increased scrutiny by all stakeholders (e.g., patients, clinicans, policy makers, funding agencies) of the potential for misleading results in trials based on unpublished scales (CONSORT 2009).

There is limited research on definitions of diarrhea and outcome measures employed in studies of acute diarrhea. Further, few reviews in the field of pediatrics have used the comprehensive review methods we have employed (Sinha 2008). Although we are

unaware of other reviews of acute diarrheal outcome measures employed in clinical trials of children, we are aware of at least three reviews involving endpoints employed in other pediatric conditions. Anand (2005) searched three databases for review articles or original data related to infant pain, ethical issues and study design (including outcome measures) (Anand 2005); Miller (2001) searched PubMed for all prospective therapeutic trials of adult and juvenile idiopathic inflammatory myopathies for disease outcome measures and the publication of validated studies to support these measures (Miller 2001); while Zhang (2001) searched five leading general medicine and pediatric journals for primary outcome measures used to evaluate therapies in RCTs of newborn infants (Zhang 2005). Our study represents the most comprehensive review of definitions and primary outcome measures employed in pediatric RCTs and, in particular, RCTs of PAD.

Our results may be limited as a product of not searching the grey literature (contact with authors or review of all citations of included studies). However, our aim was not to identify every RCT of PAD published to obtain a cumulative point-estimate around the efficacy of interventions through meta-analysis. Rather, we aimed to acquire a comprehensive sample of PAD trials for evaluation. Another potential limitation is the exclusion of non-English studies, which may have employed validated outcome measures. Given the large amount of literature in this area – to make the study feasible – we needed to build in some limitations that were reasonable. We did not search non-English studies since we were interested in identifying a validated outcome measure in English. There is reason to suppose that even if something existed in another language, one could not assume it was valid in its English format without additional validation (Sperber 2004).

The heterogeneity among RCTs of PAD identified here is an example of issues that are important across all trials and meta-analyses. We were surprised to see that these issues have not been resolved in a public health area as common and important as PAD. Our research may be useful to help promote the identification and use of standard disease definitions and the use of outcomes that have been demonstrated valid and reliable among RCTs in other fields of study. The lack of uniformity in definitions and outcome

measure assessment limits the insights that clinicians and health policy makers can glean from RCT results. We recommend that the international community of diarrhea and/or gastroenteritis investigators collaborate to resolve the limitations inherent in the studies reviewed here. Given the diversity of diarrheal definitions and primary outcome measures reported, there is a need to come to consensus in at least two key areas: 1) standard definitions of diarrhea and diarrhea resolution; and 2) standard core outcome measures. If investigators wish to employ criteria, in addition to those that have been agreed upon, this would be acceptable. However, the pooling of core criteria could be used for the purposes of meta-analysis (Clark 2007). An effort of this nature would result in uniformity among trials and be of significant benefit to trial end-users (clinicians, patients, pharmaceutical industry, and health policy makers).

Collaborative efforts in other research fields have led the charge on this important, but often neglected issue. Perhaps the best example is the Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) initiative - a group of rheumatology investigators who have paved the way toward standardizing the conduct, assessment and reporting of clinical trials in their field (Goldsmith 1993; Tugwell 1993). This initiative has also involved work toward standardizing the conduct of studies of juvenile idiopathic inflammatory myopathies (Miller 2001; Rider 1997). Standardizing outcome measures across clinical trials of any particular condition makes it easier to interpret, compare, and synthesize results of RCTs so that inferences regarding the efficacy of different interventions for PAD are not misleading. Standard criteria (definitions, outcomes) are not meant to impede the development or use of other criteria (e.g., health related quality of life measures), but would represent criteria routinely used and reported. Investigators wishing to employ other criteria in a particular trial should be encouraged to do so, but when reporting their trial, selective reporting could be avoided (i.e., outcome reporting bias) through reporting of core criteria that the international diarrhea research community have endorsed (Chan 2008; Clarke 2007). Such an effort would require consensus, guidelines, and adherence on behalf of the relevant stakeholders.

The basic tenents of RCT design emphasize the importance of randomization, allocation concealment, blinding and loss to follow-up (Jadad 1996; Schulz 1995). However, the importance of employing valid and reliable outcome measures has been relatively overlooked by key organizations such as funders and journals that endorse the CONSORT and PRISMA guidelines for reporting of trials and systematic reviews of trials. In order to ask more sophisticated research questions; head-to-head trials employing factorial or non-inferiority trial designs are needed. Trials need to also employ valid (measure what is intended) and reliable (use of an instrument that yields the same results on repeated trials) outcome measures. For instance, although ORS and vaccines have had an enormous impact worldwide in reducing the number of deaths related to diarrhea, it is unclear which ORS may be must beneficial. Given that there are therapies known to substantially reduce morbidity and mortalitiy, there is now a need for head-to-head comparative research (i.e., glucose based ORS versus rice-based ORS) to elucidate the most effective therapies to treat PAD. In addition, attention to the development of valid and reliable outcome measures for acute diarrhea trials in pediatrics is essential in establishing the clinical impact that emerging interventions such as probiotics or zinc may have (Allen 2003; Lazzerini 2008). Although placebo-controlled trials have demonstrated that probiotics as an adjunct to ORS may reduce the duration of diarrhea (Allen 2003; Johnston 2008), it is not clear which probiotic strain may be most effective. Head-to-head trials of ORS with different probiotic stains, or combination of strains, that employ valid and reliable outcome measures are needed to accurately discern if differences really exist between therapies. This is especially important given that differences between probiotic strains may be subtle as a result of different mechanisms of action of different probiotic strains (Ng 2009). If we consider the importance of a reliable primary outcome measure in the measurement of diarrhea, the consistency of stool (formed vs. loose vs. liquid) is a pivotal discriminator for determining the diagnosis of diarrhea (number of diarrhea stools in the preceding 24 hrs) and its resolution (duration of diarrhea). Without a reliable instrument to support what is and what is not considered diarrhea (such as the Stool Consistency Classification System) (Bliss 2001; 2003), inter-rater variation alone has the potential to make the difference between an intervention that

demonstrates therapeutic significance and no response, and between registered FDA approval and no approval (Cooney 2007).

As a result of the heterogeneous definitions and outcome measures and the lack of validity and reliability evidence for the outcome measures identified here, the lumping of trials of PAD using meta-analysis techniques may be misleading. Individual RCTs that are designed to control for selection, detection and loss to follow-up bias, as well as to overcome the limitation of employing outcome measures that have not been demonstrated to be reliable or valid, are paramount and will advance our understanding the therapeutic potential of PAD therapies much more than relying on meta-analysis. As pointed out in a critique of meta-analyses, "*if the same systematic biases are present across a range of studies, the only effect of meta-analysis is to reinforce them, to produce spurious statistical stability, and thereby to discourage further research"* (Shapiro 1994). Once the physiologic measures such as duration of diarrhea are determined using valid and reliable measures, subsequent outcome measures that measure components of health status important to clinicians and patients, that are not limited to physiologic aspects of health, are needed. In other words, direct measurement of how children or parents are feeling as a result of PAD and the extent to which they are able to function in daily activities, that is, health related quality of life (Guyatt 2007).

REFERENCES

Allen DA, Deadman L, Wang EEL. Standardizing the assessment of diarrhea in clinical trials: Results of an interobserver agreement study. Acta Paediatr.1994;83:179-182.

Allen SJ, Okoko B, Martinez EG, Gregorio GV, Dans LF. Probiotics for treating infectious diarrhoea. Cochrane Database of Syst Rev. 2003, Issue 4. CD003048.

Anand KJS, Aranda JV, Berde CB, Buckman S, Capparelli EV et al. Analgesia and anesthesia for neonates: study design and ethical issues. Clin Ther. 2005; 27:814-843.

Arvola T, Laiho K, Torkkeli S, Mykkanen H, Salminen S, Maunula L et al. Prophylactic Lactobacillus GG reduces antibiotic-associated diarrhea in children with respiratory infections: a randomized study. Pediatrics 1999;104:e64.

Avendano P, Matson DO, Long J, Whitney S et al. Costs associated with office visits for diarrhea in infants and toddlers. Pediatr Infect Dis J. 1993;12:897-902.

Baqui AH, Black RE, El Arifeen S, Yunus M, Chakraborty J, Ahmed S, Vaughan JP. Effect of zinc supplementation started during diarrhoea on morbidity and mortality in Bangladeshi children: community randomised trial. BMJ. 2002;325:1059.

Baqui AH, Zaman K, Persson LA, El Arifeen S, Yunus M, Begum N, Black RE. Simultaneous weekly supplementation of iron and zinc is associated with lower morbidity due to diarrhea and acute lower respiratory infection in Bangladeshi infants. J Nutr. 2003;133:4150-4157.

Baqui AH, Black RE, Yunus M, Hoque AR, et al. Methodological issues in diarrhoeal diseases epidemiology: definition of diarrhoeal episodes. Int J Epidemiol. 1991;20:1057-1063.

Barreto ML, Santos LM, Assis AM, Araújo MP, Farenzena GG, Santos PA, Fiaccone RL. Effect of vitamin A supplementation on diarrhea and acute lower-respiratory-tract infections in young children in Brazil. Lancet 1994;344:228-231.

Becker B, Kuhn U, Hardewig-Budny B. Double-blind, randomized evaluation of clinical efficacy and tolerability of an apple pectin-chamomile extract in children with unspecific diarrhea. Arzneimittelforschung 2006; 56:387-93.

Bernstein DI, Glass RI, Rodgers G, Davidson BL, Sack DA. Evaluation of rhesus rotavirus monovalent and tetravalent reassortant vaccines in US children. US Rotavirus Vaccine Efficacy Group. JAMA. 1995;273:1191-6.

Bernstein DI, Sack DA, Rothstein E, Reisinger K, Smith VE, O'Sullivan D, Spriggs DR, Ward RL. Efficacy of live, attenuated, human rotavirus vaccine 89-12 in infants: a

randomized placebo-controlled trial. Lancet 1999; 354:287-90.

Bhandari N, Bahl R, Sazawal S, Bhan MK. Breast-feeding status alters the effect of vitamin A treatment during acute diarrhea in children. J Nutr. 1997;127:59-63.

Bhandari N, Bahl R, Taneja S, Strand T, Mølbak K, Ulvik RJ, Sommerfelt H, Bhan MK. Substantial reduction in severe diarrheal morbidity by daily zinc supplementation in young north Indian children. Pediatrics. 2002;109:e86.

Bliss DZ, Larson SJ, Burr JK Savik K. Reliability of a stool consistency classification system. J Wound Ost Contin Nur. 2001:28:305-313.

Bliss DZ, Dhamani KA, Savik K, Kirk K. Tool to classify stool consistency: content validity and use by persons of diverse cultures. Nurs Health Sci. 2003;5:115-121.

Block SL, Vesikari T, Goveia MG, River SB, Adeyi BA, Dallas MJ et al. Efficacy, immunogenicity, and safety of a pentavalent human-bovine (WC3) reassortant rotavirus vaccine at the end of shelf life. Pediatrics 2007; 119:11-18.

Buscemi N, Hartling L, Vaddermeer B, Tjosvold L, Klassen TP. Single data extraction generated more errors than double data extraction in systematic reviews. J Clin Epidemiol. 2006;59:697-703.

Chan AW. Bias, spin, and misreporting: time for full access to trial protocols and results. PLoS Med. 2008;5:230.

CHOICE Study Group. Multicenter, randomized, double-blind clinical trial to evaluate the efficacy and safety of a reduced osmolarity oral rehydration salts solution in children with acute watery diarrhea. Pediatrics 2001; 107:613-618.

Clarke M. Standardising outcomes for clinical trials and systematic reviews. Trials 2007;8:39.

Clark HF, Borian FE, Bell LM, Modesto K, Gouvea V, Plotkin SA. Protective effect of WC3 vaccine against rotavirus diarrhea in infants during a predominantly serotype 1 rotavirus season. J Infect Dis. 1988;158:570-587.

Clark HF, Burke CJ, Volkin DB, Offit P, Ward RL, Bresee JS. et al. Safety, immunogenicity and efficacy in healthy infants of G1 and G2 human reassortant rotavirus vaccine in a new stabilizer/buffer liquid formulation. Pediatr Infect Dis J. 2003;22:914-920.

Cooney R, Warren B, Altman D, Abreu M, Travis S. Outcome measurement in clinical trials for ulcerative colitis: towards standardisation. Trials 2007;8(1):17.

Duffy LC, Byers TE, Riepenhoff-Talty M, La Scolea LJ, Zielezny M, Ogra PL. The effects of infant feeding on rotavirus-induced gastroenteritis: a prospective study. Am J Public Health 1986;76:259-263.

Ericsson CD, DuPont HL, Sullivan P, Galindo E, Evans DG, Evans DJ Jr. Bicozamycin, a poorly absorbable antibiotic, effectively treats travelers' diarrhea. Ann Intern Med. 1983;98:20-25.

Ericsson CD, Johnson PC, Dupont HL, Morgan DR, Bitsura JA, de la Cabada FJ. Ciprofloxacin or trimethoprim-sulfamethoxazole as initial therapy for travelers' diarrhea. A placebo-controlled, randomized trial. Ann Intern Med. 1987;106:216-220.

Faruque AS, Mahalanabis D, Haque SS, Fuchs GJ, Habte D. Double-blind, randomized, controlled trial of zinc or vitamin A supplementation in young children with acute diarrhea. Acta Paediatr. 1999; 88:154-160.

Fawzi WW, Mbise R, Spiegelman D, Fataki M, Hertzmark E, Ndossi G. Vitamin A supplements and diarrheal and respiratory tract infections among children in Dares Salaam, Tanzania. J Pediatr. 2000; 137:660-667.

Flores J, Perez-Schael I, Gonzalez M, Garcia D, Perez M, Daoud N, Cunto W, Chanock RM, Kapikian AZ. Protection against severe rotavirus diarrhea by rhesus rotavirus vaccine in Venezuelan infants. Lancet 1987;1:882-884.

Garthright WE, Archer DL, Kvenberg JE. Estimates of incidence and costs of intestinal infectious diseases in the United States. Public Health Rep. 1988;103:107-115.

Gothefors L, Wadell G, Juto P, Taniguchi K, Kapikian AZ, Glass R. Prolonged efficacy of rhesus rotavirus vaccine in Swedish children. J Infect Dis. 1989;159:753-757.

Guandalini S. Acute Diarrhea, In: Walker AW, Goulet O, Kleinman RE, Sherman PM, Shneider BL, Sanderson IR, editors. Pediatric Gastrointestinal Disease: Pathophysiology, Diagnosis, Management. Toronto: BC Decker Inc., 2004, p. 166-79.

Guyatt GH, Ferrans CE, Halyard MY, Revicki DA, Symonds TL, Varricchio CG, et al. Clinical Significance Consensus Meeting Group. Exploration of the value of health-related quality-of-life information from clinical research and into clinical practice. Mayo Clin Proc. 2007; 82:1229-1239

Hjelt K, Grauballe PC, Paerregaard A, Nielsen OH, Krasilnikoff PA. Protective effect of preexisting rotavirus-specific immunoglobin A against naturally acquired rotavirus infection in children. J Med Virol.1987;21:39-47.

Hjelt K, Nielson OH, Paerregaard A, Grauballe PC, Krasilnikoff PA. Acute gastroenteritis in children attending day-care: Diarrhea with special reference to rotavirus infections. II. Clinical manifestations. Acta Paediatr Scand. 1987;76:763-768.

Jacobs J, Jiménez LM, Gloyd SS, Gale JL, Crothers D. Treatment of acute childhood diarrhea with homeopathic medicine: a randomized clinical trial in Nicaragua. Pediatrics 1994;93:719-725.

Jacobs J, Jiménez LM, Malthouse S, Chapman E, Crothers D, Masuk M, Jonas WB. Homeopathic treatment of acute childhood diarrhea: results from a clinical trial in Nepal. J Altern Complement Med. 2000;6:131-139.

Jacobs J, Guthrie BL, Montes GA, Jacobs LE, Mickey-Colman N, Wilson AR, DiGiacomo R. Homeopathic combination remedy in the treatment of acute childhood diarrhea in Honduras. J Altern Complement Med. 2006;12:723-732.

Jadad AR, Moore RA, Carrol D, Jenkinson C, Reynolds DJ, Gavaghan DJ et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996;17:1-12.

Jadad AR, Boyle M, Cunningham C, Kim M, Schachar R. Treatment of attention-deficit/hyperactivity disorder.Evid Rep Technol Assess (Summ). 1999;11:1-341.

Joensuu J, Koskenniemi E, Pang XL, Vesikari T. Randomised placebo-controlled trial of rhesus-human reassortant rotavirus vaccine for prevention of severe rotavirus gastroenteritis. Lancet 1997;350:1205-1209.

Johnston BC, Shamseer L, da Costa BR, Vohra S. Outcome measures in clinical trials of pediatric acute diarrhea: a systematic review. Cochrane Colloquia 2008; Oct 3-7, Freiburg, Germany (http://www.imbi.uni-freiburg.de/OJS/cca/index.php/cca/article/view/6893).

Johnston BC, Supina AL, Ospina M, Vohra S. Probiotics for the prevention of pediatric antibiotic-associated diarrhea. Cochrane Database Syst Rev: Cochrane Library 2007, Issue 2. CD004827.

Kaplan MA, Prior MJ, McKonly KI, DuPont HL, Temple AR, Nelson EB. A multicenter randomized controlled trial of a liquid loperamide product versus placebo in the treatment of acute diarrhea in children. Clin Pediatr (Phila) 1999;38:579-591.

Khan WA, Saha D, Rahman A, Salam MA, Bogaerts J, Bennish ML. Comparison of single-dose azithromycin and 12-dose, 3-day erythromycin for childhood cholera: a double-blind trial. Lancet 2002;360:1722-1727.

Lazzerini M, Ronfani L. Oral zinc for treating diarrhoea in children. Cochrane Database of Syst Rev: Cochrane Library 2008, Issue 3. CD005436.

Lebak KJ, Bliss DZ, Savik K, Patten-Marsh KM.What's new on defining diarrhea in tube-feeding studies? Clin Nurs Res. 2003;12:174-204.

Linhares AC, Gabbay YB, Mascarenhas JDP, de Freitas RB, Oliveira CS, Bellesi N.

et al. Immunogenicity, safety and efficacy of tetravalent rhesus-human, reassortant rotavirus vaccine in Belem, Brazil. Bull World Health Org. 1996;74:491-500.

Long KZ, Santos JI, Rosado JL, Lopez-Saucedo C, Thompson-Bonilla R, Abonce M, et al. Impact of vitamin A on selected gastrointestinal pathogen infections and associated diarrheal episodes among children in Mexico City, Mexico. J Infect Dis. 2006;194:1217-1225.

Marshall M, Lockwood A, Bradley C, Adams C, Joy C, Fenton M. Unpublished rating scales: a major source of bias in controlled trials of treatments for schizophrenia. Br J Psychiatry 2000;176:249-252.

Mathers DC, Loncar D. Projections of Global Mortality and Burden of Disease from 2002 to 2030. PloS Med. 2006;3:e442.

Miller FW, Rider LG, Chung YL, Cooper R, Danko K et al. Proposed preliminary core set measures for disease outcome assessment in adult and juvenile idiopathic inflammatory myopathies. Rheumatology 2001;40:1262-1273.

Mitra AK, Akramuzzaman SM, Fuchs GJ, Rahman MM, Mahalanabis D. Long-term oral supplementation with iron is not harmful for young children in a poor community of Bangladesh. J Nutr. 1997;127:1451-1455.

Nakamura S, Sarker SA, Wahed MA, Wagatsuma Y, Oku T, Moji K. Prebiotic effect of daily fructooligosaccharide intake on weight gain and reduction of acute diarrhea among children in a Bangladesh urban slum: a randomized double-masked placebo-controlled study. Tropical Med Health 2006; 34:125-131.

Narkeviciute I, Rudzeviciene O, Leviniene G, Mociskiene K, Eidukevicius R. Management of Lithuanian children's acute diarrhea with Gastrolit solution and dioctahedral smectite. Eur J Gastroenterol Hepatol. 2002;14:419-424.

Ng SC, Hart AL, Kamm MA, Stagg AJ, Knight SC. Mechanisms of action of probiotics: recent advances. Inflamm Bowel Dis. 2009;15:300-310.

Parashar UD, Hummelman EG, Bresee JS, Miller MA, Glass RI. Global illnesses and death caused by rotavirus disease in children. Emerg Infect Dis. 2003;9:565-572.

Raghupathy P, Ramakrishna BS, Oommen SP et al. Amylase-resistant starch as adjunct to oral rehydration therapy in children with diarrhea. Journal of Pediatric Gastroenterology & Nutrition 2006; 42:362-368.

Rider LG, Giannini EH, Harris-Love M, Joe G, Isenberg D, Pilkington C, et al. Defining Clinical Improvement in Adult and Juvenile Myositis. J Rheumatol. 2003 30:603-617.

Riepenhoff-Talty M, Bogger-Goren S, Li P, Carmody PJ, Barrett HJ, Ogra PL. Development of serum and intestinal antibody response to rotavirus after naturally acquired rotavirus infection in man. J Med Virol. 1981;8:215-222.

Rosenfeldt V, Michaelsen KF, Jakobsen M et al. Effect of probiotic Lactobacillus strains in young children hospitalized with acute diarrhea. Pediatr Infect Dis J. 2002; 21:411-416.

Ruiz-Palacios GM, Perez-Schael I, Velazquez FR et al. Safety and efficacy of an attenuated vaccine against severe rotavirus gastroenteritis. N Engl J Med. 2006; 354:11-22.

Ruuska T, Vesikari T. Rotavirus disease in Finnish children: use of numerical scores for clinical severity of diarrhoeal episodes. Scand J Infect Dis. 1990;22:259-267.

Ruuska T, Vesikari T, Delem A, André FE, Beards GM, Flewett TH. Evaluation of RIT 4237 bovine rotavirus vaccine in newborn infants: correlation of vaccine efficacy to season of birth in relation to rotavirus epidemic period. Scand J Infect Dis. 1990;22:269-278.

Ruuska T, Vesikari T. A prospective study of acute diarrhea in Finnish children from birth to 2 1/2 years of age. Acta Paediatr Scand. 1991;80:500-507.

Saha D, Khan WA, Karim MM, Chowdhury HR, Salam MA, Bennish ML. Single-dose ciprofloxacin versus 12-dose erythromycin for childhood cholera: a controlled trial. Lancet 2005; 966:1085-1093.

Sarker SA, Sultana S, Fuchs GJ, Alam NH, Azim T, Brüssow H, Hammarström L. Lactobacillus paracasei strain ST11 has no effect on rotavirus but ameliorates the outcome of nonrotavirus diarrhea in children from Bangladesh. Pediatrics2005;116:e221-228.

Salinas B, Pérez Schael I, Linhares AC, Ruiz Palacios GM, Guerrero ML, Yarzábal JP, et al. Evaluation of safety, immunogenicity and efficacy of an attenuated rotavirus vaccine, RIX4414: A randomized, placebo-controlled trial in Latin American infants. Pediatr Infect Dis J. 2005;24:807-816.

Savarino SJ, Hall ER, Bassily S, Wierzba TF, Youssef FG, Peruski LF Jr, et al. Introductory evaluation of an oral, killed whole cell enterotoxigenic Escherichia coli plus cholera toxin B subunit vaccine in Egyptian infants. Pediatr Infect Dis J. 2002;21:322-330.

Sazawal S, Black RE, Bhan MK, Bhandari N, Sinha A, Jalla S. Zinc supplementation in young children with acute diarrhea in India. N Engl J Med. 1995;333:839-844.

Sazawal S, Black RE, Bhan MK, Jalla S, Sinha A, Bhandari N. Efficacy of zinc supplementation in reducing the incidence and prevalence of acute diarrhea--a community-based, double-blind, controlled trial. Am J Clin Nutr. 1997;66:413-418.

Sazawal S, Dhingra U, Dhingra P, Hiremath G, Kumar J, Sarkar A, Menon VP, Black RE. Effects of fortified milk on morbidity in young children in north India: community based, randomised, double masked placebo controlled trial. BMJ 2007; 20;334:140.

Sempértegui F, Estrella B, Camaniero V, Betancourt V, Izurieta R, Ortiz W, Fiallo E, Troya S, Rodríguez A, Griffiths JK. The beneficial effects of weekly low-dose vitamin A supplementation on acute lower respiratory infections and diarrhea in Ecuadorian children. Pediatrics 1999;104:e1.

Shapiro S. Meta-analysis/Shmeta-analysis. Am J Epidemiol. 1994;140:771-778.

Sperber AD. Translation and validation of study instruments for cross-cultural research. Gastroenterology 2004;126:S124–128.

Sharieff W, Bhutta Z, Schauer C, Tomlinson G, Zlotkin S. Micronutrients (including zinc) reduce diarrhea in children: The Pakistan Sprinkles Diarrhoea Study. Arch Dis Child. 2006; 91:573-579.

Sinha I, Jones L, Smyth RL, Williamson PR. A systematic review of studies that aim to determine which outcomes to measure in clinical trials of children. PLoS Med. 2008: 5:e96.

Sur D, Gupta DN, Mondal SK, Ghosh S, Manna B, Rajendran K, Bhattacharya SK. Impact of zinc supplementation on diarrheal morbidity and growth pattern of low birth weight infants in kolkata, India: a randomized, double-blind, placebo-controlled, community-based study. Pediatrics. 2003;112(6 Pt 1):1327-1332.

Thibault H, Aubert-Jacquin C, Goulet O. Effects of long-term consumption of a fermented infant formula (with Bifidobacterium breve c50 and Streptococcus thermophilus 065) on acute diarrhea in healthy infants. J Pediatr Gastroenter Nutr. 2004;39:147-152.

Vesikari T, Clark HF, Offit PA, Dallas MJ, DiStefano DJ, Goveia MG, et al. Effects of the potency and composition of the multivalent human-bovine (WC3) reassortant rotavirus vaccine on efficacy, safety and immunogenicity in healthy infants. Vaccine 2006;24:4821-4829.

Wright J, Gundry SW, Conroy RM. A review of changing episode definitions and their effects on estimates of diarrhoeal morbidity. J Health Popul Nutr. 2007;25:448-455.

World Health Organization. Water related diseases: Diarrhea. 2006; Available

at: http://www.who.int/water_sanitation_health/diseases/diarrhea/en/ Accessed January 20, 2009.

World Health Organization. Diarrhoea (definition): Available at: http://www.who.int/topics/diarrhoea/en/. Accessed January 20, 2009.

Weizman Z, Asli G, Alsheikh A. Effect of a probiotic infant formula on infections in child care centers: Comparison of two probiotic agents. Pediatrics 2005; 115:5-9.

Zaman K, Yunus M, Rahman A, Chowdhury HR, Sack DA. Efficacy of a packaged rice oral rehydration solution among children with cholera and cholera-like illness. Acta Paediatr. 2001; 90:505-510.

Figure 3-1: Flow Diagram of Search Results

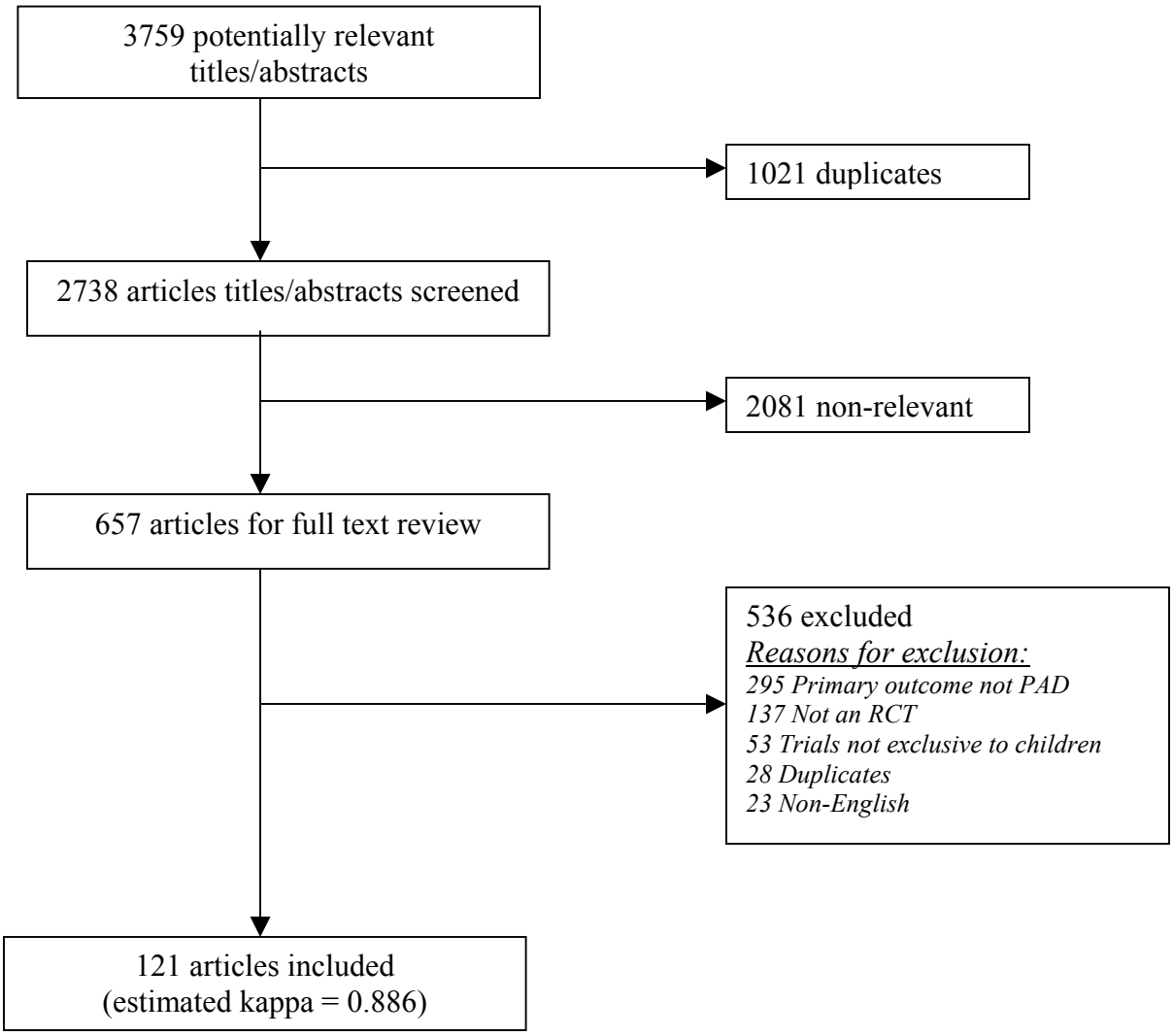


Table 3-1: *Interventions*

Most common interventions administered	***# of trials***
Oral rehydration	33
Probiotic	15
Zinc	14
Vitamin A	9
Vaccine	8
Formula	5
Antibiotic	4
Zinc & Vitamin A	4
Colostrum	3
Homeopathy	3
Loperamide	2
Prebiotic	2
Smectite	2

Table 3-2: *Definitions*

Definitions of diarrhea	***# of trials***
≥ 2 loose or liquid stools/ 24 hrs	2
≥ 3 loose or watery or liquid or semi-liquid stools/ 24 hrs (WHO criteria)‡	37
≥ 3 loose stools / 24 hrs or 1 blood or mucoid stools/ 24 hrs ¶	15
≥ 3 loose/watery stools/ 48 hrs	2
≥ 4 loose/unformed/watery/liquid stools/ 24 hrs	12
≥ 5 loose/watery/liquid/ 24 hrs	8
Mother's judgment & criteria (e.g., ≥ 3 loose/watery stools/ 24 hrs)	5
Gastroenteritis (4 or more criteria, one of which was ≥3 loose/watery stools/ 24 hrs)	3
Trials having defined diarrhea using 1 or 2 criteria (& not fitting into the above categories) †	33

‡ To meet WHO criteria studies must have stated liquid or watery to be included and not stated anything about blood or mucus

¶ Did not contain the term liquid or watery and/or contained the term blood or mucous. Two of the 15 trials employed a previously validated definition: ≥ *3 loose stools or ≥1 bloody stools/ 24 hrs* (Baqui 2002; Baqui 2003)

† By comparison, the other studies always used frequency and consistency over a pre-specified time period.

Table 3-3: *Outcome Measures*

Primary outcome measures	***# of trials***
Duration of diarrhea	65
Stool output/volume	46
Incidence of diarrhea	15
Stool frequency	7
Prevalence of diarrhea	7
Incidence rate (i.e., child-periods/density)	6
Intake (i.e., formula, ORS)	5
Duration of infection (i.e., rotavirus)	3
Prevalence of infection (i.e., rotavirus)	3
Body weight	3

Note: 54 trials reported more than one primary outcome measure

Table 3-4: *Grading Systems Employed as Outcome Measures*

	Co	NoS	DF	DD	BlM	SE	Gas	VF/ Se	VD	Fe	FeD	De	Tx	AbP	Be	BeD	ADL	WB
Study																		
Barretto‘94				√														
Bernstein‘99			√	√				√	√	√		√	√					
Block‘07†			√	√				√	√	√	√				√	√		
CHOICE‘01												√						
Fayad‘99												√						
Jacobs‘94	√		√					√		√		√		√				
Kaplan‘99								√						√				
Linhares‘96			√	√					√	√		√	√					
Long‘06†	√	√			√					√								
Salazar-Lindo’04												√						
Salinas‘05			√	√				√	√	√		√	√					
Savarino‘02																	√	√
Sharieff‘06								√	√	√	√							
Ruiz-Palocioi‘06			√	√				√	√	√		√	√					
Vesikari‘06			√	√				√	√	√	√				√	√		
Weizman‘05	√	√				√	√								√		√	
Total	**3**	**2**	**7**	**7**	**1**	**1**	**1**	**8**	**7**	**9**	**3**	**8**	**4**	**2**	**3**	**2**	**2**	**1**

Co = Stool Consistency; **NoS** = Number of Stools; **DF** = Diarrhea Frequency; **DD** = Duration of Diarrhea; **BlM** = Blood or Mucus; **SE** = Stooling Effort; **Gas** = Gas; **VF** = Vomiting Frequency or Vomiting Severity; **VD** = Vomiting Duration Scale; **Fe** = Fever; **FeD** = Fever Duration; **De** = Dehydration; **Tx** = Treatment (i.e., rehydration, hospitalization); **AbP** = Abdominal Pain/cramping; **Be** = Behavioral Symptoms (i.e., crying, irritable, lethargic, restless, seizure); **BeD** = Behavioral Symptom Duration; **ADL** = Activities of Daily Living (i.e., appetite, sleep); † = Authors’ state the grading system employed was valid

CHAPTER 4: Development and Content Validity of the Pediatric Acute Diarrheal Diseases Scale: A Modified Delphi Consensus Approach

4.1 ABSTRACT

Background
Diarrheal diseases rank second among conditions afflicting children, resulting in substantial global mortality and morbidity. Despite the disease burden, there is limited consensus on the definition of pediatric acute diarrhea, its remission, or criteria for evaluating its severity and this is an important methodological limitation of research in this area.
Objectives
In children (0-5 years of age): 1) How should acute diarrhea and its resolution be defined? 2) What items should be included in a scale to evaluate the severity of pediatric acute diarrhea? 3) Given these items, what response scale should be used to obtain reliable data that can be validly interpreted?
Methods
A modified Delphi consensus procedure was employed to develop a new instrument for evaluating the severity of PAD. The study involved steering committee discussions (phase 1) and two electronic surveys (phase 2 and 3) of leading experts in measurement and clinical gastroenterology. Definitions of diarrhea were established and items and their respective response formats were developed to construct the International Pediatric Acute Diarrheal Diseases Scale (IPADDS).
Results
Our instrument is based on feedback and consensus from 19 (61.3% response rate) expert clinician-scientists who have conducted 134 RCTs of PADD. These experts agreed on the inclusion of five attributes containing 13 items. Attributes proposed for the IPADDS multi-faceted index include: *Diarrhea Frequency* and *Duration, Vomiting Frequency* and *Duration, Fever*, *Restrictions in Normal Daily Activities* and *Dehydration*.
Conclusion
This study represents the first in a series of sequential steps toward the development of a valid and reliable multi-faceted outcome measure for measuring the severity of PADD in RCTs. A numerical scoring system needs to be added and further empirical evidence of reliability and validity is required.

4.2 INTRODUCTION

Pediatric acute diarrhea is a leading cause of child hospitalization and occurs most frequently during the first two years of life, with infants less than 6 months of age and most at risk for severe diarrhea (Avendano 1993; Molbak 1997; Fagundes-Neto 1999). Worldwide, there are an estimated 1.4 billion episodes of diarrhea per year in children less than 5 years of age (Kosek 2003; Parashar 2003). Although mortality associated with PAD has decreased substantially in recent years, it still accounts for approximately 17% of all deaths in this age group (Bryce 2005). Despite a reduction in mortality, there has not been a decline in morbidity attributable to diarrhea (Bern 1992; Wright 2007).

Based on published trials, there is no consensus on the definition of PAD, its resolution, or criteria for evaluating its severity. In a recent systematic review, we identified 121 RCTs that reported 62 unique definitions of diarrhea, 64 unique definitions of diarrhea resolution, and 62 unique primary outcomes (Johnston 2008). Thirty-one trials used grading systems to support outcome measure evaluation and three of these trials stated that their grading systems were valid (Long 2006; Faruque 1999; Block 2007). However, none of the trials (or their citations) reported evidence of their validation. As such, the external validity of these trials is limited and the opportunity to conduct meaningful knowledge synthesis is greatly impeded. Consequently, three research questions were addressed in the present study: (i) How should PAD and its resolution be defined? (ii) What items should be included in a scale to evaluate the severity of PAD? and (iii) Given these items, what response scale should be used to obtain reliable data that can be validly interpreted?

4.3 OVERVIEW OF METHODS

Using a modified Delphi procedure involving unstructured discussions and two structured surveys, the definition of PAD and an accompanying instrument for measuring PAD severity were developed (Hasson 2000; Sackman 1974). The items and their respective response formats were developed to form a measurement instrument for PAD. The study was completed in three sequential phases: (i) steering committee discussions to develop the initial definitions of PAD and the accompanying items; (ii) collection of content validity

evidence by an external panel of local, national and international experts and revision of the initial definitions and items in the instrument; and (iii) assessment of the utility of the final instrument by the same panel of experts.

4.4 PHASE I: Development of Definition and Scale

4.4.1 Methods

The aim of the study was to develop an instrument to measure the severity of PAD (i.e., a disease activity index) for employment in RCTs. In conjunction with the results of our systematic review of PAD (Johnston 2008), steering committee discussions were the basis for item generation, revision and item scaling. To help with the development of the definition and identification of what items should be included, a steering committee with expertise in PAD and/or psychometrics was established. The members of this committee together with their areas of expertise and their relevant degrees are in Table 4-1 below.

Table 4-1: *Steering Committee*

	Expertise	**Qualifications**
Expert 1	Pediatric Emergency Medicine	MD, MSc
Expert 2	Measurement of Diarrhea, Dietetics	RD, PhD
Expert 3	Measurement of Dehydration, Pediatric Emergency Medicine	MD
Expert 4	Measurement of Pain, Pediatric Clinical Psychology	PhD
Expert 5	Psychometrics	PhD
Expert 6	Gastroenterology, Randomized Trials	MD
Expert 7	Randomized Trials, Pharmacology	MSc, PharmD
Expert 8	Randomized Trials, General Pediatrics	MD, MSc
Expert 9	Randomized Trials, Pediatric Acute Diarrhea	ND, PhD (cand.)

Before developing an instrument, it is usual to have a constitutive definition that describes the domain of interest prior to developing the items to include in the instrument. However, in this case, although previous instruments to measure the severity of PAD were found lacking, when taken together, the items provided a starting point to develop the needed definition and to construct the initial set of items that reflected the definition. We generated

a list of items for consideration in the definition of PAD, developed a draft definition, and then adopted or adapted existing items or developed new items to form the initial draft of the disease activity index.

Definition

Items that could be used to help develop the definition of PAD were selected from (i) trials included in a systematic review of RCTs in PAD (Johnston 2008); (ii) previous scales used to evaluate diarrhea (Bliss 2001; Clark 1988; Ericsson 1983; Flores 1987; Hjelt 1986; Jacobs 1994; Lewis 1997; Riepenhoff-Talty 1981; Ruuska & Vesikari 1990; Whelan 2008); and (iii) papers reporting the common clinical presentation of PAD (Carnerio 2005; Giaquito 2007; Guandalini 2000; Guandalini 1988; Vernacchio 2006).

We originally set out to study probiotics for the prevention of pediatric AAD (see Chapter 2), but uncovered the need for the development of a validated primary outcome measure. We systematically reviewed the literature on primary outcomes employed in RCTs of PAD (Chapter 3) and found definitions and outcome measures that sometimes included vomiting (often referred to as gastroenteritis) and other times did not. The list of items in the left column (Table 4-2 below) was discussed with the members of the steering committee. The steering committee members together with the principal investigator (BCJ) discussed whether or not each item was relevant to the measurement of the severity of PAD. The steering committee agreed on all items but one. In particular, the members were divided on whether or not vomiting should be included in the definition of the severity of PAD. The issue was centered on the difference between PAD and pediatric acute gastroenteritis (PAG). The decision taken was to develop two definitions, one for diarrhea and one for gastroenteritis. The definition of PAD was based on the first 13 items listed in the left column of Table 4-2. The definition for PAG was based on the 16 items listed in the right column of Table 4-2.

Table 4-2: *Preliminary List of Items for PAD and PAG*

PAD	**PAG**
Stool frequency	Stool frequency
Stool consistency	Stool consistency
Stool volume	Stool volume
Diarrhea duration	Diarrhea duration
Blood in stool	Blood in stool
Dehydration	Dehydration
Abdominal pain	Abdominal pain
Pain & discomfort	Pain & discomfort
Restrictions in normal daily activities	Restrictions in normal daily activities
Nocturnal awakening	Nocturnal awakening
Appetite	Appetite
Fever	Fever
Nausea	Nausea
	Retching frequency
	Vomiting frequency
	Vomiting duration

4.4.2 Results and Justification

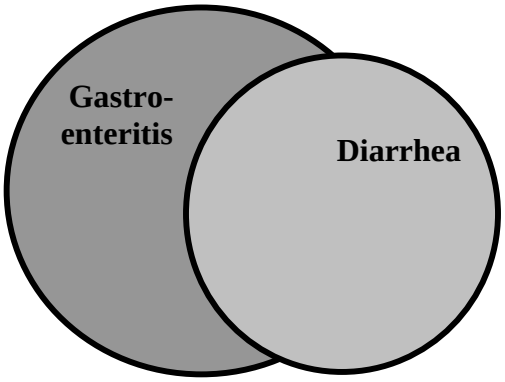
From the table, it is evident that pediatric gastroenteritis encompasses all of the same attributes as PAD, with the addition of attributes related to vomiting. At the same time and given that the duration of diarrhea is the most common primary outcome employed in RCTs of diarrheal diseases (Johnston 2008), the decision was taken to include definitions of the resolution of PAD and PAG. The four definitions are provided in Table 4-3.

Table 4-3: *Definitions of PAD and PAG and their Resolution*

Pediatric Acute Diarrhea
Pediatric acute diarrhea is defined by an increase in the frequency of bowel movements and a change in the consistency of stool (i.e., diminished degree of firmness). Diarrhea may be associated with dehydration, pain or discomfort, restrictions in normal daily activities of the child or caregiver, and fever.
Resolution of Pediatric Acute Diarrhea
The duration of acute diarrhea typically resolves in less than 14 days and resolution is marked by (a) production of 2 consecutive normal stools or; (b) production of one normal stool followed by 12 hours with no stool production; or (c) no stool production for a period of 12 hours.
Pediatric Acute Gastroenteritis
Pediatric acute gastroenteritis is defined by an increase in the frequency of bowel movements, and a change in the consistency of stool (i.e., diminished degree of firmness) and vomiting. Gastroenteritis may be associated dehydration, pain or discomfort, restrictions in normal daily activities of the child or caregiver and fever.
Resolution of Pediatric Acute Gastroenteritis
The duration of acute gastroenteritis typically resolves in less than 14 days and resolution is marked by (a) production of 2 consecutive normal stools and absence of vomiting or; (b) production of one normal stool followed by 12 hours with no stool production or vomiting; or (c) no stool production or vomiting for a period of 12 hours.

Differentiation between PAD and PAG is difficult clinically. According to the WHO, the definition of diarrhea is: *three or more loose or liquid stools per day, or more frequently than is normal for the individual* (http://www.who.int/topics/diarrhoea/en/). Antibiotic-associated diarrhea is also defined using similar frequency and consistency parameters (Johnston 2007). On the other hand, according to European Society for Paediatric Gastroenterology, Hepatology, and Nutrition and the European Society for Paediatric Infectious Disease, the term "gastroenteritis" implies: *a decrease in the consistency of stools (loose or liquid) and/or an increase in the frequency of evacuations (typically ≥3 in 24 hours), with or without fever or vomiting* (Guarino 2008). Although the definitions are similar in that they contain an operational definition of diarrhea, gastroenteritis implies that a patient has the symptoms of diarrhea with or without fever or vomiting. In other words, gastroenteritis has attributes in addition to diarrhea. These additional attributes are unlikely to occur in those diagnosed with AAD. Given the difference, diarrhea and gastroenteritis can be said to be different but related, constructs as shown in Figure 4-1. For these reasons we chose to develop two separate definitions and respective scales.

Figure 4-1: Overlapping Constructs of Acute Gastroenteritis and Diarrhea



Item Development

Once the definitions of PAD and PAG had been developed, the items were reexamined and revised as needed to fit with the definitions and to form the initial draft of the International Pediatric Acute Diarrhea Scale (IPADS) and the International Acute Gastroenteritis Scale (IPAGS). The potential items for inclusion in these scales were discussed and refined using personal and electronic communication with the steering committee members.

After multiple discussions via phone, email, and in person with each member of the steering committee, we reached consensus on six items for inclusion in the IPADS and eight items for inclusion in IPAGS. As shown in Table 4-7, three of the subscales contained more than one item. The subscales for the IPADS, with the corresponding number of items in each subscale provided in brackets, included 1) diarrhea frequency (one); 2) diarrhea duration (two); 3) dehydration (four); 4) fever (one); 5) pain and discomfort (five), and; 6) restrictions in normal daily activities (one). The two additional items for the IPAGS were vomiting frequency (one) and vomiting duration (one).

Each item was accompanied by a set of response options. A scoring scale with a score point attached to each response option was also included given the intent to develop cut-

scores regarding the severity acute diarrhea or acute gastroenteritis. The possible scores for each item or subscale ranged from zero to 10, with higher values indicating a more severe condition. For example, the following response options for assessing a child's activities of daily living were awarded the score in brackets: Activities not disturbed (0), Child less playful/social (2), Child or caregiver sleep or daily activities disturbed (4), Child or caregiver unable to attend to homemaking duties, daycare/school, or work (6), Visit to healthcare practitioner, and (8) Admitted to hospital (10). The total number of points across the six subscales for the IPADS and the eight subscales of the IPAGS were, respectively, 60 and 80.

With respect to both the IPADS and the IPAGS items that measured *Diarrhea Frequency* and *Diarrhea Duration*, it was necessary to clarify the nature of stool. The *Stool Consistency Classification System* (Bliss 2001) was adopted for this purpose. This is the only instrument, to our knowledge, with demonstrated adequate content validity (Bliss 2003) and high reliability (Bliss 2001). Using 12 adult stool specimens, subjects in each of three groups (20 nurses, 20 nursing students, and 20 lay persons) classified the consistency of specimens. The percent agreement for inter-rater reliability was high, ranging from 67% to 100%. Using a weighted kappa statistic, agreement beyond chance among the 60 observers on all stool specimens using word and diagram descriptions was 0.74 (p <0.0001). For test-retest reliability, percent agreement was no less than 72% with a weighted kappa of 0.75 (p <0.001) for the consistency classifications of all stool specimens on day 1 and day 2 (Bliss 2001). The series of four diagrams together with their verbal descriptions shown in Table 4-4 preceded the diarrhea frequency and diarrhea duration items. It was felt that by providing the diagrams together with a verbal description of each diagram, a more accurate and consistent measurement of diarrhea consistency and duration would result. Additional stool consistency scales considered for adoption included the Bristol Stool Chart, the King's Stool Chart and the Bergstrom's Stool Consistency tool (Lewis 1997; Whelan 2004; Young 1999). Only the Bergstrom's was developed based on children. However, the methods for its development were limited in that only two experts were used to judge the content validity; inter-rater agreement, 0.81, was calculated using percent agreement rather than Cohen's kappa,

which also adjusts for chance agreement, and the eight stool consistency categories were felt to be both burdensome for raters and too numerous if the objective was to employ a reliable tool for interpreting stool consistency.

Table 4-4: *Description of Stool Consistency Classification System (Ver.1)*

1) Stool Consistency Classification System (Adopted from Bliss *et al*. J. Wound Ostomy. Contin. Nurs. 2001)			
Hard and Formed	**Soft but Formed**	**Loose & Unformed**	**Liquid**
			
Having a hard or firm texture and retaining a definite shape like a banana, cigar or marbles	Retaining same general shape in the collection bag; does not spread all over the bottom of bag, or has a texture that appears like peanut butter	Lacking any shape of its own; spreads over the bottom of the collection bag**;** having a texture that appears like hot cereal	Like water

We also incorporated two items (*Pain* and *Dehydration*) having demonstrated evidence of validity as stand alone scales. We chose to include a measure of pain because previous observational studies of the clinical presentation of PAD have demonstrated that abdominal pain is the fifth most common symptom reported; 35% of over 300 children with infectious diarrhea (Guandalini 1988; 2000) and 19% of 611 diarrheal episodes in children with mostly non-infectious diarrhea (no likely pathogen was found in almost 80% of cases) (Vernacchio 2006). The *Pain* item included in the IPADS and IPAGS was adopted from Merkel et al. (1997). The details for including this item are summarized in Table 4-5. Additional *Pain* scales considered for adoption included the Faces Pain Scale – Revised (Hicks 2001), Faces Pain Scale (1990), Preverbal, Early Verbal Pediatric Pain Scale (Schultz 1999), University of Wisconsin Children's Hospital Pain Scale (Soetenga 1999) and the Wong-Baker Faces Scale (Wong & Baker 1998).

Table 4-5: *Proposed Pain Item*

Face Legs Activity Cry Consolability (FLACC) **Scale** (Merkel et al. 1997)			
Age/sample	**Type of pain**	**Categories & Scale range**	**Measurement Properties**
2 months – 7 years, n = 89 (developmentally delayed children excluded) Children 4 – 19 years of age with cognitive impairment, n = 52	Variety of surgical procedures	Five categories of behavior (face, legs, activity, cry, consolability) each scored on a 0-2 scale Total score 0 –10	Inter-rater reliability r = 0.94 Criterion validity: FLACC vs. global ratings of pain r = 0.41, p<0.005; Objective Pain Score vs. FLACC r = 0.80, p<0.001 Construct validity was supported by significant reductions in pain scores after analgesic administration Cognitively Impaired: Inter-rater reliability (ICC, ranging from 0.76 to 0.90). Criterion validity (rho = 0.65-0.87; p <0.001). Construct validity via analgesic administration (6.1 +/- 2.6 vs. 1.9 +/- 2.7; p <0.001) (Malviya et al. 2006; Malviya et al. 2005)

The *Dehydration* item was adapted from a previously validated scale of clinical dehydration (Freidman 2004; Goldman 2008). The adaptation involved revising the scoring scale so that it ranged from 0 to 10, rather than 0 to 8. Investigators developed and validated a clinical dehydration scale with four items (general appearance, eyes, mucous membranes, and tears) in children 1 month to 5 years with symptoms of acute gastroenteritis at a single tertiary pediatric center in Canada. The first study, a prospective observational study, enrolled 137 children with infectious diarrhea who were evaluated for 12 clinical characteristics of dehydration. The final dehydration scale consisted of the four clinical characteristics stated above and the measurement properties were as follows: reliability as assessed by intra-class correlation was 0.77; validity as assessed by Pearson's correlation coefficient was 0.36 to 0.57 and responsiveness to change as assessed by Wilcoxon signed rank test was significant ($p<0.01$) (Freidman 2004). The second study enrolled 205 children and demonstrated that three dehydration categories (none, some, moderate to severe) were positively associated with length of hospital stay (245 ± 181; 397 ± 302; 501 ± 389, $p<0.001$) and the proportion of children receiving intravenous rehydration (15%; 49%; 80%, $p<0.001$). The results of this study support the use of the *Clinical Dehydration Scale* and the three dehydration severity

categories for diagnostic and therapeutic purposes in children 1 month to 5 years. Additional dehydration scales considered were the Fortin-Parent Scale (Fortin 1978), the Santosham Scale (Santosham 1987), and the WHO Scale (WHO 2005); however, all have limited evidence of reliability and validity.

The remaining IPADS items included *Fever* and *Normal Daily Activities.* Although these items were found in previous scales, the items in the IPADS contain unique qualifiers and response options. *Fever* is a symptom that regularly accompanies PAD (Guandalini 2000; Guandalini 1988; Vernacchio 2006) and is a symptom of considerable parental importance (Betz 2006; Crocetti 2001). It is most common in children <5 years of age, with the prevalence being highest in children 6 to 24 months of age (Kline 1999). Invasive pathogens that infect the distal ileum and colon such as *Campylobacter jejuni, Salmonella species, Yersinia enterocolitica, Escherichia coli* and *Shigella species*, can all be clinically present with fever. *Rotavirus,* which invades the proximal small intestine and is the single most pervasive cause of infectious diarrhea worldwide, is also highly correlated with fever (Guandalini 2000).

There are four modes of taking temperature (rectal, ear, oral and axillary) as shown in Table 4-6. Previous scales have employed rectal and oral temperatures (Jacobs 1994; Ruiz 2006; Vesikari 2006). However, many parents are uncomfortable with consenting to having their child subjected to a rectal temperature. For anatomical reasons, tympanic temperature is not reliable in children up to 2 years of age. In addition, it is unusual to find tympanic thermometers outside of emergency rooms in developed countries. Oral temperature is considerably difficult to obtain in young children as they need to keep the thermometer in their mouth with the tongue depressed for 3-4 minutes (standard thermometer) and 15-20 seconds with a digital thermometer. Oral temperature may be challenging to acquire in unwilling children, especially those less than three years of age, and there is limited access to digital thermometers in the developing world. For these reasons, axillary temperature, although generally regarded as less reliable than rectal, was adopted to monitor fever in children up 5 years (Avner 2009; Leduc 2009). However, axillary body temperature can be taken while the child is asleep and a recent study of 90

inpatient and ambulatory Italian children (<1 year, 1-5 years, >5 years) demonstrated that a relatively inexpensive infrared thermometer can measure body temperature accurately and reliably at various body sites, including the axilla (Osio 2007). Normal axillary temperature is defined according to the Canadian Pediatric Society (see Table 4-6), and is generally considered to be one degree Celsius below rectal.

Table 4-6: *Normal Temperature Ranges*

Measurement method	**Normal temperature range**
Rectal	36.6°C to 38°C (97.9°F to 100.4°F)
Ear	35.8°C to 38°C (96.4°F to 100.4°F)
Oral	35.5°C to 37.5°C (95.9°F to 99.5°F)
Axillary	34.7°C to 37.3°C (94.5°F to 99.1°F)

Canadian Task Force on Preventive Health - Strength of Recommendation B, II (39)
Paediatrics and Child Health 2000 (re-endorsed in 2008)

Although we aimed to devise response options that increase in severity as a child moves from the response options on the left to the response options on the right, the *Fever* item is the exception. For example, *Fever*, although a potential contributor to dehydration, is relatively less important as a risk factor for dehydration and hospitalization when compared to the fluid loss as a result of diarrhea or vomiting. In fact, serious cases of PADD can result in lower or even hypothermic temperatures if the child is severely dehydrated. In addition, body temperature varies as much as 0.5 Celsius from the mean under normal circumstances, with the lowest temperatures in the early morning (4:00 am to 8:00 am) and reaching its height in the early evening (4:00 pm to 6:00 pm) (Mackowiak 1992). Given that temperature can vary as a result of the time of day and ambient air temperature, it proved difficult to propose gradients of temperature that reflect increasing severity of disease. Given that we were unable to locate a fever scale based on gradients backed by empirical evidence, we decided to choose two temperatures, a normal temperature and a high-grade temperature, and partition the values in the middle as mild and moderate. As indicated above we adopted the recommendations from the Canadian Pediatric Society as well as the American Academy of Pediatrics regarding normal axillary temperature (34.7 to 37.3 Celsius) (Avner 2009; Leduc 2009) and for the high-grade temperature we employed ≥39 Celsius, based on febrile children 2 to 36 months without localizing signs of infection who are at the greatest risk of bacteremia (Al-Rashed 2008; Osman 2002; Peters 2004). The high-grade temperature has also been used in the Ruuska and Vesikari Scale (1990).

It is recognized that no matter how gradients are justified, clinical appearance rather than height of fever is a more accurate predicator of serious illness (Avner 2002; Trautner 2006). Future research might consider the possibility of weighting fever lower than attributes such as diarrhea and vomiting on the overall score scale.

Restrictions in *Normal Daily Activities* was employed as an outcome in one RCT included in our systematic review (Savarino 2002) and may reflect what is important to the child and parent (Guyatt 2004). The daily activity item and subscale was constructed since we were unable to identify a stand-alone scale that had been validated. In addition, the subscales in instruments such as the Irritable Bowel Syndrome-36 (IBS-36) and the Impact II, which were designed to measure disease specific health-related quality of life in children with Inflammatory Bowel Disease (Loonen 2002), and the generic health related Child Health Questionnaire (CHQ) were inappropriate. The IBS-36 is not validated in children and the CHQ subscale and Impact II subscale are limited in scope and like, the IBS-36, are validated as part of a multi-dimensional scale. Although each of these instruments is well established, they have been validated as multi-dimensional instruments and are often scored with this in mind. For these reasons we constructed the *Normal Daily Activities* item with six response options. The initial draft of the IPADS is presented in Table 4-7.

Table 4-7: *International Pediatric Acute Diarrhea Scale (Ver.1)*

1) Stool Consistency Classification System‡ (Adopted from Bliss *et al*. J. Wound Ostomy. Contin. Nurs. 2001)			
Hard and Formed	**Soft but Formed**	**Loose & Unformed**	**Liquid**
			
Having a hard or firm texture and retaining a definite shape like a banana, cigar or marbles	Retaining same general shape in the collection bag; does not spread all over the bottom of bag, or has a texture that appears like peanut butter	Lacking any shape of its own; spreads over the bottom of the collection bag**;** having a texture that appears like hot cereal	Like water

2) Diarrhea Frequency				
	Diarrhea Stools in Preceding 24 hours			
Stool Consistency	0 stools	1 – 2 stools	3 – 4 stools	≥5 stools
Loose & Unformed stools(s)	*0 score*	*1 score*	*5 score*	*9 score*
Liquid Stool(s)	*0 score*	*2 score*	*6 score*	*10 score*

3) Diarrhea Duration				
	Number of Days with Diarrhea (in previous 13 days)			
Stool Consistency	0 days	1 – 2 days	3 – 4 days	≥5 days
Loose & Unformed stool(s)	*0 score*	*1 score*	*5 score*	*9 score*
Or Liquid Stool(s)	*0 score*	*2 score*	*6 score*	*10 score*

4) Dehydration (Adopted from Friedman *et al.* J. Pediatrics 2004)			
	Score = 0 each box	***Score = 1.5 per box***	***Score = 2.5 per box***
General Condition	□ Normal	□ Thirsty, restless or lethargic but irritable when touched	□ Drowsy, limp, cold, sweaty, +/ comatose
Eyes	□ Normal	□ Slightly sunken	□ Very sunken
Mucus Membranes (tongue)	□ Moist	□ "Sticky"	□ Dry
Tears	□ Tears	□ Decreased tears	□ Absent tears

5) Fever				
	Temperature			
Axillary	34.7 to 37.3°C	37.4 to 38.1°C	38.2 to 38.9°C	≥39.0°C
	0 score	*2 score*	*6 score*	*10 score*

6) Pain & Discomfort (Adopted from Merkel et al. Pediatr. Nurs. 1997)			
	Score = 0 per box	***Score = 1 per box***	***Score = 2 per box***
Face	□ No particular expression or smile	□ Occasional grimace or frown, withdrawn, disinterested	□ Frequent or constant quivering chin, clenched jaw
Legs	□ Normal position or relaxed	□ Uneasy, restless, tense	□ Kicking, or legs drawn up
Activity	□ Lying quietly, normal position, moves easily	□ Squirming, shifting back and forth, tense	□ Arched, rigid or jerking
Cry	□ No cry (awake or asleep)	□ Moans or whimpers; occasional complaint	□ Crying steadily, screams or sobs, frequent complaints
Consolability	□ Content, relaxed	□ Reassured by occasional touching hugging or being talked to, distractible	□ Difficult to console or comfort

7) Activities of Daily Living						
	Score = 0	***Score = 2***	***Score = 4***	***Score = 6***	***Score = 8***	***Score = 10***
(e.g., child or caregiver activities: playing, daycare or school, work)	Activities not disturbed	Child less playful/social	Child/caregiver sleep or daily activities disturbed	Child/caregiver unable to attend to homemaking duties, daycare/school, or work	Visit to health-care practitioner	Admitted to hospital

‡ Stool Consistency Classification System is not scored per se, but used to score Diarrhea Frequency and Diarrhea Duration

Regarding the IPAGS, the six items included in the IPADS together with two additional items were used to measure PAG (see Table 4-8). The two new items measured the number of vomiting episodes per day (*Vomiting Frequency*) and the duration of vomiting (*Vomiting Duration*). Each of these items, although found in previous scales (Clark 1988; Ruuska & Vesikari 1990; Sharieff 2006), employed newly constructed response options and a different scoring system.

Table 4-8: *International Pediatric Acute Gastroenteritis Scale (Ver.1b)*

7) Vomiting Frequency				
	Episodes of Vomiting in Preceding 24 hours			
Vomiting	0 episodes	1 – 2 episodes	3 - 4 episodes	≥5 episodes
	0 score	*2 score*	*6 score*	*10 score*
8) Vomiting Duration				
	Number of Days with Vomiting (in the previous 13 days)			
Vomiting	0 days	1 – 2 days	3 - 4 days	≥5 days
	0 score	*2 score*	*6 score*	*10 score*

Note: the IPAGS contains all of the IPADS items as well as the additional vomiting items depicted here

4.5 PHASE II: Empirical Evidence of Content Validity

4.5.1 Methods

A second independent panel of experts (the outer panel of experts) was recruited to review and assess the definitions and items included in the initial drafts of the IPADS and IPAGS. A questionnaire was developed for this purpose. The members of the outer panel of experts first reviewed and assessed the definitions. They then assessed each item in terms of its relevance to the definition to which it was referenced and the appropriateness of the accompanying scale. Once they had completed the item-relevance task, they were asked to consider the items they deemed to be relevant and assess how well this set represented the definition to which the items were referenced. In this way, data about both relevance of each item and representativeness of the relevant items were obtained.

Characteristics of Outer Panel Experts

Thirty-one physicians and nurses with expertise in acute diarrhea and/or gastroenteritis were approached by email and asked to participate. To identify the expertise of potential

participants, three 'panels' were identified *a priori* according to their overall experience in clinical practice and research. Panels were identified as A, B or C (see Table 4-9).

Table 4-9: *Expertise*

Expert Criteria
Expertise in at least one of the following clinical areas or one of the following research areas: *Clinical* 1. pediatric emergency medicine 2. pediatric gastroenterology *Research* 3. measurement (validity and/or reliability of outcome measures) 4. randomized control trials
Overall experience in clinical practice and research
Panel A: Experts who have been in clinical practice 15 years or more and completed at least three RCTs of PAD/G (as a co-investigator) or two RCTs of PAD/G as a primary or senior investigator.
Panel B: Experts who have been in clinical practice for 10 –14 years and completed at least two RCTs of pediatric acute diarrhea (as a co-investigator) or one RCT of PAD/G as a primary or senior investigator.
Panel C: Experts who have been in clinical practice for at least 5 years and have been a co-investigator on any clinical trial or been involved in measurement (i.e., reliability, validity) research.

Given that Panel A had relatively more experience in both research and clinical practice, the responses from these panel members were given greater weight than the members of Panel B or Panel C. If there was disagreement between panels on the relevance of items, then the information from Panel A was used.

Recruitment of Outer Panel of Experts

Using electronic mail, 31 experts (10 from Panel A, 10 from Panel B, and 11 from Panel C) were initially contacted and invited to participate in the study. They were requested to indicate their willingness and availability to participate within a forthcoming two-week period. A survey consisting of a content validity rating survey (described in detail below) and a background questionnaire was sent to the willing experts in mid-September 2008. Non-responders were sent electronic mail reminders up to three times over the two-week data collection period ending September 2008.

Content Validity Rating Survey

An online survey questionnaire was administered, consisting of three parts: (i) draft definitions of PAD and PAG and their resolution; (ii) a form for assessing the proposed items for the measurement of PAD and PAG severity and their resolution; and (iii) a background questionnaire to gain information about the expertise of participants. Approximately 20 minutes was required to complete the entire online survey questionnaire.

Item Response Format

A Likert-type 5-point response option was selected to assess the definitions and items (Likert 1932; Likert 1970). Likert (1932) asserted that participants presented with a response scale anchored only at the endpoints would first attach meaning to the interior points by partitioning the distance between the points into equal intervals and then choose the closest option to their true position. Thus, the response options were treated as an equal interval scale (see also Lam & Klockars 1982).

The panels of experts were asked to indicate the degree to which they agreed:

a) with the proposed definitions of PAD and PAG and the resolution of each (1 = Strongly disagree to 5 = Strongly agree);
b) that the set of four diagrams represented stool consistency in children less or equal to five years of age (1 = Not representative to 5 = Very representative);
c) that each of the eight proposed items (six for PAD and eight for PAG) were relevant to the definition to which they were referenced (1 = Not relevant to 5 = Very relevant); and
d) that the proposed scoring scale for each of the items was appropriate (1 = Not appropriate to 5 = Very appropriate).

A space was provided for each item in which panel members could provide comments regarding the content or wording of the item and scoring scale. After the panel members had completed their ratings of the items and scoring scales referenced to PAD, they were asked to re-consider the items they found to be relevant (rating of 4 or 5). They were then

asked if their set of relevant items represented the definition of PAD and its resolution using a five-point scale (1= Not representative to 5 = Very representative). If they rated the representativeness with three or less, they were asked to indicate what was missing and to provide a corresponding item. They were asked to do the same for PAG.
The final section contained a set of questions developed to obtain background information on each of the experts. A copy of the survey questionnaire is provided in Appendix 4-1. The survey was approved by the Health Research Ethics Board at the University of Alberta. Participants were not identified by name and the analysis of the data involved group results only.

Data analysis

To assess inter-expert agreement and the possible presence of outlying panelists for the IPADS and IPAGS, the absolute value of the discrepancy between the rating for each panel member and the median of all panel members' ratings was calculated for each item and summed across all items in the scale (JDM_{js}). The formula for JDM_{js} (Rogers 2008) is as follows:

$$JDM_{js} = \sum_{k=1}^{K} \left| X_{kj} - Md_k \right|,$$

where: X_{kj} is the rating given by panelist j to item k;
Md_k is the median of the ratings given by the J judges/panelists to item k; and
K is the number of items in subscale s.

Ideally there will be perfect agreement among the panel members on all items and each panel members' discrepancy from the median is zero. If the JDM_{js} for a panel member exceeded the range of JDM_{js}'s for the remaining panel members, the panel member's ratings were considered to be outlying. There are no strict criteria for judging deviancy other than it be substantially removed from the ratings of the remaining panelists. Outlying responses from panelists sometimes occur because they do not understand the task they are required to complete and/or the domain that is being assessed. Outlying panel member's scores were then examined further to establish whether the panel member provided comments that explained the outlying scores and/or suggestions for item revision.

Item relevance was examined by assessing: 1) the *degree of ambiguity* among panelists' ratings for each item and 2) the *central tendency* of the panelists' ratings for each item. *Degree of item ambiguity* was assessed by examining the range, *R*, of the panelists' ratings for each of the items (Rogers 2008). For item *k*,

$$R_k = X_{kj_H} - X_{kj_L} + 1,$$

where X_{kj_H} and X_{kj_L} are, respectively, the highest and lowest ratings on each item. The value of R_k should ideally be 1. R_k=1 when the highest and lowest ratings are the same. Substantial ambiguity in the fit of an item is represented by large values of R_k. Due to the judgmental nature of this process, unacceptable R_k values could not be determined *a priori.* A large observed R_k raised questions about the fit of the item to the definition to which it was referenced and motivated further examination for insight into why the item was ambiguous.

The central tendency of the ratings for each item was then examined to assess whether panel members believe there was a fit between the item and the definition to which it was referenced. The median of the panelists' ratings for each item was used. Larger median values indicate better fit or more relevant items. Since items were assessed using a scale ranging from 1 to 5, items required a median value of at least 3.50 in order to meet the criteria of item relevance. Panel A ratings were considered to determine which items remained and which items were deleted.

To assess item representativeness, the panel members were requested to further to rate how representative the diagrams of stool consistency were and how representative the set of "relevant" items (i.e., those items rated as a "4" or "5") were of the overall constructs – PAD and PAG. Adequate representation of each construct was assumed if the majority of Panel A experts rated their set of relevant PAD and PAG items, respectively, representative of the corresponding definitions.

Regarding the incorporation of comments provided by the panelists, the following procedures were followed: 1) review of comments of the outlying panelists for each item; 2) review of comments made by retained panelists for each item with a median below 3.50 and/or high item ambiguity; and 3) content analysis techniques of the recorded comments to identify key themes (Mayan 2009; Morse 1995). The comment(s) were evaluated by the principal investigator (BCJ) and shared with the members of the steering committee. Based on feedback from the steering committee, the principal investigator chose to retain, delete or revise the definitions and items referenced to each definition. Comments were interpreted as an opportunity to: 1) understand why experts may have selected particular item ratings on the respective rating scales; 2) evaluate the wording of items (experts from different cultural or clinical backgrounds may interpret wording differently); and 3) identify possible problems regarding the participants' expertise for evaluating research-specific nature of the test items.

For data captured from the background questionnaire (see Appendix 4-2), namely the level of expertise in PAD, a median, range and sum were calculated for continuous scaled variables and number and percentages for categorical variables. At no time during data handling or analysis did the investigators share expert participants' names or contact information. All data were analyzed using SPSS 14.0 and Microsoft Excel 2007.

4.5.2 Results and Justification

Response rate and expertise of panelists

Twenty-five (80.6%) of the identified 31 experts indicated their willingness to participate. Of those who indicated willingness to participate, 7/9 (77.8%) Panel A experts responded, 2/5 (40.0%) Panel B experts responded, and 10/11 (90.9%) Panel C experts responded. Since only two Panel B experts responded, Panel C was collapsed into Panel B. One Panel B expert and one Panel C expert were reassigned to Panel A based on the *a priori* criteria, leaving 9 experts in Panel A and 10 experts in Panel B. Respondents included 13 pediatricians, two nurses and four members that described themselves as "other" (family physician, infectious disease, physician-researcher, and researcher). Eight had a sub-specialty in gastroenterology, four in emergency care, and seven in other areas (infectious

disease, intensive care, or public health). Ten respondents (53%) had been an investigator on a combined total of 134 RCTs of PAD or PAG. In particular, the nine Panel A respondents had been investigators on a total of 128 (96%) of the 134 RCTs and each had 15 or more years of clinical experience. Regarding studies evaluating reliability and/or validity, six Panel A experts and two Panel B experts had been investigators on a total of 20 measurement studies (see Table 4-10). Five experts, three from Panel A and two from Panel B, were not actively engaged in clinical practice likely as a result of full-time research responsibilities.

Table 4-10: *Summary of Panelists' Expertise*

	Country	**Profession**	**Specialty**	**Years in Practice**	**Practice setting**	**Avg no. child/wk**	**No. DD RCTs**	**No meas. studies**
Panel A†								
Expert 1	US	Nurs. P	Emer	>20	Comm	30	5	1
Expert 2	Italy	Pedi	Gast	16-20	Acad	5	6	1
Expert 3	US	GP	NA	>20	Acad	0	4	0
Expert 4	Peru	Pedi	Gast	>20	Comm	30	11	1
Expert 5	Poland	Pedi	Gene	16-20	Acad	40	10	0
Expert 6	US	Int. M	Inf. D	>20	Acad	0	55	5
Expert 7	India	GP	Pub. H	>20	Comm	0	20	1
Expert 8	US	Pedi	Gast	>20	Acad	25	5	0
Expert 9	US	Pedi	Gast	16-20	Acad	15	12	1
Median(r)						*15 (40)*	*10 (51)*	*1 (5)*
Sum						*145*	*128*	*10*
Panel B*								
Expert 1	Israel	Pedi	Emer	<5	Acad	20	0	5
Expert 2	Canada	Nurs	P. ICU	11-15	Acad	0	0	5
Expert 3	Canada	Pedi	Emer	>20	Acad	5	0	0
Expert 4	Canada	Pedi	Emer	5-10	Acad	100	6	0
Expert 5	Canada	Padi	Gast	>20	Acad	1	0	0
Expert 6	Canada	Pedi	Emer	5-10	Hosp	30	0	0
Expert 7	Canada	Pedi	Gast	11-15	Acad	5	0	0
Expert 8	Canada	Pedi	Emer	11-15	Acad	48	0	0
Expert 9	US	Rese	NA	11-15	Acad	0	0	0
Expert 10	US	Pedi	Inf. D	11-15	Acad	15	0	0
Median(r)						*10 (100)*	*0 (6)*	*0 (5)*
Sum						*224*	*6*	*10*

Nur.P = Nurse Practitioner; **Pedi** = Pediatrician; **GP** = General Practitioner; **Int.M** = Internal Medicine; **Rese** = Researcher; **Comm** = Community; **Acad** = Academic; **Hosp** = Hospital

†Panel A: Experts who have been in pediatric clinical practice 15 years or more and completed at least three RCTs of Diarrheal Diseseas (DD) (as a co-investigator) or two RCTs of DD as a primary or senior investigator.

***Panel B:** Experts who did not meet the threshold for Panel A

Definitions

The results of the expert panel members' ratings for each of the definitions are presented in Tables 4-11 and 4-12. The JDM_{js} values for both Panel A and B indicate that there were no outlying panelists for the definitions of PAD and PAG (see Table 4-11).

Table 4-11: *JDM_{js}: Definitions and Overall Outlying Experts*

	Diarrhea	**Diarrhea resolution**	**Gastroenteritis**	**Gastroenteritis resolution**	***JDM_{js}***
Panel A					
Expert 1	0	2	0	0	2
Expert 2	1	2	3	1	7
Expert 3	0	0	1	-	N/A
Expert 4	1	1	3	1	6
Expert 5	0	1	2	0	3
Expert 6	0	1	1	0	2
Expert 7	0	0	0	2	2
Expert 8	1	1	1	3	6
Expert 9	0	0	1	3	4
Panel B					
Expert 1	0	2	0	1	3
Expert 2	2	2	0	0	4
Expert 3	1	0	2	2	5
Expert 4	0	0	0	0	0
Expert 5	0	1	2	0	3
Expert 6	0	0	0	0	0
Expert 7	0	0	0	0	0
Expert 8	1	1	1	0	3
Expert 9	1	1	3	3	8
Expert 10	2	1	0	2	5

JDM_{js} = Outlying judge/panelist calculation; **Note:** scores for definitions represent the absolute value of the discrepancy between the rating for each panel member and the median of all panel members'

Table 4-12 indicates that the central tendency ratings were low for three of the four definitions (definition of diarrhea, definition of diarrhea resolution, and definition of gastroenteritis) for Panel A. Panel B results indicate that central tendency ratings were low for two of the four definitions (definition of diarrhea and definition of diarrhea resolution). The accompanying *R*-values for Panel A were moderate for the definition of diarrhea and high (4 or 5) for the remaining definitions (diarrhea resolution, gastroenteritis, and

gastroenteritis resolution). The Panel B *R*-values were all high ("4" for each definition) (see Table 4-12).

Table 4-12: *Central Tendency and Item Ambiguity of Definitions*

	Diarrhea	**Diarrhea resolution**	**Gastroenteritis**	**Gastroenteritis resolution**
Panel A				
Expert 1	2	4	2	4
Expert 2	1	4	5	5
Expert 3	2	2	1	-
Expert 4	3	1	5	5
Expert 5	2	3	4	4
Expert 6	2	3	1	4
Expert 7	2	2	2	2
Expert 8	1	1	1	1
Expert 9	2	2	1	1
Median (R)	*1.8(3)*	*2.33(4)*	*1.75(5)*	*3.83(5)*
Mean	*1.89*	*2.44*	*2.44*	*3.25*
% 4 or 5	*0%*	*22%*	*33%*	*63%*
Panel B				
Expert 1	2	4	4	3
Expert 2	4	4	4	4
Expert 3	1	2	2	2
Expert 4	2	2	4	4
Expert 5	2	3	2	4
Expert 6	2	2	4	4
Expert 7	2	2	4	4
Expert 8	3	3	3	4
Expert 9	1	1	1	1
Expert 10	4	1	4	2
Median (R)	*2.1(4)*	*2.25(4)*	*3.67(4)*	*3.67(4)*
Mean	*2.3*	*2.4*	*3.2*	*3.2*
% 4 or 5	*20%*	*20%*	*60%*	*60%*

Note: scores for definitions represent ratings on a 1 to 5 Likert-type scale

Using content analysis techniques, four themes were identified from the comments and suggestions made by the Panel A members (see Table 4-13). Two themes were identified from the comments made by the Panel B members (see Table 4-14).

Table 4-13: *Themes from Panel A Comments on Definitions*

PAD definition
"Need some type of quantification of increase in the number of stools over daily/weekly normal"
"Need some degree of illness"
"May work as a clinical definition but semi-quantitative terms (e.g., "diminished") make it problematic from a research prospective"
PAD resolution - "normal"
"What is a "normal" stool?"
"What is normal?"
PAD resolution – "24 hours"
"Would use 24 hrs"
"No diarrhea for 24 or 48 hours"
"Should change 12 to 24 hrs"
PAG definition
"I would not think 2 separate definitions (gastroenteritis and diarrhea) are needed"
"I don't think there is a need for separate gastroenteritis scale. Vomiting may or not be part of diarrhea and is simply a symptom not really different in severity"

Table 4-14: *Themes from Panel B Comments on Definitions*

PAD definition
"Can also be defined as a decrease in consistency and an increase in frequency to >3/day"
"Better to indicate a number of stools more than normal"
PAD resolution
"What is a "normal" stool?"
"Return to child's normal bowel movement patterns?"

Based on the comments provided by the Panel A and B experts, the ratings for the definition of PAD appear to have been low because it was not made clear to the panelists that the definitions were constitutive, as opposed to operational. "Constitutive" definitions are used to generate the items (e.g., diarrhea frequency, dehydration, fever) to include in the scale. The items in the scale are then used to operationalize (or quantify) the definition. Regarding the definition of resolution of PAD, the central tendency ratings were higher likely as a result of the fact that we provided more of a quantitative definition. However, quantification was still the major theme as four experts from Panels A and B commented on what a "normal" stool was and three Panel A members took issue with the number of hours without diarrheal stools, suggesting that 24 to 48 hours would be more valid. Our previous systematic review indicated that 24 hours was the most common diarrhea-free

period employed in definitions of diarrhea resolution (Johnston 2008). With respect to the definition of PAG, the central tendency ratings from Panel A appear to have been low as a result of separating PAG from PAD. This finding was supported by comments made for items (see below).

Based on the low medians for the definitions and the major themes that emerged, we collapsed the definitions for PAD and PAG and made two revisions to the "resolution of diarrhea": listed below.

- We revised the definition of PAD to read: *production of 2 consecutive normal stools (i.e., "soft and formed" or "hard and formed" stool)*
- We revised the definition of PAD resolution to read: *normal stool production (or no stool production) and no vomiting for a period of 24 hours*

Table 4-15: *Revised Definitions*

Pediatric Acute Diarrhea & Gastroenteritis
Acute diarrhea is marked by an increase in the frequency of bowel movements and a change in the consistency of stool (i.e., "loose and unformed" or "liquid" stool), and is commonly, but not necessarily, associated with vomiting. Diarrhea may also be associated with fever, restrictions in normal daily activities of the child and dehydration.
Resolution of Pediatric Acute Diarrhea & Gastroenteritis
Acute diarrhea typically lasts less than 7 days and not longer than 14 days and resolution is marked by (a) production of 2 consecutive normal stools (i.e., "soft and formed" or "hard and formed" stool) and absence of vomiting or; (b) production of one normal stool followed by 12 hours with no stool production or vomiting or; (c) normal stool production (or no stool production) and no vomiting for a period of 24 hours.

Item relevance

The results of the expert panel members' ratings for each of the items within each scale are presented in Tables 4-16 and 4-19. Items in the IPADS included: *Diarrhea Frequency, Diarrhea Duration, Dehydration, Pain, Normal Daily Activities*, and *Fever*. *Stool Consistency*, used to qualify *Diarrhea Frequency* and *Duration of Diarrhea*, is also included in the tables. The IPAGS items included each of the IPADS items with the addition of *Vomiting Frequency* and *Vomiting Duration*.

Table 4-16: *Items and Overall Outlying Experts*

	Co	**DF**	**DD**	**De**	**Pa**	**Ac**	**Fe**	***JDM***	**VF**	**VD**	***JDM***
Panel A											
Expert 1	1	3	3	3	1	3	3	**17**	1	0.5	**18.5**
Expert 2	0	3	3	0	2	1	1	7	3	2.5	**12.5**
Expert 3	0	-	-	1	0	0	0	N/A	-	-	N/A
Expert 4	0	0	0	1	1	0	0	2	1	0.5	3.5
Expert 5	1	0	0	1	0	1	0	3	1	1.5	5.5
Expert 6	0	1	0	1	2	1	1	6	0	0.5	6.5
Expert 7	0	1	1	1	2	1	1	7	1	1.5	9.5
Expert 8	1	0	0	1	1	1	2	6	0	2.5	8.5
Expert 9	0	1	1	2	1	0	1	6	1	0.5	7.5
Panel B											
Expert 1	0.5	0	0.5	1	0.5	0	1	3.5	1	0.5	5
Expert 2	0.5	1	2.5	0	0.5	1	2	7.5	1	3.5	**12**
Expert 3	0.5	1	1.5	0	1.5	1	2	7.5	0	0.5	8
Expert 4	0.5	0	1.5	2	1.5	1	0	6.5	0	0.5	7
Expert 5	0.5	1	1.5	0	0.5	0	2	5.5	0	0.5	6
Expert 6	1.5	1	0.5	1	0.5	0	0	4.5	1	0.5	6
Expert 7	0.5	1	0.5	1	1.5	1	1	6.5	0	0.5	7
Expert 8	1.5	3	2.5	0	0.5	0	0	7.5	0	0.5	8
Expert 9	1.5	1	0.5	0	0.5	1	1	5.5	2	1.5	9
Expert 10	0.5	3	0.5	0	0.5	1	2	7.5	0	3.5	11

Co = Stool Consistency; **DF** = Diarrhea Frequency/day; **DD** = Duration of Diarrhea; **De** = Dehydration; **Pa** = Pain; **Ac** = Activity; **Fe** = Fever; **VF** = Vomiting Frequency; **VD** = Duration of Vomiting; ***JDM*** = Outlying judge/panelist calculation; Blank **(-)** = Missing Data; **Note:** scores below items represent the absolute value of the discrepancy between the rating for each panel member and the median of all panel members'

The JDM_{js} values indicate that there were two outlying panel members on Panel A and one outlying member on Panel B (see Table 4-16). We examined which individual items the panelists seemed to find problematic and looked for any comments that may have accompanied these items. The three comments provided by the Panel B member and four comments provided by the two Panel A members' and are provided in Table 4-17 and 4-18, respectively.

Table 4-17: *Comments from Panel B Outlier*

Diarrhea Duration
"I don't think number of days matters. Often, the longer it is going on the less severe the dehydration and clinical implications"
Fever
"Fever phobia is outdated and not relevant. We do not treat children differently whether fever is present or not"
Vomiting Duration
"Not relevant…the impact will be captured in the dehydration score otherwise 1-2 per day does not change anything even if <5 days"

Duration of Diarrhea is an item that will remain given that it was the most common primary outcome identified in our systematic review (Chapter 3). Regarding *Fever*, the accompanying text following Table 4-22 (below), discusses a theme that emerged from Panel B and why this item was retained. With respect to *Vomiting Duration*, if the impact of vomiting is captured in the dehydration score, this can be explored when the instrument is piloted in a sample of children with acute diarrhea. If this item is redundant, empirical evidence can be used to justify its deletion.

Table 4-18: *Comments from Panel A Outliers*

Diarrhea Frequency
"Too complicated, especially in a community setting"
Diarrhea Duration
"Loose/unformed and liquid can be combined - they often occur on the same day"
Dehydration
"Better to use what is recommended by WHO to maintain consistency"
Fever
"The exercise would perhaps be useful if the scenario under which the scores are likely to be used can be made more specific e.g. in outpatient facilities, research settings, etc"

It was felt that the comments on *Diarrhea Frequency* and the *Duration of Diarrhea* were convincing regarding their "complication" and the collapse of loose/unformed stools & liquid stools into one category. We decided to disregard the comments on *Dehydration* and *Fever* for the following reasons. Unlike the WHO scoring systems for dehydration (below), the scoring system we adapted for dehydration has considerable evidence of validity and reliability (Goldman 2008; Friedman 2004). Further, Gorlick (1997) and Duggan (1996) have reported similar items as clinical predictors of dehydration (Gorlick 1997; Duggan 1996). We chose not to adopt the WHO classification instrument (Table 4-18), as it does not incorporate mucous membranes or tears, which are measures that have been demonstrated to be reliable and valid (Gorelick 1997; Steiner 2004; Freidman 2004). In study settings where the majority of children may have clinical features of mild to moderate dehydration the WHO instrument may have a limited ability to distinguish children with "some dehydration" (or mild dehydration) who need only minimal rehydration from children with "some dehydration" (or moderate dehydration) who require rehydration (Freedman 2006).

Table 4-19: *WHO Assessment of Dehydration (2005)*

	A	B	C
Condition	Well, alert	Restless, irritable	Lethargic or unconscious
Eyes	Normal	Sunken	Sunken
Thirst	Drinks normally, not thirsty	Thirsty, drinks eagerly	Drinks poorly, or not able to drink
Skin Pinch	Goes back quickly	Goes back slowly	Goes back very slowly
Decide	The patient has **no signs of dehydration**	If the patient has two or more signs in B, there is **some dehydration**	If the patient has two or more signs in C, there is **severe dehydration**

Regarding the comment on *Fever*, it was made clear that the proposed instruments were for research purposes (in particular, employment in RCTs in inpatient and outpatient settings). We then interpreted this comment as an oversight by the expert panel member.

To confirm our inferences above as well as to address a number of stand alone comments provided by these panelists, we contacted each of the two outlying panel members and the member who had not responded to a number of the questions on item-relevance. Since the three experts (two outliers, one none responder) had conducted the majority of the RCTs (87 of 134 combined RCTs conducted by panels), we thought it was vital to involve these panelists. We also contacted one additional Panel A expert who is a well-recognized author on the pathogenesis and clinical epidemiology of PAD. Given that the major comments provided by the three outlying/non-response experts suggested that diarrhea and gastroenteritis should not be separated, we chose to additionally contact this expert. To do this, we first developed a structured questionnaire. We then attempted to set-up a phone call with each of these experts requesting clarity with respect to their survey responses and endorsement of the revisions that the steering committee had agreed upon. We were able to correspond with three of the four experts (two outliers and the expert on the pathogenesis and clinical epidemiology of PAD). One expert was available via phone and two via email only. In the phone discussion the principal investigator (BCJ) went through each of the questions and recorded the expert's responses. The questionnaire with responses was then sent to this expert for his/her review and to provide an opportunity to revise the responses in the case of any missed information or misinterpretations. For the experts available via

email only, the structured questionnaire was sent to them and they were requested to return it within one week.

Based on our contact with these experts (see Appendix 4-3) we retained the ratings by each of the experts and predicted that the *R*-values for the items these panelists rated low would be high (Table 4-20). These experts endorsed the previous revisions to the definitions listed above and the four "item" revisions below:

- For the *Diarrhea Frequency* and *Diarrhea Duration* items, the collapse of loose/unformed stools & liquid stools into one category
- Revised the *Dehydration* item to include *Capillary refill* and *Skin turgor*
- Reduced the *Normal Daily Activities* item from four response options to three
- Removal of the scoring system for each item (0 to 10)

Of these revisions, the suggestion to add *Capillary refill* and *Skin turgor* has empirical evidence to support the modification of the instrument. In particular, a systematic review of seven studies of the precision and accuracy of clinical signs for evaluating dehydration in children (1 month to 5 years) concluded that the two most useful signs in predicting ≥5% dehydration were capillary refill and abnormal skin turgor (Steiner 2004). Personal communication with the Dr. Freidman, principal investigator on the development and validation of the Clinical Dehydration Scale that we orginally adopted, indicated that *Capillary refill* and *Skin turgor* were not included in the final dehydration scale as a result of relatively low measurement properties as compared to the items that were included (see also Friedman 2004). Dr. Friedman further indicated that the low measurement properties was likely because there was only one child with severe dehydration in the study sample (which is often the case in North American studies). However, in developing countries, severe dehydration as a result of diarrheal diseases is much more prevalent and had Friedman's orginal scale been validated in such a population it is reasonable to assume that these items would have scored higher, and, like the previous literature has indicated, have been the best predictors of ≥5% dehydration (Streiner 2004)

Regarding the scoring system, this can be developed via a theoretical perspective and/or an empirical study. Although the IPADS and IPAGS from the first survey had questions on a scoring system developed from a theoretical perspective, based on feedback from the steering committee it was decided to eliminate the proposed scoring system, which the outlying experts and pathogenesis/clinical epidemiology expert endorsed. As indicated in the discussion section, further study is needed to develop a scoring scale for each item that accurately reflects the degree of severity of the attribute being measured.

The item ambiguity and median values for all nine items are reported in Table 4-20. To meet item relevance criteria, a minimum median score of 3.50 with low R_k was required.

Table 4-20: *Central Tendency of Ratings*

	Co	**DF**	**DD**	**De**	**Pa**	**Ac**	**Fe**	**VF**	**VD**
Panel A									
Expert 1	4	1	1	1	2	1	1	3	3
Expert 2	5	1	1	4	1	3	3	1	1
Expert 3	5	-	-	5	3	4	4	-	-
Expert 4	5	4	4	5	4	4	4	3	4
Expert 5	4	4	4	3	3	5	4	5	5
Expert 6	5	3	4	5	5	5	5	4	4
Expert 7	5	5	5	5	5	5	5	5	5
Expert 8	4	4	4	3	4	5	2	4	1
Expert 9	5	5	5	2	2	4	3	5	3
Median (R)	*4.75(2)*	*3.83(5)*	*4.0(5)*	*4.0(5)*	***3.25(5)***	*4.33(5)*	*4.0(5)*	*4.0(5)*	*3.50(5)*
Mean	*4.67*	*3.375*	*3.50*	*3.67*	***3.22***	*4.00*	*3.44*	*3.75*	*3.25*
% 4 or 5	*100%*	*63%*	*75%*	*56%*	***44%***	*78%*	*56%*	*63%*	*50%*
Panel B									
Expert 1	4	4	3	4	4	4	2	4	3
Expert 2	3	3	1	5	4	5	1	4	1
Expert 3	4	5	5	5	2	3	5	5	5
Expert 4	3	4	2	3	2	3	3	5	4
Expert 5	3	5	5	5	4	4	5	5	5
Expert 6	2	3	4	4	3	4	3	4	5
Expert 7	4	3	4	4	2	3	4	5	5
Expert 8	2	1	1	5	4	4	3	5	5
Expert 9	5	5	4	5	3	5	2	3	3
Expert 10	4	1	1	5	4	5	1	5	1
Median (R)	*3.5(4)*	*3.5(5)*	*3.5(5)*	*4.67(3)*	*3.5(3)*	*4.0(3)*	***2.83(5)***	*4.67(3)*	*4.5(5)*
Mean	*3.40*	*3.40*	*3.00*	*4.50*	*3.20*	*4.00*	***2.90***	*4.50*	*3.70*
% 4 or 5	*50%*	*60%*	*50%*	*90%*	*50%*	*70%*	***30%***	*90%*	*60%*

Co = Stool Consistency; **DF** = Diarrhea Frequency/day; **DD** = Duration of Diarrhea; **De** = Dehydration; **Pa** = Pain; **Ac** = Activity; **Fe** = Fever; **VF** = Vomiting Frequency; **VD** = Duration of Vomiting; Blank **(-)** = Missing Data; **Note:** scores below each item represent ratings on a 1 to 5 Likert-type scale

Given that the JDM_{js} values indicated that there were two outlying Panel A experts (Table 4-16), as predicted earlier, the *R*-values (indicating item ambiguity) (Table 4-20) for some items were also high. Inspection of the central tendency for Panel A experts revealed that the median score for eight of nine items were 3.50 or higher, with six items (66.6%) having a median score of ≥4.0. The item with the highest median score was *Stool Consistency*, followed by four items with the same median score (*Duration of Diarrhea, Dehydration, Fever* and *Vomiting Frequency*). The the median score for *Pain* was the lone score below 3.50 at 3.25.

Using content analysis techniques, three themes were identified from Panel A experts' comments (Table 4-21).

Table 4-21: *Themes from Panel A Comments on Items*

Duration of Diarrhea
"Disagree dichotomy between loose/unformed and liquid"
"Loose/unformed and liquid can be combined – they often occur on the same day"
Dehydration
"Consider capillary refill time"
"Prolonged capillary refill time is not included"
Pain & Discomfort
"Most kids with acute diarrhea don't have much pain – maybe a little irritable due to cramping & nausea"
"Not much point in this"
"Unclear what the relevance pertains to. Usually scoring is done to define severity which would lead to specific action"

Based on the major themes that emerged from the Panel A comments the following revisions were made:

- Collapsed the stool consistency response options "*Loose & Unformed* and *Liquid Stool(s)*" into one category (also confirmed by contact with outlying panelists)
- Revised the *Dehydration* item by adding *Capillary Refill* (also confirmed by contact with outlying panelists)
- Removed the *Pain & Discomfort* subscale (also confirmed by central tendency ratings)

Inspection of the central tendency for Panel B experts revealed that the median score for eight of nine items were 3.50 or higher, with four items (44.4%) having a median score of ≥4.0. The the median score for *Fever* was the lone score below 3.50 at 2.83. Using content analysis techniques, one theme was identified from the Panel B expert comments regarding this item (Table 4-22).

Table 4-22: *Themes from Panel B Comments on Items*

Fever
"Not very relevant as a sign for severity of diarrhea"
"Relevant to sepsis and complications and as a clue to rotavirus enteritis, but not to AD per se"
"Fever phobia is outdated and not relevant. We do not treat children differently whether fever is present or not"

As mentioned above, Panel A ratings took priority over Panel B ratings. In particular, Panel A rated *Fever* as 4.0. Since *Fever* was considered relevant by the Panel A experts, is a symptom that is a clinical feature of 68% of children with invasive bacterial or parasitic acute diarrhea and 26% of children with rotavirus diarrhea (Guandalini 2004), and is a symptom of considerable parental importance (Betz 2006; Crocetti 2001), the steering committee was comfortable retaining this item.

Item Representativeness

We asked the panel members if the items that they identified as relevant (4 or 5) represented the PAD and PAG definitions (Tables 4-23 and 4-24 below). Six of nine (66.6%) Panel A experts rated the PAD items as representative of PAD, whereas only four of the nine (44.4%) experts rated the PAG items as representative of PAG. All Panel B expert members (100%) rated the PAD items as representative of PAD, whereas only seven of the ten (70%) experts rated the PAG items as representative of PAG.

Table 4-23: *PAD Representativeness*

	E1	**E2**	**E3**	**E4**	**E5**	**E6**	**E7**	**E8**	**E9**	**E10**	**Md (*R*)**	**Mean**
Panel A	3	3	-	4	5	5	5	5	5	NA	4.70(3)	4.38
Panel B	4	4	5	4	5	4	4	4	4	4	4.13(2)	4.20

Table 4-24: *PAG Representativeness*

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	Md (*R*)	Mean
Panel A	3	1	3	4	1	5	5	1	4	NA	3.25(5)	3.00
Panel B	3	4	5	4	5	4	3	3	5	4	4.16(3)	4.00

E = Expert; **Md** = Median; (***R***) = range; **Note:** scores below "Experts" represent ratings on a 1 to 5 Likert-type scale

The same Panel A members that rated PAD and PAG as not representative were identical to the two outlying members and the member that did not indicate relevance ratings for four items. Given that we were able to contact two of three of these panel members, based on their responses, we suspect that lower ratings for the representativeness of the items for PAGs is likely a result of the two constructs being separated. Although we cannot confirm this, as we did not ask this direct question to the panel members, it could be deduced that had the constructs and the accompanying eight items appeared as one definition, their ratings would not have been so low. If experts rated the representativeness with three or less, they were asked to indicate what was missing and identify the missing attribute(s). Using content analysis techniques one theme was identified from Panel A members regarding variable(s) recommended for addition to the PAD construct (Table 4-25).

Table 4-25: *Panel A Suggestions for Additional PAD Variables*

Vomiting
"Vomiting and its severity"
"Vomiting should also be included"

Using content analysis techniques, three themes were identified from Panel B members regarding variable(s) recommended for addition to the PAD construct (Table 4-26).

Table 4-26: *Panel B Suggestions for Additional PAD Variables*

Blood in Diarrhea
"Presence of blood in diarrhea"
"Is the presence of blood, mucus a consideration"
Weight Loss
"Weight loss"
"Weight loss but this will be difficult to document prefusion"
Urine Output
"Urine output"
"Poor urine output – reduce, marked reduction, absent in preceding 24 hours"

Using content analysis techniques, no themes were identified from Panel A members and two themes were identified from Panel B members regarding variable(s) recommended for addition to the PAG construct (Table 4-27).

Table 4-27: *Panel B Suggestions for Additional PAG Variables*

Blood in Diarrhea
"Blood pr"
"Presence of blood in diarrhea"
Weight Loss
"Weight loss"
"Weight loss"

Common themes among Panel A and B members was the addition of *Bloody Diarrhea*. In primary care in the UK, the annual incidence of pediatric bacterial infections is estimated to be 1.5 per 1000 and about 50-75 per 100 000 children will develop *Bloody Diarrhea* with these infections. In the developing world, bacterial and amoebic *Bloody Diarrhea* are much more likely to occur. Bacterial intestinal infections that cause bloody diarrhea are *Campylobacter jejuni, Salmonella* species, *Yersinia enterocoliticia*, *Shiga* toxin producing *E. coli* (e.g., 0157H7), and *Shigella* species. Each of these is known as an invasive pathogen, since they destroy the cell walls lining the intestines resulting in bloody stools. The most serious of these infections include *Shiga* toxin producing *E. coli* in the form of hemolytic uremic syndrome and *Shigella* species, where the illness may be life threatening, with septicemia (Murphy 2008).

Although *Campylobacter jejuni* and *Salmonella,* each associated with *Bloody Diarrhea*, are common causes of PAD, Guandalini (2000) identified the most common clinical features of acute diarrhea in 287 European children with a mean age 12.3 months. *Bloody Diarrhea* was relatively uncommon, occurring in 8.3% of cases (Guandalini 2000). Given the prospective studies mentioned above, it appears that bloody stools is not a common clinical feature of acute diarrhea in developed countries, and, more importantly, is not associated with the severity of acute diarrhea. Data on the clinical presentation of PAD in developing countries is unknown.

Regarding *Weight Loss* and *Urine Output*, each of these are measures of dehydration. However, these measures have not been demonstrated to be as reliable and valid as the six items (general condition, eyes, mucous membranes, tears, skin turgor, capillary refill) we have chosen for the measurement of dehydration (Friedman 2004; Steiner 2004). Although many definitions of dehydration are supported by a change or loss in body weight, it is rare that the clinician will have a recent accurate estimate of the child's baseline weight, especially in children >1 year of age, who do not regularly visit their primary healthcare provider (Goldman 2008). Steiner (2004) reported a systematic review of seven studies involving 11 clinical signs, none of which included *Urine Output* (Steiner 2004).

The final revisions in Phase II were based on the literature regarding risk factors for dehydration, which is the most serious clinical consequence of PAD. Three case-control studies were identified, the first involving 379 children aged up to 2 years admitted to the Infectious Disease Hospital in Calcutta (Bhattacharya 1995), the second involving 774 children less than 5 years of age admitted to the Diarrhea Treatment Unit in Nagpur, India (Zodpey 1998), and the third involving 240 Bangladeshi children under 2 years of age (Ahmed 2002). The first study was comprised of 243 children with moderate to severe dehydration and 136 controls with absent to mild dehydration. After controlling for confounders, multivariate analysis identified frequent diarrhea (>8/day) and vomiting (>2/day) as significant risk factors for the development of life-threatening dehydration (Bhattacharya 1995). These findings were reproduced in a hospital-based case-control study of South East Indian children less than 5 years of age. The study included an even larger sample size by way of 387 cases of diarrhea with moderate to severe dehydration and 387 controls with mild or no dehydration (Zodpey 1998). The third study compared 80 children with some to severe dehydration with 160 age-matched controls having no signs of dehydration. Thirty-eight risk factors were studied for probable influence on the development of dehydration and using stepwise logistic regression a combination of vomiting during diarrheal episode, oral rehydration at home, mother's dirty fingernails, and residing more than 3 km from the hospital were found to provide the maximum sensitivity and specificity for predicting dehydration, with vomiting being the most significant factor

(Ahmed 2002). Based on the empirical evidence and discussion among the steering committee the following revision was made:

- Revised the response options for *Diarrhea Frequency* and *Vomiting Frequency*

Based on the overall feedback from the group of 19 experts in PADG, the steering committee, as well as findings in the literature on risk of dehydration as a result of PADG, each of the revisions indicated above was made. These revisions are in italics and underlined in the Table 4-28.

Table 4-28: *International Pediatric Acute Diarrhea and Gastroenteritis Scale (Ver.2)*

1) Stool Consistency Classification System				
Hard and Formed	**Soft but Formed**	**Loose & Unformed**	**Liquid**	
				
Having a hard or firm texture an retaining a definite shape like a banana, cigar or marbles	Retains its general shape; has a texture that appears like *butter*	Lacking any shape of its own having a texture that appears like hot cereal	Like water	
2) Diarrheal Frequency				
	Diarrhea Stools in Preceding 24 hours			
Stool Consistency	0 stools	1 – *4* stools	*5 – 8* stools	≥*9* stools
Loose & Unformed Or Liquid Stool(s)	□	□	□	□
3) Duration of Diarrhea				
	Number of Days with Diarrhea (in previous 13 days)			
Stool Consistency	0 days	1 – 2 days	3 - 4 days	≥5 days
Loose & Unformed Or Liquid Stool(s)	□	□	□	□
4) Vomiting Frequency				
	Episodes of Vomiting in Preceding 24 hours			
Vomiting	0 episodes	*1 episode*	*2 episodes*	*≥3 episodes*
Yes	□	□	□	□
5) Duration of Vomiting				
	Number of Days with Vomiting (in the previous 13 days)			
Vomiting	0 days	1 – 2 days	3 - 4 days	≥5 days
Yes	□	□	□	□
6) Fever				
	Temperature			
Axillary	34.7 to 37.3°C	37.4 to 38.1°C	38.3 to 38.2°C	≥39.0°C
Yes	□	□	□	□
7) Normal Daily Activities				
Restrictions	□	□	□	□
Child's normal daily activities (e.g., *eating,* sleeping, playing, daycare or school)	*Normal*	*Disturbed*	*Unable to participate*	*Admitted to hospital*
8) Dehydration				
	Signs & Symptoms			
General *Behavior*	□ Normal	□ Thirsty, restless or lethargic but irritable when touched	□ Drowsy, limp, cold, sweaty, +/- comatose	
Eyes	□ Normal	□ Slightly sunken	□ Very sunken	
Mucus Membranes (tongue)	□ Moist	□ "Sticky"	□ Dry	
Tears	□ Tears	□ Decreased tears	□ Absent tears	
Skin Turgor	□ *Immediate*	□ *Slow (≤ 2 sec)*	□ *Very slow (>2 sec)*	
Capillary Refill	□ *<1.5 sec*	□ *1.5 to 3 sec*	□ *>3 sec*	

Regarding the scoring system, this can be developed via a theoretical perspective and/or empirically. Although the IPADS and IPAGS from the first survey had questions on a scoring system developed from a theoretical perspective, in revising the instruments, we decided to eliminate the proposed scoring system, which the outlying experts and pathogenesis/clinical epidemiology expert endorsed.

4.7 PHASE III: Endorsement and Appropriateness of Scale

4.7.1 Methods

The experts who responded to Phase II were re-contacted during Phase III. A survey form consisting of results from the first questionnaire (including the group median and range for each of the items and respective scales) was provided to each participant, changes to the definitions and International Pediatric Acute Diarrhea & Gastroenteritis Scale (IPADGS) items were highlighted and participants were then asked about the appropriateness of the multi-attribute scale in three research settings: an inpatient setting, an outpatient setting, and a community setting. In addition, there was an opportunity for participants to provide any last comments regarding suggested modifications (see Appendix 4-4). Non-responders were sent electronic mail reminders up to two times over the two-week data collection period ending mid-December 2008.

4.7.2 Results and Justification

Fifteen of 19 (78.9%) experts responded. Eight Panel A members responded and seven Panel B members responded. Of those respondents, all agreed that IPADGS is appropriate for use in RCTs in an inpatient setting, 13 (86.7%) agreed that it is appropriate for use in an RCT in an outpatient setting and 8 (53.3%) agreed it appropriate for use in an RCT in a community setting.

Based on the feedback from the panelists, the following three revisions were made.

- The IPADGS is for research purposes only, not clinical practice (which should be duly noted on the instrument). To be appropriately employed, study personnel (parents, nurses, physicians) would need to be instructed on how to accurately measure each of the items (such as skin turgor/pinch, capillary refill)

- For the *Dehydration* dimension, insert “pinch” next to skin turgor, (to read: “skin turgor/pinch”)
- For the *Fever* dimension, insert Fahrenheit measures below Celsius measures and revise the first response option (34.7 to 37.3 Celsius) to ≤37.3 Celsius

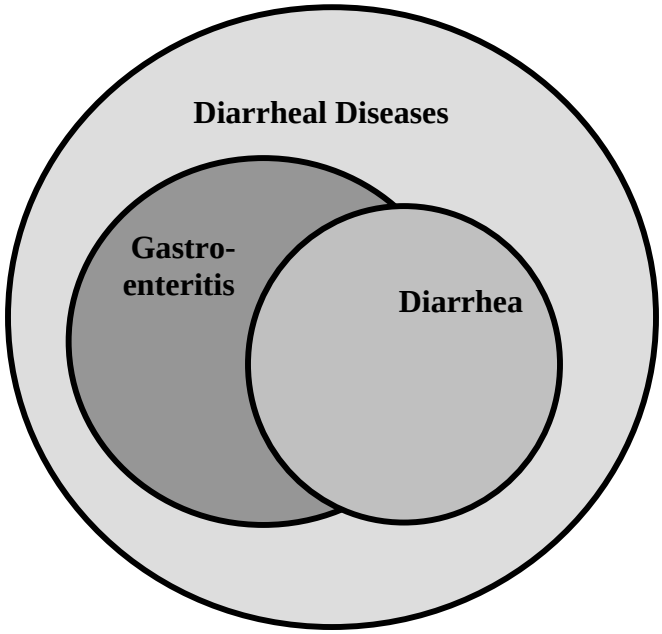
Revised Definition and Scale Name

The need to distinguish diarrhea from gastroenteritis by definition and, from a more practical viewpoint, in a measurement instrument is a challenging issue. As indicated in Figure 4-1 above, the two constructs have significant overlap. The initial intent of this study (Chapter 3 and 4) was to develop an instrument for evaluating the severity of acute diarrhea (i.e., a disease activity index for PAD) in children up to 5 years of age. The genesis of including the term gastroenteritis was a result of our systematic review of PAD literature. This review revealed definitions of diarrhea as well as the three disease activity indexes (Diarrheal Disease Index, Ruuska & Vesikari Scale, and Clark Scale) each of which included vomiting (see Table 4-30). Definitions that included vomiting were often identified as gastroenteritis, as opposed to diarrhea.

Although gastroenteritis is often recognized in the scientific community as diarrhea with or without vomiting, by virtue of its pathophysiologic definition, it means inflammation of the stomach and intestines. It appears that historically it was thought that inflammation of these organs resulted in diarrhea and/or vomiting. However, modern imaging techniques have proven that inflammation is not consistent across the many types of diarrhea such as rotavirus or cholera related diarrhea. While "itis" does mean inflammation for whatever body part is being referred to, it is difficult to overlook how a term is commonly used by a large community of clinicians and scientists. Indeed, popular use also contributes to the definition of a word, not only a word's historical origins. Thus, the term diarrhea versus gastroenteritis is open to different interpretations. For example, from a theoretical point of view gastroenteritis could mean three separate things: 1) inflammation of stomach and intestines; 2) proven bacterial or viral component; or 3) the clinical presentation of both diarrhea and vomiting. Further, for some it may be a combination of some or all of these components. From a classification point of view, this is not ideal.

Given this, rather than the previous definitional term "diarrhea and gastroenteritis," and scale name "IPADGS", we propose IPADDS (International Pediatric Acute Diarrheal Diseases Scale). The advantages of using "diarrheal diseases" over "diarrhea and gastroenteritis" include that fact that the term is inclusive of both gastroenteritis and diarrhea and accepted internationally by the WHO. The disadvantage is that we drop the term gastroenteritis, a term that is well-recognized and commonly used in the medical community.

Figure 4-2: Overlapping Constructs Diarrheal Diseases, Gastroenteritis and Diarrhea



This decision is supported by one of our Panel A experts, a well-known researcher on the topic:

> *"I don't think one can make the differentiation between acute diarrhea and acute gastroenteritis. The term "gastroenteritis" - although admittedly widely used - should be deleted, as the presence of gastric and small intestinal inflammation (as implied in the term) is far from constant in the episodes of acute diarrhea in children. Examples: Rotavirus, Norovirus, Cholera, E. coli, Campylobacter, Yersinia (no gastric involvement); Cholera (no enteritis)"*

Based on feedback from one of the Panel A members and consultation with the steering committee, the following revisions were made.

- Revise title of definition to "Pediatric Acute Diarrheal Diseases" (PADD)
- Revise title of scale to "International Pediatric Acute Diarrheal Diseases Scale"

The definitions and accompanying scale proposed appear in Table 4-29 and 4-30 below.

Table 4-29: *Definitions*

Pediatric Acute Diarrheal Diseases
For children up to 5 years of age, acute diarrhea is marked by an increase in the frequency of bowel movements above normal for the individual and a change in the consistency of stool (i.e., "loose and unformed" or "liquid" stool), and is commonly, but not necessarily, associated with vomiting. Diarrhea may also be associated with fever, restrictions in normal daily activities of the child and dehydration.
Resolution of Pediatric Acute Diarrheal Diseases
For children up to 5 years of age, acute diarrhea typically lasts less than 7 days and not longer than 14 days and resolution is marked by (a) production of 2 consecutive normal stools (i.e., "soft and formed" or "hard and formed" stool) and absence of vomiting or; (b) production of one normal stool followed by 12 hours with no stool production or vomiting or; (c) normal stool production (or no stool production) and no vomiting for a period of 24 hours.

Table 4-30: *International Pediatric Acute Diarrheal Diseases Scale* (Ver.3)*

1) Stool Consistency Classification System				
Hard and Formed	**Soft but Formed**	**Loose & Unformed**		**Liquid**
				
Having a hard or firm texture and retaining a definite shape like a banana, cigar or marbles	Retains its general shape; has a texture that appears like butter	Lacking any shape of its own; having a texture that appears like hot cereal		Like water
2) Diarrhea Frequency				
	Diarrhea Stools in Preceding 24 hours			
Stool Consistency	0 stools	1 – 4 stools	5 – 8 stools	≥9 stools
Loose & Unformed, Or Liquid Stool(s)	□	□	□	□
3) Diarrhea Duration				
	Number of Days with Diarrhea (in previous 13 days)			
Stool Consistency	0 days	1 – 2 days	3 - 4 days	≥5 days
Loose & Unformed, Or Liquid Stool(s)	□	□	□	□
4) Vomiting Frequency				
	Episodes of Vomiting in Preceding 24 hours			
Vomiting	0 episodes	1 episode	2 episodes	≥3 episodes
	□	□	□	□
5) Vomiting Duration				
	Number of Days with Vomiting (in the previous 13 days)			
Vomiting	0 days	1 – 2 days	3 - 4 days	≥5 days
	□	□	□	□
6) Fever				
	Temperature			
Axillary	≤37.3°C *(≤99.1°F)*	37.4 to 38.1°C *(99.3 to 100.6°F)*	38.2 to 38.9°C *(100.7 to 102°F)*	≥39.0°C *(≥102.2°F)*
	□	□	□	□
7) Normal Daily Activities				
	□	□	□	□
(e.g., eating, sleeping, playing daycare or school)	Normal	Disturbed	Unable to participate	Admitted to hospital
8) Dehydration				
	Signs & Symptoms			
General Condition	□ Normal	□ Thirsty, restless or lethargic but irritable when touched	□ Drowsy, limp, cold, sweaty, +/- comatose	
Eyes	□ Normal	□ Slightly sunken	□ Very sunken	
Mucus Membranes (tongue)	□ Moist	□ "Sticky"	□ Dry	
Tears	□ Tears	□ Decreased tears	□ Absent tears	
Skin Turgor/Pinch	□ Immediate	□ Slow (≤2 sec)	□ Very slow (>2 sec)	
Capillary Refill	□ <1.5 sec	□ 1.5 to 3 sec	□ >3 sec	

** for research purposes only*

4.8 DISCUSSION

Diarrheal diseases rank second among conditions afflicting children. Despite the burden of disease and the many existing randomized controlled trials of interventions to prevent and treat pediatric diarrheal diseases, there is no consensus on how to define and measure PADD. This is an important methodological limitation in this field. As such, we surveyed leading experts and front line clinicians on the items for inclusion in a multi-faceted scale for evaluating the efficacy of interventions in RCTs of pediatric acute diarrheal diseases. Overall, the members of the expert panels agreed on the inclusion of 5 attributes containing 13 items (signs or symptoms). Attributes proposed for the IPADDS were: *Diarrheal Frequency* and *Duration, Vomiting Frequency* and *Duration, Dehydration* and *Fever*, as well as, a measure of *Restrictions in Normal Daily Activities*.

Although early in its development, this is the first instrument for measuring PADD having used a systematic, transparent, and reproducible approach. The IPADDS, unlike other instruments for evaluating the severity of PADD, is based on a systematic review of the definitions of diarrhea and primary outcome measures in RCTs of PADD (Johnston 2008). The results of this review were used to inform the development of the IPADDS. Additionally, we employed panels of international, national, and local experts to collect and analyze content evidence for the IPADDS and to judge the utility of the final draft of the IPADDS. Based on a systematic review by Sinha (2008), a critical review of the methods of the 13 other pediatric collaborations working toward standardization of outcomes demonstrated that only two groups reported the use of a systematic review followed by a consensus study regarding items for inclusion in outcome measures (Lux 2004; Smith 1996).

Our consensus techniques may be limited as a result of employing a modified rather than the full version of the classic Delphi method. The classic Delphi method, originally developed by the Rand Corporation, involves up to four rounds and, beginning with the second round, participants are provided with both the overall summary of the ratings for each question from all participants (median and range) and their own ratings (Sackman 1974). This gives participants the opportunity to revise their original ratings in light of

what the group score indicated. As compared to the classic Delphi consensus techniques, our consensus methods involved individual discussions between the principal investigator with steering committee members followed by two surveys, each with a different purpose. Given that most of the combined group scores for each item met our *a priori* cut-off (3.50) after survey one, we interpreted this to represent consensus. However, the high values for item ambiguity mitigate this simple interpretation. In the final survey we provided just the overall scores and a summary of our revisions to the IPADDS and asked about the usability of the instrument in different study settings, without providing the opportunity to gain further consensus. We feel that our modified Delphi method was reasonable given our considerable background work (a systematic review and the use of individual discussions with eight steering committee experts with expertise in the relevant attributes of diarrheal diseases) before proceeding to our electronic surveys. We are not aware of empirical evidence demonstrating that the classic Delphi method of conducting consensus is superior to other methods (Burns 2008; Hasson 2000). If no empirical evidence exists to support one method over the other, the central concern in assessing item content-relevance for instrument construction is the use of a systematic approach. In generating item content-relevance, it is recommended that investigators report on six key issues. These are: 1) characteristics of expert participants and provide justification as to why panelists were selected; 2) number of experts per panel (minimum of 5 experts should provide a sufficient control for chance agreement) (Lynn 1986); 3) use of separate panelists for item generation and item content-relevance; 4) methods used to rate item content relevance and item representativeness; 5) qualitative or statistical methods used to assess experts' ratings and; 6) selection criteria to determine the selection, modification or deletion of items (Dunn 1999). We fulfilled each of these criteria. However, with respect to the "selection criteria" although the selected items met our cut-off point of 3.5, many items had high *R*-values, which reflected item ambiguity. To understand the high *R*-values, which were due to extreme ratings by some panelists, we contacted the outlying and non-responding panel Panel A members, each of whom were leading international opinion leaders in the field of PADD RCTs. Based on further discussion with these experts (see Appendix 4-3), we retained the ratings by all panel members (Table 4-20).

IPADDS Compared to Previous Scales

The IPADDS is more comprehensive than previous scales for measuring the severity of PADD (Table 4-31). For example, the frequency of stools and vomiting, both objective observations, can now be readily quantified. However, the consistency of the stools and the degree of dehydration may be just as important in deciphering the severity of PADD and without a reliable and valid classification, items can be a cause of measurement error. Our previous systematic review identified 10 other grading systems (single and multi-faceted ordinal and continuous scales) for supporting or rating the evaluation or severity of PADD without formal evidence of reliability or validity. Despite this apparent limitation, two of these grading systems have gained acceptance in the diarrhea/gastroenteritis research community (Clark 1988; Ruuska & Vesikari 1990). These two systems have been used in the majority of vaccine trials for the prevention of rotavirus diarrhea with sometimes minor modifications that have not been validated. The IPADDS is comprised of similar attributes. However, to improve the accuracy of measuring the severity of PADD, we have incorporated items previously demonstrated to be valid and reliable. These items include the Stool Consistency Classification System for measuring diarrhea, and for measuring dehydration: capillary refill time, tears (present/decreased/absent), and moisture of mucus membranes. For example, regarding the measurement of diarrhea, the consistency of stool (formed vs. loose vs. liquid) is a pivotal discriminator for determining the diagnosis of diarrhea (i.e., diarrhea stools in the preceding 24 hrs) and its resolution (duration of diarrhea). Without reliable and valid evidence to support what is and what is not considered diarrhea (such as a Stool Consistency Classification System), inter-rater variation alone has the potential to make the difference between a therapy that demonstrates therapeutic significance and no response, and between registered FDA approval and no approval (Cooney 2007).

Table 4-31: *Grading Systems for Pediatric Acute Diarrheal Diseases*

Items *(sign, symptom, behavior)*	18 points (Jacobs *et al.*)	20 points (Ruuska & Vesikari)	24 points (Clark *et al.)*	IPADDS (Johnston *et al.*)
Diarrhea				
1. Stool consistency classification system	No	Yes	No	Yes*
2. No. stools per day	Yes	Yes	Yes	Yes‡
3. Duration, in days	No	Yes	Yes	Yes
Vomiting				
4. No. emesis per day	No	Yes	Yes	Yes‡
5. Duration, in days	No	Yes	Yes	Yes
6. Intensity (i.e., some, frequent)	Yes	No	No	No
Dehydration				
7. Eyes	Yes*	No	No	Yes*
8. Thirst	Yes	No	No	No
9. Skin turgor/pinch	Yes*	No	No	Yes*
10. General Condition	Yes*	No	Yes*	Yes*
11. Duration of poor general condition, in days	No	No	Yes	No
12. Mucus membranes (tongue)	No	No	No	Yes*
13. Tears	No	No	No	Yes*
14. Capillary refill	No	No	No	Yes*
15. Weight loss	No	Yes	No	No
Fever				
16. Degrees (C)	Yes 'not reported'	Yes 'rectal'	Yes 'rectal'	Yes 'axillary'
17. Duration	No	No	Yes	No
Behavioral Symptoms				
18. Activities of daily living	No	No	No	Yes
Abdominal Pain				
19. Intensity	Yes	No	No	No
Treatment				
20. Dehydration/ Hospitalization	No	Yes	No	No
Total # items	**8**	**7**	**8**	**13**

* evidence of validity and reliability; ‡ response options based on risk of dehydration or hospitalization

The IPADDS may be more sensitive to some types of diarrheal diseases than others. For example, for children suffering from AAD, it will be unlikely that they will also suffer from the additional IPADDS attributes such as vomiting, fever, and dehydration. Comparatively, this could result in low scores on the IPADDS in contrast to children suffering from rotavirus diarrhea, who also commonly experience vomiting, fever, and dehydration. To increase the utility of the IPADDS, we decided that the instrument should be sensitive to all potential types of PADD, rather than specific to just some sub-types.

Future Research

Future research directions should involve the evaluation of the properties of the response options, the weighting of the attributes with regards to their importance in contributing to the total overall score, and establishing an overall numeric scoring system (McDowell 2006; Streiner & Norman 2002). After further discussions with my steering committee and my external examiner for my PhD, it was decided that the response options for five items (*Diarrheal Frequency/24 hrs, Diarrhea Duration, Vomiting Frequency/24 hrs, Vomiting Frequency, Fever*) should employ continuous response options (as opposed to the current response options developed by the principle investigator based on a review of the literature, that was not systematic, but subsequently endorsed by Panels A and B). There are four key advantages to using continuous response for each of the five items mentioned above.

1. Since the current response options may result in a ceiling effect (i.e., the response options in the right column for *Diarrheal Frequency* (≥9 diarrheal stools) theoretically reflect more severe diarrhea), continuous response options will prevent data loss by not forcing continuous data into categories. For example, it is not uncommon for infants with acute diarrhea to pass greater than 10 diarrheal stools per day and continuous response options would effectively capture this.

2. The current response options for each of these items are uneven in size (*Diarrhea Frequency/24 hrs* response options are: 0, 1-4, 5-8 and ≥9) and the use of continuous options will potentially avoid this problem.

3. They allow for greater options for statistical analysis. For instance, the current response options are not amenable to parametric statistics. However, by collecting raw data on a continuous scale (e.g., *Diarrhea Frequency/24 hrs*) from a representative sample of children presenting to pediatric emergency with AD, the continuous data will allow for the development of response options (mild, moderate, severe) using regression modeling (i.e., logistic regression or multiple discriminate regression analysis). In these types of regression models the independent/predictor variable needs to be continuous;

whereas, the dependent/outcome variable is required to be discrete or ordinal (mild, moderate, and severe). With continuous data used for the predictor variable, the outcome variable (mild, moderate, severe) could be based on the global impression according to an expert clinician. Regression analysis can then be run to see how the continuous data predict the outcome variables (mild, moderate and severe). The advantage of using regression modeling is that the response options could be further developed based on empirical data from actual children with acute diarrhea, as opposed to expert consensus.

4. By employing continuous response options and exploring regression modeling to further develop each of these items, the IPADDS would be substantially different and independent from previously used instruments (Diarrheal Disease Index, Ruuska & Vesikari Scale, Clark Scale).

Once the score scales for the items have been established, weighting of the five attributes needs to be considered so as to obtain a total score that is reliable and can be validly interpreted as an index of severity of diarrheal disease. Next, overall cut-scores will need to be established in the total score distribution to separate classes of severity. Following these steps, the instrument will need to be piloted to understand any potential differences in comprehension, problems with layout, methods of delivery, and any cultural or other feasibility issues. Finally, the instrument will need to be tested for properties such as inter-rater reliability and criterion validity.

REFERENCES

Ahmed FU, Karim E. Children at risk of developing dehydration from diarrhoea: a case-control study. J Trop Pediatr. 2002;48:259-263.

Avendano P, Matson DO, Long J, Whitney S, Matson CC, Pickering LK. Costs associated with office visits for diarrhea in infants and toddlers. Pediatr Infect Dis J. 1993;12:897-902.

Avner JR. Acute fever. Pediatr Rev. 2009;30:5-13.

Avner JR, Baker MD. Management of fever in infants and children. Emerg Med Clin North Amer. 2002;20:49-67.

Bern C, Martines J, De Zoysa I, Glass R. The magnitude of the global problem of diarrhoeal disease: a ten year update. Bull World Health Organ. 1992;70:705-714.

Betz MG, Grunfeld AF. 'Fever phobia' in the emergency department: a survey of children's caregivers. Eur J Emerg Med. 2006;13:129-33.

Bhattacharya SK, Bhattacharya MK, Manna B, Dutta D, Deb A, Dutta P, Goswami AG, Dutta A, Sarkar S, Mukhopadhaya A, et al. Risk factors for development of dehydration in young children with acute watery diarrhoea: a case-control study. Acta Paediatr. 1995;84:160-164.

Bieri D, Reeve RA, Champion GD, Addicoat L, Ziegler JB. The Faces Pain Scale for the self-assessment of the severity of pain experienced by children: development, initial validation, and preliminary investigation for ratio scale properties. Pain 1990 41:139-150.

Block SL, Vesikari T, Goveia MG, Rivers SB, Adeyi BA, Dallas MJ, Bauder J, et al. Efficacy, immunogenicity, and safety of a pentavalent human-bovine (WC3) reassortant rotavirus vaccine at the end of shelf life. Pediatrics 2007;119:11-18.

Bliss DZ, Dhamani KA, Savik K, Kirk K. Tool to classify stool consistency: content validity and use by persons of diverse cultures. Nurs Health Sci. 2003;5:115-121.

Bliss DZ, Larson SJ, Burr JK Savik K. Reliability of a stool consistency classification system. J Wound Ost Contin Nur. 2001:28:305-313.

Bryce J, Boschi-Pinto C, Shibuya K, Black RE, WHO Child Health Epidemiology Reference Group. WHO estimates of the causes of death in children. Lancet 2005;365:1147-1152.

Burns KE, Duffett M, Kho ME, Meade MO, Adhikari NK, Sinuff T, Cook DJ; ACCADEMY Group. A guide for the design and conduct of self-administered surveys of clinicians. CMAJ 2008;179:245-252.

Carneiro NB, Diniz-Santos DR, Fagundes SQ, Neves LL, Reges RM, Lima EK, et al. Clinical and epidemiological aspects of children hospitalized with severe rotavirus-associated gastroenteritis in Salvador, BA, Brazil. Braz J Infect Dis. 2005;9:525-528.

Clark HF, Borian FE, Bell LM, Modesto K, Gouvea V, et al. Protective effect of WC3 vaccine against rotavirus diarrhea in infants during a predominantly serotype 1 rotavirus season. J Infect Dis. 1988;158:570-587.

Cooney RM, Warren BF, Altman DG, Abreu MT, Travis SPL. Outcome measurement in clinical trials for ulcerative colitis: towards standardisation. Trials 2007;8:17.

Crocetti M, Moghbeli N, Serwint J. Fever phobia revisited: have parental misconceptions about fever changed in 20 years? Pediatrics 2001;107:1241-1246.

Duggan C, Refat M, Hashem M, Wolff M, Fayad I, Santosham M. How valid are clinical signs of dehydration in infants? J Pediatr Gastroenterol Nutr. 1996;22:56-61.

Dunn JGH, Bouffard M, Roger WT. Assessing item content-relevance in sport psychology scale-construction research: Issues and recommendations. Meas Phys Educ Exer. 1999;3:15-36.

Ericsson CD, DuPont HL, Sullivan P, Galindo E, Evans DG, Evans DJ Jr. Bicozamycin, a poorly absorbable antibiotic, effectively treats travelers' diarrhea. Ann Intern Med. 1983;98:20-25.

Fagundes-Neto U, de Andrade JA. Acute diarrhea and malnutrition: lethality risk in hospitalized infants. J Am Coll Nutr. 1999;18:303-308.

Faruque AS, Mahalanabis D, Haque SS, Fuchs GJ, Habte D. Double-blind, randomized, controlled trial of zinc or vitamin A supplementation in young children with acute diarrhoea. Acta Paediatr. 1999;88:154-60.

Flores J, Perez-Schael I, Gonzalez M, Garcia D, Perez M, Daoud N, et al. Protection against severe rotavirus diarrhoea by rhesus rotavirus vaccine in Venezuelan infants. Lancet 1987;1:882-884.

Fortin J, Parent MA. Dehydration scoring system for infants. J Trop Pediatr Environ Child Health. 1978;24:110-114.

Freedman SB, Adler M, Seshadri R, Powell EC. Oral ondansetron for gastroenteritis in a pediatric emergency department. N Engl J Med. 2006;354:1698-1705.

Friedman JN, Goldman RD, Srivastava R, Parkin PC. Development of a clinical dehydration scale for use in children between 1 and 36 months of age. J Pediatr. 2004;145:201-207.

Giaquinto C, Van Damme P, Huet F, Gothefors L, Maxwell M, Todd P, da Dalt L; REVEAL Study Group. Clinical consequences of rotavirus acute gastroenteritis in Europe, 2004-2005: the REVEAL study. J Infect Dis. 2007;195:S26-S35.

Goldman RD, Friedman JN, Parkin PC. Validation of the Clinical Dehydration Scale for Children With Acute Gastroenteritis. Pediatrics 2008;122:545-549.

Gorelick MH, Shaw KN, Murphy KO. Validity and reliability of clinical signs in the diagnosis of dehydration in children. Pediatrics 1997;99:e6.

Guandalini S. Acute Diarrhea, In: Walker AW, Goulet O, Kleinman RE, Sherman PM, Shneider BL, Sanderson IR, editors. Pediatric Gastrointestinal Disease: Pathophysiology, Diagnosis, Management. Toronto: BC Decker Inc., 2004, p.166-79.

Guandalini S, Mazzerella G, Fontana M. Childhood acute diarrhea in Italy. Pediatr Res. 1988;24:A414.

Guandalini S, Pensabene L, Zikri MA, Dias JA, Casali LG, Hoekstra H, et al. Lactobacillus GG administered in oral rehydration solution to children with acute diarrhea: a multicenter European trial. J Pediatr Gastroenterol Nutr. 2000;30:54-60.

Guarino A, Albano F, Ashkenazi S, Gendrel D, Hoekstra JH, Shamir R, Szajewska H. Paediatric Infectious Diseases evidence-based guidelines for the management of acute gastroenteritis in children in Europe: executive summary. Pediatr Gastroenterol Nutr. 2008; 46:S81-S122.

Guyatt G, Montori V, Devereaux PJ, Schünemann H, Bhandari M. Patients at the center: in our practice, and in our use of language. ACP J Club. 2004;140:A11-A12.

Hasson F, Keeney S, McKenna H. Research guidelines for delphi survey technique. J. Advanced Nurs. 2000;32:1008-1015.

Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement. Pain 2001;93:173-83.

Hjelt K, Grauballe PC, Andersen L, Schiøtz PO, Howitz P, Krasilnikoff PA. Antibody response in serum and intestine in children up to six months after a naturally acquired rotavirus gastroenteritis. J Pediatr Gastroenterol Nutr. 1986;5:74-80.

Jacobs J, Jiménez LM, Gloyd SS, Gale JL, Crothers D. Treatment of acute childhood diarrhea with homeopathic medicine: a randomized clinical trial in Nicaragua. Pediatrics 1994;93:719-725.

Johnston BC, Shamseer L, da Costa BR, Vohra S. Outcome measures in clinical trials of pediatric acute diarrhea: a systematic review. Cochrane Colloquia 2008; Oct 3-7, Freiburg, Germany (http://www.imbi.uni-freiburg.de/OJS/cca/index.php/cca/article/view/6893).

Kline MW, Lorin MI. Infectious Diseases: Fever. In: McMillian JA et al., editors. Oski's Pediatrics: Principles and Practice. Philadephia: Lippincott Williams & Wilkins, 3rd Edition, 1999. p. 842-850.

Kosek M, Bern C, Guerrant RL. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. Bull World Health Organ. 2003;81:197-204.

Lam TCM, Klockars AJ. Anchor point effects on the equivalence of questionnaire items. J Educ Meas. 1982; 9:317-322.

Leduc DG, Woods S. Temperature Measurement in Paediatrics. Paed Child Health 2009; 5:273-276.

Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. Scand J Gastroenterol. 1997;32:920–924.

Likert R. A technique for the measurement of attitudes. Arch Psychology 1932; 140, 44-53.

Likert R. A technique for the measurement of attitudes. In: Summers GF, editors. Attitude measurement. Chicago: Rand McNally 1970. p.149-58.

Long KZ, Montoya Y, Hertzmark E, Santos JI, Rosado JL. A double-blind, randomized, clinical trial of the effect of vitamin A and zinc supplementation on diarrheal disease and respiratory tract infections in children in Mexico City, Mexico. Am J Clin Nutr. 2006;83:693-700.

Loonen HJ, Grootenhuis MA, Last BF, de Haan RJ, Bouquet J, Derkx BH. Measuring quality of life in children with inflammatory bowel disease: the impact-II (NL). Qual Life Res. 2002;11:47-56.

Lux AL, Osborne JP. A proposal for case definitions and outcome measures in studies of infantile spasms and West syndrome: consensus statement of the West Delphi group. Epilepsia 2004;45:416-1428.

Lynn MR. Determination and quantification of content validity. Nurs Res. 1986;35:382-385.

Mackowiak PA, Wasserman SS, Levine MM. A critical appraisal of 98.6 degrees F, the upper limit of the normal body temperature, and other legacies of Carl Reinhold August Wunderlich. JAMA. 1992;268:1578-1580.

Mayan, M. Essentials of Qualitative Inquiry. Walnut Creek, CA: Leftcoast Press. 2009. p. 93-98.

Merkel SI, Voepel-Lewis T, Shayevitz JR, Malviya S.The FLACC: a behavioral scale for scoring postoperative pain in young children. Pediatr Nurs. 1997;23:293-297.

McDowell I. Measuring Health: A Guide to Rating Scales and Questionnaires, 3rd Edition, Oxford: Oxford University Press Inc. 2006. p 10-46.

Molbak K, Jensen H, Ingholt L, Aaby P. Risk factors for diarrheal disease incidence in early childhood: a community cohort study from Guinea-Bissau. Am J Epidemiol. 1997;146:273-282.

Morse JM, Field PA: Qualitative Research Methods for Health Professionals, Second edn. London: Sage Press. 1995.

Murphy MS. Management of bloody diarrhoea in children in primary care. BMJ. 2008;336:1010-1015.

Osio CE, Carnelli V. Comparative study of body temperature measured with a non-contact infrared thermometer versus conventional devices. The first Italian study on 90 pediatric patients. Minerva Pediatr. 2007;59:327-336.

Osman O, Brown D, Beattie T, Midgley P. Management of febrile children in a paediatric emergency department. Health Bull (Edinb). 2002;60:33-39.

Peters RP, Zijlstra EE, Schijffelen MJ, Walsh AL, Joaki G, Kumwenda JJ, et al. A prospective study of bloodstream infections as cause of fever in Malawi: clinical predictors and implications for management. Trop Med Int Health. 2004;9:928-934.

Parashar UD, Bresee JS, Glass RI. The global burden of diarrhoeal disease in children. Bull World Health Organ. 2003;81:236.

Riepenhoff-Talty M, Bogger-Goren S, Li P, Carmody PJ, Barrett HJ, Ogra PL. Development of serum and intestinal antibody response to rotavirus after naturally acquired rotavirus infection in man. J Med Virol. 1981;8:215-22.

Rogers WT. Educational Psychology 507: Classical Test Theory. Unpublished textbook, University of Alberta. 2008.

Ruiz-Palacios GM, Perez-Schael I, Velazquez FR, Abate H, Breuer T, Costa-Clemens SA, et al. Safety and efficacy of an attenuated vaccine against severe rotavirus gastroenteritis. N Engl J Med. 2006; 354:11-22.

Ruuska T, Vesikari T. Rotavirus disease in Finnish children: use of numerical scores for clinical severity of diarrhoeal episodes. Scand J Infect Dis. 1990;22:259-267.

Sackman H. Delphi Assessment: Expert Opinion, Forecasting, and Group Process. A Report Prepared for the United States Air Force Project Rand. April 1974. p. 1-118.

Santosham M, Brown KH, Sack RB. Oral rehydration therapy and dietary therapy for acute childhood diarrhea. Pediatr Rev. 1987;8:273-278.

Savarino SJ, Hall ER, Bassily S, Wierzba TF, Youssef FG, Peruski LF et al. Introductory evaluation of an oral, killed whole cell enterotoxigenic Escherichia coli plus cholera toxin B subunit vaccine in Egyptian infants. Pediatr Infect Dis J. 2002;21:322-230.

Schultz AA, Murphy E, Morton J, Stempel A, Messenger-Rioux C, Bennett K. Preverbal, Early Verbal Pediatric Pain Scale (PEPPS): development and early psychometric testing. J Pediatr Nurs. 1999;14:19-27.

Sharieff W, Bhutta Z, Schauer C, Tomlinson G, Zlotkin S. Micronutrients (including zinc) reduce diarrhoea in children: the Pakistan Sprinkles Diarrhoea Study. Arch Dis Child. 2006;91:573-579.

Sinha I, Jones L, Smyth RL, Williamson PR. A systematic review of studies that aim to determine which outcomes to measure in clinical trials in children. PLoS Med. 2008; 29;5:e96.

Smith MA, Leeder SR, Jalaludin B, Smith WT. The asthma health outcome indicators study. Aust N Z J Public Health 1996;20:69-75.

Soetenga D, Frank J, Pellino TA. Assessment of the validity and reliability of the University of Wisconsin Children's Hospital Pain Scale for Preverbal and Nonverbal Children. Pediatr Nurs. 1999;25:670-676.

Steiner MJ, DeWalt DA, Byerley JS. Is this child dehydrated? JAMA 2004;291:2746-2754.

Streiner DL, Norman GR. Health measurement scales: A practical guide to their development and use. Oxford: Oxford University Press; 2002.

Trautner BW, Caviness AC, Gerlacher GR, Demmler G, Macias CG. Prospective evaluation of the risk of serious bacterial infection in children who present to the emergency department with hyperpyrexia (temperature of 106 degrees F or higher). Pediatrics 2006;118:34-40.

Vernacchio L, Vezina RM, Mitchell AA, Lesko SM, Plaut AG, Acheson DWK. Diarrhea in American infants and young children in the community setting: Incidence, clinical presentation and microbiology. Pediatr Infect Dis J. 2006;25:2-7.

Vesikari T, Clark HF, Offit PA, Dallas MJ, DiStefano DJ, et al. Effects of the potency and composition of the multivalent human-bovine (WC3) reassortant rotavirus vaccine on efficacy, safety and immunogenicity in healthy infants. Vaccine 2006;24:4821-4829.

Whelan K, Judd PA, Preedy VR, Taylor MA. Covert assessment of concurrent and construct validity of a chart to characterize fecal output and diarrhea in patients receiving enteral nutrition. J Parenter Enteral Nutr. 2008;32:160-168.

Whelan K, Judd PA, Taylor MA. Assessment of fecal output in patients receiving enteral tube feeding: validation of a novel chart. Eur J Clin Nutr. 2004;58:1030-1037.

Wong D, Baker C. Pain in children: comparison of assessment scales. Pediatr Nurs. 1988;14:9-17.

World Health Organization: Treatment of diarrhea: A manual for physicians and other senior health workers, 2005: http://whqlibdoc.who.int/publications/2005/9241593180.pdf Accessed on March 17, 2009.

Wright J, Gundry SW, Conroy RM. A review of changing episode definitions and their effects on estimates of diarrhoeal morbidity. J Health Popul Nutr. 2007;25:448-455.

Young R, MSc Thesis in Nursing, University of Nebraska, 1999.

Zodpey SP, Deshpande SG, Ughade SN, Kulkarni SW, Shrikhande SN, Hinge AV. A prediction model for moderate or severe dehydration in children with diarrhoea. J Diarrhoeal Dis Res. 1999;17:10-16.

CHAPTER 5: Summary, Conclusions and Implications

5.1 STATEMENT OF STUDY PURPOSE

The initial purpose of this doctoral dissertation was to systematically evaluate the randomized trial literature on the use of probiotics to prevent pediatric AAD and to address the methodological limitations of these trials. Contingent upon what was identified, we planned to develop and conduct a RCT to address these limitations. The initial systematic review identified the absence of a standardized definition or a valid and reliable primary outcome measure for PAD. Based on this, a second systematic review of definitions and primary outcome measures in RCTs of PAD was conducted confirming what was uncovered in our first systematic review. This review revealed three disease activity indexes (Diarrheal Disease Index, Ruuska & Vesikari Scale, and Clark Scale), each having limited evidence of validity and no published evidence of reliability for evaluating PAD morbidity. Recognizing the importance of a valid and reliable primary outcome measure to avoid misleading results regarding the efficacy of interventions evaluated in RCTs, and the need for standardized definitions and outcome measures when pooling trials, we then aimed to address this deficiency. Consequently, for the third study we did not conduct an RCT. Rather, the purpose of this study was to develop and begin to collect validity evidence to support the IPADDS, developed for use in RCTs to assess PAD morbidity in children ≤5 years of age. The IPADDS represents the first pediatric acute diarrhea scale developed using sound systematic and scientific methods. Development of a validated outcome measure for diarrhea will be an important methodological advance, which may improve the internal and external validity of therapeutic trials in this area.

5.2 SUMMARY OF METHODS

Four main procedures were involved in this doctoral work. First, a systematic review and meta-analysis were conducted of RCTs involving probiotics (any specified strain or dose) co-administered with antibiotics (any agent) to assess the potential of probiotics to reduce the incidence of antibiotic-associated diarrhea in children (0 to 18 years of age). Six major electronic databases were searched from inception to 2006, along with three specialized registries and reference lists from included and relevant articles, to identify

relevant studies. In addition, letters were sent to authors of included trials, nutra/pharmaceutical companies, and experts in the field requesting additional information on ongoing or unpublished trials. Methodological quality assessment and data extraction were conducted independently by two authors. Dichotomous data (incidence of diarrhea, adverse events) were combined using pooled relative risks, and continuous data (mean duration of diarrhea, mean daily stool frequency) were summarized as weighted mean differences, along with their corresponding 95% confidence intervals. Adverse events were summarized using risk difference. For overall pooled results on the incidence of diarrhea, *a priori* sensitivity analyses included per protocol versus ITT, random versus fixed effects, and methodological quality criterion. Subgroup analyses were conducted on probiotic strain, dose, definition of antibiotic-associated diarrhea, and antibiotic agent.

Second, a systematic review of RCTs conducted in children involving acute diarrhea as their primary outcome measure was completed. Objectives included documentation of how: 1) acute diarrhea is defined; 2) acute diarrhea is assessed; and 3) clinimetric properties of the outcome measures are reported. To identify relevant studies, we searched four major databases from inception to February 2007 for English-language RCTs in children <19 years of age measuring acute diarrhea as a primary outcome measure. To identify additional articles that utilized or supported the development of measurement instruments, we searched the bibliographies of included RCTs in an attempt to uncover evidence of validity and trace the lineage of their development. Searches were screened using titles of papers and when available, abstracts. The full texts of the selected articles were retrieved and two reviewers independently assessed each article for inclusion. Using a standardized data extraction form, two reviewers independently extracted data items of each included article. Descriptive statistics were employed to illustrate the characteristics of trials measuring acute diarrhea as their primary outcome measure.

Third, to develop the initial instrument to measure the severity of PAD and pediatric acute gastroenteritis (PAG) for employment in RCTs, a steering committee composed of eight experts with relevant clinical and research expertise was established. Based on the

results of the second systematic review, we generated a list of items for consideration in the definition of PAD and PAG, developed draft definitions, and then adopted or adapted existing items and their respective response formats, or developed new items to form the initial draft of the International Pediatric Acute Diarrhea Scale (IPADS) and the International Pediatric Acute Gastroenteritis Scale (IPAGS). We aimed to develop a disease activity index (scale) that incorporated the key attributes and their accompanying signs and symptoms (items) that are common to the clinical presentation of acute childhood diarrhea. For each item, rather then choosing a discrete dichotomous response option, which may obscure important gradients of change for patients and clinicians, we used a continuous approach (i.e., scaling symptoms into a continuum based on symptom severity). In other words, each item was to be accompanied by ordered response options that reflected an increase in severity. The aim of each item was to order the response options so that risk of dehydration and/or hospitalization increased sequentially across the ordinal scale.

Fourth, to refine the initial instruments to measure the severity of PAD and PAG for employment in RCTs we: (i) collected empirical evidence of content validity using independent external panels of local, national and international; and (ii) using the same external panels, assessed the utility of the final instrument. Content validity evidence was collected via an online survey. The members of the external panel of experts first reviewed and assessed the definitions. They then assessed each item in terms of its relevance to the definition to which it was referenced and assessed the appropriateness of the accompanying response options. Once they had completed the item-relevance and appropriateness task, they were asked to consider the items they deemed to be relevant and assess how well this set represented the definition to which the items were referenced. Next, experts were asked to fill out a background questionnaire. This was used to determine their experience with RCTs, measurement and evaluation and determine their clinical background. Based on the results of the analysis of the survey responses, we contacted three experts (two with outlying values and one who did not respond to a substantial portion of survey questions) who had conducted the majority of the RCTs (87 of 134). Based on comments regarding diarrhea versus gastroenteritis, one

additional Panel A expert was contacted who is a well-recognized author on the pathogenesis and clinical epidemiology of PAD and PAG. Finally, experts who responded to the survey on content validity were re-contacted in the second round of the modified Delphi regarding the appropriateness of the revised scale in different research settings. The results from the first survey were provided to all participants (including the group median and range for each of the items and respective scales), and changes to the definitions and scale items were highlighted. Participants were asked about the appropriateness of the scale in three research settings: an inpatient setting, an outpatient setting, and a community setting. There was also an opportunity for participants to provide any final comments regarding suggested modifications.

5.3 SUMMARY OF MAJOR FINDINGS

5.3.1 Probiotics for the Prevention of Pediatric AAD

1. In our meta-analysis of probiotics for the prevention of pediatric AAD, ten studies met the inclusion criteria. To allow for a heterogeneous definition of diarrhea, data (as a binary endpoint) were included based on the primary authors' definition of the presence/absence of diarrhea. Nine studies (n = 1946) reported incidence of diarrhea. Trials included treatment with either *Lactobacilli spp*., *Bifidobacterium spp*., *Streptococcus spp*., or *S. boulardii*, alone or in combination. The per protocol analysis for 9/10 trials (n = 1,946) reporting on the incidence of diarrhea showed statistically significant results favouring probiotics over active/non active controls (Relative Risk 0.49; 95% CI 0.32 to 0.74, I^2 = 71%). However, ITT analysis showed non-significant results overall (Relative Risk 0.90; 95% CI 0.50 to 1.63, I^2 = 93%).

2. This review included two trials (Benhamou 1999; Erdeve 2004) not included in previous meta-analyses of probiotics for pediatic AAD (Johnston 2006; McFarland 2006; Sazawal 2006) and was the first published meta-analysis in the child or adult population to conduct an ITT analysis. Concerning previous meta-analyses in the general population, our results are consistent with the results of four earlier meta-analyses: D'Souza et al. (9 RCTs included) used a per-protocol analysis (Odds Ratio 0.37; 95% CI 0.26, 0.53) (D'Souza 2002); whereas, Cremonini et al. (7 RCTs) reported that trials with a loss to follow-up of 15% or greater would be excluded (RR 0.40; 95% CI 0.27, 0.57)

(Cremonini 2002). Sazawal 2006 (19 RCTs) and McFarland 2006 (25 RCTs) also reported similar results (Relative Risk 0.48; 95% CI 0.35, 0.65, I^2 = 53% and Relative Risk 0.43; 95% CI 0.31, 0.58, p <0.001), respectively.

3. Five of ten trials monitored for adverse events (n = 647) and none reported a serious adverse event. Meta-analysis demonstrated no statistically significant differences in the incidence of adverse events (Risk Difference 0.02; 95% CI -0.02 to 0.06).

4. *Probiotic strain:* Two of nine trials administered *Lactobacillus GG* (*Lactobacillus casei spp rhamnosus*) (Arvola 1999; Vanderhoof 1999), while three studied the yeast *Saccharomyces boulardii* (Benhamou 1999; Erdeve 2004; Kotowska 2005). The summary statistic from the *Saccharomyces boulardii* trials (n = 1,328) was not statistically significant (Relative Risk 0.45; 95% CI 0.14 to 1.48, I^2 = 88.1%). Combined results from two studies (n = 307) were statistically significant indicating a protective effect for *Lactobacillus GG* (RR 0.30; 95% CI 0.15 to 0.58, I^2 = 0%). *Lactobacillus sporogenes* also provided evidence of efficacy (Relative Risk 0.47; 95% CI 0.29 to 0.77) for the prevention of childhood AAD (LaRosa 2003).

5. *Dose*: The daily dosage of probiotic(s) varied greatly (2 to 40 billion Colony Forming Units (CFU)/day). Eight of nine studies that reported incidence of diarrhea data provided dosage information. The *a priori* subgroup analyses on dose compared <5 billion CFU/day versus ≥5 billion CFU/day. Five studies providing children with 5 - 40 billion bacteria/yeast cells per day showed evidence for the preventative effects of probiotics (Relative Risk 0.35; 95% CI 0.25 to 0.47); whereas, three studies providing <5 billion CFU bacteria/yeast per day demonstrated non-significant results when combined (Relative Risk 0.89; 95% CI 0.53 to 1.48, I^2 = 61.4%). A chi square test revealed statistically significant dose related heterogeneity
(p = 0.0004).

6. *Definition:* The criteria for diarrhea varied amongst the studies and only eight studies defined diarrhea. Five studies (Arvola 1999; Benhamou 1999; Correa 2005; Erdeve 2004; Kotowska 2005) defined diarrhea synonymous to World Health Organization criteria: three or more abnormally loose bowel movements per 24 to 48 hours (Relative Rsik 0.45; 95% CI 0.23, 0.91). Tankanow 1990 defined diarrhea as one or more abnormally loose bowel movements per 24 hours (Relative Risk 0.96; 95% CI 0.61,

1.50). LaRosa 2003 defined diarrhea as at least two liquid bowel movements per 24 hour period (Relative Risk 0.47; 95% CI 0.29, 0.77). Vanderhoof 1999 defined diarrhea as two or more loose stools on two or more occasions throughout the study period (Relative Risk 0.29; 95% CI 0.13, 0.63).

7. Although the primary outcome in the meta-analysis was based on the incidence of diarrhea, as defined by the authors' of included RCTs, there were seven unique definitions among included trials. None of the trials included justified their choice of definition. Future trials would benefit from standard definitions of diarrhea and a valid and reliable primary outcome measure for AAD that is sensitive to change and reflects what treatment effect clinicians, parents, and children consider important.

5.3.2 Definitions & Primary Outcome Measures in PAD RCTs

1. The subsequent systematic review of RCTs identifying PAD as a primary outcome measure yielded 121 RCTs reporting one or more primary endpoints related to PAD. Authors used 62 different definitions of acute diarrhea and 64 different definitions of what constituted resolution of diarrhea. Trials used 62 different primary outcome measures related to diarrhea, the most common of which was duration of diarrhea (n = 65 trials). Thirty-one trials used grading systems (e.g., scale, scoring system or questionnaires with one or more signs, symptoms, behaviors) to support outcome measure assessment. Of these, three trials stated that their grading system was valid; however, none of the trials (or their citations) reported any evidence of this.

2. Even in what would be considered methodologically sound clinical trials (median methodological quality on 0-5 Jadad Scale was 3.0 (range = 1 to 5), definitions of diarrhea, outcome measure selection, and grading systems employed in PAD RCTs are heterogeneous.

3. The problem of heterogeneity is two-fold. Given that there were 62 distinguishable primary outcome measures employed among 121 trials represents the potential for difficulty when trying to interpret, compare or synthesis results. The difficulties are substantial, but the problem gets considerably worse when one considers that in many instances the primary outcome measure (e.g., "duration of diarrhea") is measured in

different ways (62 distinguishable definitions of diarrhea and 64 distinguishable definitions of what constitutes the resolution of diarrhea).

4. The difficulties with the interpretation, comparison and synthesis of results compounds if we consider that grading systems (e.g., a stool consistency scale), or lack thereof, to support the evaluation of outcome measures may not have evidence of reliability or validity.

5. Given these results, the external validity of the included RCTs is limited and the opportunity to conduct meaningful knowledge synthesis is impeded. Similarly, a systematic review of studies that aim to determine which outcome measures to measure in clinical trials of children demonstrated that few studies have been done to address the choice of endpoints for clinical trials, and in most pediatric subspecialties no research has been conducted (Sinha 2007).

6. Uniform definitions of diarrhea and outcome measures would be of significant benefit to trial end-users (clinicians, patients, pharmaceutical industry, and health policy makers). Standardizing definitions and outcome measures across clinical trials of any particular condition makes it easier to interpret, compare, and synthesis results of RCTs so that inferences regarding the efficacy of different interventions for PAD are not misleading. Standard criteria are not meant to impede the development or use of other criteria, but would represent criteria routinely used and reported. Investigators wishing to employ other criteria in a particular trial should be encouraged to do so, but when reporting their trial, selective reporting could be avoided (i.e., outcome reporting bias) through reporting of core criteria (Clarke 2007; Chan 2008). Such an effort would require consensus, guidelines, and adherence on behalf of the relevant stakeholders.

7. A second finding from our systematic review of RCTs involving PAD was the lack of valid and reliable outcome measures. Given the public health importance of acute diarrhea and the fact that there are therapies known to substantially reduce morbidity and mortality (ORS, vaccines), there is now a need for head-to-head comparative trials (e.g., non-inferiority trials of glucose based ORS versus rice-based ORS) to elucidate the most effective therapies to treat and prevent this condition. To avoid misleading results valid and reliable outcomes are needed.

8. Defintions and outcome measures supported by consensus techniques would require additional empirical evaluation of the measurement properties relative to different diarrheal diseases.

5.3.3 Development of Definitions and Scale to Evaluate PAD

1. Given that the systematic review of RCTs of PAD found definitions and outcome measures that sometimes included vomiting (often referred to as gastroenteritis) and other times did not; the decision taken was to develop two definitions, one for diarrhea and one for gastroenteritis.

2. Based on the definitions for PAD and PAG, we then adopted or adapted existing items and their respective response formats, or developed new items with an accompanying scoring scale to form the initial draft of the IPADS and the IPAGS. Consensus among the steering committee was reached on six items for inclusion in the IPADS (*Diarrhea Frequency, Duration of Diarrhea, Fever, Pain & Discomfort, Normal Daily Activities, Dehydration*) and eight items for inclusion in IPAGS (*Vomiting Frequency* and *Vomiting Duration* in addition to each of the items listed for IPADS).

5.3.4 Empirical Evidence of Content Validity of Scale

1. The modified Delphi consensus study involving an independent external panel of 19 local, national and international experts resulted in the collapsing of the definitions for PAD and PAG (PADG) and the collapse of the IPADS and IPAGS into one scale (IPADGS). Upon further discussions with the steering committee it was decided to drop the term "gastroenteritis" as, by virtue of the pathophysiological definition of gastroenteritis, inflammation of the stomach and small intestines is not consistent across many types of diarrhea such as *Cholera* and *Rotavirus*. It was decided that the term "diarrheal diseases" was more fitting. The advantages of using "diarrheal diseases" over "diarrhea and gastroenteritis" include that fact that the term is inclusive of both gastroenteritis and diarrhea and is accepted internationally by the World Health Organization.

2. The proposed instrument, the IPADDS, contains items to measure both diarrhea and vomiting. The external experts agreed on the inclusion of five attributes containing 13

items (i.e., signs, symptoms, behaviors). Attributes proposed for the IPADDS include: *Diarrhea Frequency* and *Duration, Vomiting Frequency* and *Duration, Fever*, *Restrictions in Normal Daily Activities*, and *Dehydration.*

3. The IPADDS is more comprehensive than three previous scales we identified in our systematic review for measuring the severity of PADD (15-point modified Diarrheal Index Score, the 20-point Ruuska & Vesikari Scale, and the 24-point scale proposed by Clark (1988)) (Table 31, Chapter 4). The IPADDS is comprised of similar attributes. However, we have used consensus techniques to justify which attributes and response formats should be included, and to improve the accuracy of measuring these attributes we have attempted to incorporate items previously demonstrated to be valid and reliable (such as the Stool Consistency Classification System, dehydration items).

4. Given the developments to date, the use of the IPADDS is restricted. Additional research is required (see Implications for Future Research).

5.4 MAJOR LIMITATIONS OF STUDIES

Based on current understandings, the major limitations of each study method are as follows. First, the systematic review of probiotics was meta-analyzed based on authors' chosen definition of the incidence of diarrhea. The definitions varied (7 definitions among 9 studies meta-analyzed), so an argument can be made against the choice to pool the studies. A subgroup analysis of the different definitions demonstrated that three of four subgroups were statistically significant. In particular, one subgroup involving a single study employed the most conservative definition of diarrhea (one or more abnormally loose bowel movements per 24 hours), and the results were non-significant (RR 0.96; 95% CI 0.61, 1.50) (Tankanow 1990). Since a trend between the definition of diarrhea used in the included trials and efficacy was not found (increased probiotic efficacy in trials with more conservative definitions of AAD e.g., one or more loose stools per day), we felt that the pooling of definitions was appropriate. Given the results from chapter 3 and 4, our systematic review results (chapter 2) are still appropriate (when interpreted with the cautionary notes in the discussion section of chapter 2 regarding the lack of a valid and reliable outcome measures employed in included trials). In hindsight, from a theoretical standpoint, it may be that the confidence interval accompanying the

relative risk (RR 0.49; 95% CI 0.32 to 0.74, per-protocol), could possibility be wider, or narrower, than reported based on the fact that no trials used a valid and reliable primary outcome for diarrhea. Future updates of this Cochrane review should continue to pool studies, but also conduct subgroups to further explore any statistical heterogeneity and continue to call for standard definitions and valid and reliable outcomes in future trials.

Second, regarding our subsequent systematic review of definitions and primary outcome measures in RCTs of PAD, we did not search the non-English literature, observational studies or contact authors of included trials. It is possible that a valid and reliable definition, outcome measure or grading system was missed as a result of this decision. However, our aim was not to identify every RCT of PAD published to obtain a cumulative point-estimate around the efficacy of interventions through meta-analysis. Rather, we aimed to acquire a comprehensive sample of PAD trials for evaluation. Our results decisively demonstrate the lack of uniformity in the field regarding the lack of standard definitions of diarrhea and reported evidence of validity or reliability of outcome measures and grading systems. Although we did not ask participants directly, our subsequent study involving surveys of many of the leading international clinician-scientists in the field did not turn up any additional definitions, outcome measures or grading systems.

Third, it can be argued that our consensus techniques may be limited as a result of employing a modified rather than the full version of the classic Delphi method (Sackman 1974). However, we followed published guidelines involving the six key issues recommended when generating item content-relevance and representativeness. These are: 1) the number of experts per panel (minimum of 5 experts should provide a sufficient control for chance agreement) (Lynn 1986); 2) report characteristics of expert participants and provide justification as to why panelists were selected; 3) use of separate panelists for item generation and item content-relevance; 4) methods used to rate item content relevance and item representativeness; 5) qualitative or statistical methods used to assess experts' ratings and; 6) *a priori* criteria to determine the selection, modification or deletion of items (Dunn 1999).

Fourth, the modifed Delphi study was limited to the evaluation of the relevance and representativeness of the items (content validity). Although we developed a scoring system in phase 1, we decided to remove this based on its limitations (theroretically-driven as opposed to data-driven). Further empirical study is needed to develop a scoring scale for each item that accurately reflect the degree of severity of the attribute being measured.

5.5 CONCLUSION

The fact that hundreds of different probiotic strains, or, combination of strains, and doses are readily available in pharmacies, health food and grocery stores as capsules and yogurts; it is imperative that we continue to gather the necessary evidence to best advise patients, families, and clinicians who wish to use probiotics for the treatment or prevention of diarrhea. Our meta-analysis demonstrated that for the prevention of AAD, ≥5 billion CFU of *Lactobacillus GG* or *S. boulardii* demonstrate the most promise. However, even in what would be considered methodologically sound clinical trials, outcome measures (e.g., the incidence of diarrhea) employed in these trials used heterogeneous criteria for defining diarrhea. Consensus among stakeholders on what definition of diarrhea should be used to operationalize the outcome measure "incidence of diarrhea" is needed. As well, reliable and valid grading systems (e.g., Stool Consistency Classification System) are needed to assess the incidence of diarrhea (Bliss 2001; Bliss 2003). For these reasons, the current data are promising, but it is premature to routinely recommend probiotics for the prevention of pediatric AAD.

The lack of consensus in the field was further confirmed in our systematic review of the definitions, primary outcome measures and grading systems employed in 121 RCTs of PAD. It was concluded that the lack of uniformity regarding the definitions of diarrhea and the lack of valid and reliable primary outcome measures limits the insights that clinicians, patients and health policy makers can glean from RCT results. Uniform definitions of diarrhea would be of significant benefit to trial end-users as it would facilitate easier interpretation, comparison, and synthesis of results of RCTs. Moreover,

it was found that the difficulties as a result of heterogeneous definitions were compounded given the lack of reported validity/reliability evidence for outcome measures and grading systems (e.g., a stool consistency scale) employed in trials. Given the public health importance of acute diarrhea and the fact that there are therapies known to substantially reduce morbidity and mortality (ORS, vaccines), there is now a need for head-to-head comparative trials (e.g., non-inferiority trials of glucose based ORS versus rice-based ORS) to elucidate the most effective therapies to treat and prevent this condition. To avoid misleading results valid and reliable outcomes are needed.

Using results from our systematic reviews, an eight member steering committee and an independent modified Delphi consensus procedure involving 19 additional leading clinician-scientists in the field, we have made steps towards developing a grading system (the IPADDS) with preliminary properties of validity for RCTs that local, national and international clinicians-scientists have endorsed. Before employing the IPADDS to evaluate the therapeutic potential of different interventions (e.g., probiotics for AAD), the IPADDS requires further research including the development of a scoring system, and additional empirical evidence of reliability and validity. Each of the proposed next steps will need to be considered relative to the child population in question. For example, numeric scoring and disease severity criteria may differ depending on the sub-population (e.g., AAD *versus* rotavirus diarrhea) and additional evidence of validity will need to be collected specific to the population under question.

Finally, while we started this work with a focus on probiotic evaluation in pediatric AAD, our findings (lack of standard definitions, lack of valid and reliable outcomes, development of a new outcome measure for PADD) are relevant to the entire field of PAD, and not exclusive to probiotics and AAD alone.

5.6 IMPLICATIONS FOR PRACTICE

5.6.1 Meta-analysis of Probiotics for Preventing AAD

1. The current data are promising, but inconclusive. The use of *Lactobacillus GG*, or *S. boulardii* at a dose of 5 to 40 billion CFU/day appear to hold promise as an option for co-

administration with antibiotics, but there is insufficient evidence to recommend their routine use at this time.

2. No serious adverse events were reported in the included studies. However, safety is better assessed from the population-based studies that have been published (Saxelin 1996).

5.7 IMPLICATIONS FOR RESEARCH

5.7.1 RCTs of Probiotics for Preventing AAD

Future RCTs need to:

1. Administer probiotic strains and doses with the most promising evidence (e.g., *Lactobacillus GG*, or *S. boulardii* at 5 - 40 billion CFU/day);

2. Determine the effect of age (i.e., infant, child, adolescent), and antibiotic duration (e.g., 5 days versus 10 days) on probiotic efficacy;

3. Define potential adverse events *a priori* and monitor for these adverse reactions accordingly;

4. Limit losses to follow-up and analyze results using ITT (given that our ITT analysis was non-significant, likely as a result of substantial losses to follow-up in four included trials);

5. Identify and employ a primary outcome measure for AAD that reflects what clinicians and patients (or their proxies) consider important with regards to its signs and symptoms (e.g., stool frequency and consistency, functional status).

5.7.2 Definitions and Outcome Measures in PADD RCTs

1. We recommend that the international community of diarrhea/gastroenteritis investigators collaborate to standardize the design, assessment and reporting of outcome measures in PAD trials. Consensus techniques should be used to determine: 1) definitions of diarrhea, definitions of diarrhea resolution and; 2) core set of outcome measures, including physiologic, clinical and health related quality of life measures, that are important to the major stakeholders who will use the outcome measures. Clinicians and patients (or their proxies) should be involved in obtaining consensus on the criteria and outcome measures that represent important change.

2. Upon consensus of definitions of diarrhea and outcome measures, empirical evidence of the reliability and validity of the criteria, outcome measures and grading systems is needed.

5.7.3 Development of IPADDS for use in RCTs

Further steps are needed to refine the IPADDS and collect needed reliability and validity results. For example,

1. A scoring system needs to be developed that accurately reflects the degree of severity across each proposed response scale;

2. A system for weighting the attributes needs to be considered, developed and validated;

3. Different methods of operationalizing the IPADDS need to be explored and disease severity criteria that are important to clinicians and patients (or their proxies) need to be developed;

4. Strategies for the interpretation the IPADDS, including whether the IPADDS yields scores that reflect a compensatory or conjunctive model of interpretation, the definitions of different levels of severity and the setting of cut-scores in the score distribution(s) of the IPADDS that differentiate these levels, need to be defined and operationalized;

5. Additional criterion related validity evidence is needed prior to the use of the IPADDS in RCTs;

6. A system for analyzing change/responsiveness on the IPADDS needs to be developed and validated

7. The IPADDS needs to be further evaluated for reliability among parents and clinicians. For example, inter-rater reliability among parents of IPADDS items needs to be determined, including *Diarrheal Frequency* and *Diarrhea Duration* using the Stool Consistency Classification System to support item evaluation; and

8. Each of the above suggestions should be considered in context of the population under investigation. For example, a trial that involves probiotics (as an adjunct to antibiotics) for the prevention of AAD will likely have different cut-scores and score distributions than a trial of children given probiotics and followed for the risk of rotavirus diarrhea.

REFERENCES

Arvola T, Laiho K, Torkkeli S, Mykkanen H, Salminen S, Maunula L et al. Prophylactic Lactobacillus GG reduces antibiotic-associated diarrhea in children with respiratory infections: a randomized study. Pediatrics 1999;104:e64.

Benhamou PH, Berlier P, Danjou G, Plique O, Jessueld D, DuPont C. Antibiotic-associated diarrhoea in children: A computer monitored double-blind outpatients trial comparing a protective and a probiotic agent. Med Chir Dig. 1999;28:163-168.

Bliss DZ, Dhamani KA, Savik K, Kirk K. Tool to classify stool consistency: content validity and use by persons of diverse cultures. Nurs Health Sci. 2003;5:115-121.

Bliss DZ, Larson SJ, Burr JK Savik K. Reliability of a stool consistency classification system. J Wound Ost Contin Nur. 2001:28:305-313.

Clarke M. Standardising outcomes for clinical trials and systematic reviews. Trials 2007;8:39.

Chan AW. Bias, spin, and misreporting: time for full access to trial protocols and results. PLoS Med. 2008;5:230.

Correa NB, Peret Filho LA, Penna FJ, Lima FM, Nicoli JR. A randomized formula controlled trial of Bifidobacterium lactis and Streptococcus thermophilus for prevention of antibiotic-associated diarrhea in infants. J Clin Gastroenterol. 2005;39:385-389.

Cremonini F, DiCaro S, Nista EC, Bartolozzi F, Capelli G, Gasbarrini G, Gasbarrini A. Meta-analysis: the effect of probiotic administration on antibiotic associated diarrhea. Aliment Pharmacol Ther. 2002;16:1461-1467.

D'Souza A, Rajkumar C, Cooke J, Bulpitt CJ. Probiotics in prevention of antibiotic associated diarrhea: meta-analysis. BMJ 2002;324:1361-1364.

Erdeve O, Tiras U, Dallar Y. The probiotic effect of Saccharomyces boulardii in a pediatric age group. J Trop Pediatr. 2004;50:234-236.

Kotowska M, Albrecht P, Szajewska H. Saccharomyces boulardii in the prevention of antibiotic-associated diarrhoea in children: a randomized double-blind placebo-controlled trial. Aliment Pharmacol Ther. 2005;21:583-590.

LaRosa M, Bottaro G, Gulino N, Gambuzza F, Di Forti F, Ini G, et al. Prevention of antibiotic-associated diarrhea with Lactobacillus sporogens and fructo-oligosaccharides in children: A multicentric double-blind vs placebo study. Minerva Pediatr. 2003;55:447-452.

McFarland LV. Meta-Analysis of probiotics for the prevention of antibiotic-associated diarrhea and the treatment of Clostridium difficle disease. Am J Gastroenterol. 2006;101:812-822.

Sazawal S, Hiremath G, Dhingra U, Malik P, Deb S, Black RE. Efficacy of probiotics in acute diarrhea: a meta-analysis of masked, randomized, placebo-controlled trials. Lancet Infect Dis. 2006;6:374-382.

Johnston BC, Supina AL, Ospina M, Vohra S. Probiotics for the prevention of pediatric antibiotic-associated diarrhea. Cochrane Database Syst Rev: Cochrane Library 2007, Issue 2. CD004827.

Streiner DL, Norman GR. Health measurement scales: A practical guide to their development and use. Oxford: Oxford University Press; 2002.

Sackman H. Delphi Assessment: Expert Opinion, Forecasting, and Group Process. A report Prepared for the United States Air Force Project Rand. April 1974, p. 1-118.

Lynn MR. Determination and quantification of content validity. Nurs Res. 1986;35:382-385.

Dunn JGH, Bouffard M, Roger WT. Assessing item content-relevance in sport psychology scale-construction research: Issues and recommendations. Meas Phys Educ Exer. 1999:3;15-36.

Ruuska T, Vesikari T. Rotavirus disease in Finnish children: use of numerical scores for clinical severity of diarrhoeal episodes. Scand J Infect Dis. 1990;22: 259-267.

Sinha I, Jones L, Smyth RL, Williamson PR (2008) A systematic review of studies that aim to determine which outcomes to measure in clinical trials in children. PLoS Med. 5: e96.

Vanderhoof JA, Whitney DB, Antonson DL, Hanner TL, Lupo JV, Young RJ. Lactobacillus GG in the prevention of antibiotic-associated diarrhea in children. J Pediatr. 1999;135:564-568.

APPENDIX 2-1 (Systematic Review of Probiotics for Pediatric AAD)

MEDLINE Search Strategy (search dates 1966 to August 29, 2006)

1. exp PROBIOTICS/tu or probiotic$.tw.
2. exp LACTOBACILLUS/ or lactobacill$.tw. or "l acidophilus".tw. or "l casei".tw. or bifidobacter$.mp. or "b infantis".tw. or "b bifidum".tw. or "b longum".tw. or saccharomyce$.mp. or "s boulardii".tw. or clostridium butyricum.tw. or clostridium difficile.mp. or "streptococcus thermophilus".tw. or enterococcus faecium.mp.
3. exp antibiosis/ or biotherapeutic agent$.tw.
4. or/1-3
5. exp Anti-Bacterial Agents/ or antimicrobial$.tw. or antibiotic$.tw.
6. ((antimicrobial or anti microbial or antimycrobial or antimycobacteri$ or antibacteri$ or bacteriocid$) adj3 agent$).tw.
7. 5 or 6
8. exp DIARRHEA/ or diarrhea.tw. or diarrhoe$.tw. or diarhe$.tw. or diahoe$.tw. or dysenter$.tw. or gastro enteritis$.tw. or gastroenteriti$.tw.
9. and/4,7-8
10. child/ or infant/ or adolescence/ or exp infant, new born/ or exp child, preschool/
11. (child$ or newborn$ or adolescen$ or infan$).tw.
12. (preschool$ or pre-school$).tw.
13. teen$.tw.
14. (kindergarten$ or kindergarden$).tw.
15. elementary school$.tw.
16. secondary school$.tw.
17. nursery school$.tw.
18. high school$.tw.
19. highschool$.tw.
20. youth$.tw.
21. (baby$ or babies$ or preemie$ or premature$).tw.
22. (schoolchild$ or "school child$").tw.
23. (schoolage$ or school age$).tw.
24. toddler$.tw.
25. pubert$.tw.
26. (pre-pubescen$ or prepubescen$ or post-pubescen$ or postpubescen$).tw.
27. (kid or kids or boy$ or girl$).tw.
28. juvenile.tw.
29. or/10-28
30. 9 and 29
31. RANDOMIZED CONTROLLED TRIAL.pt.
32. CONTROLLED CLINICAL TRIAL.pt.
33. RANDOMIZED CONTROLLED TRIALS/
34. RANDOM ALLOCATION/
35. DOUBLE BLIND METHOD/
36. SINGLE-BLIND METHOD/
37. or/31-36

38. ANIMAL/ not HUMAN/
39. 37 not 38
40. CLINICAL TRIAL.pt.
41. exp CLINICAL TRIALS/
42. (clin$ adj25 trial$).ti,ab.
43. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).ti,ab.
44. PLACEBOS/
45. placebo$.ti,ab.
46. random$.ti,ab.
47. RESEARCH DESIGN/
48. or/40-47
49. 48 not 38
50. 49 not 39
51. COMPARATIVE STUDY/
52. exp EVALUATION STUDIES/
53. FOLLOW UP STUDIES/
54. PROSPECTIVE STUDIES/
55. (control$ or prospectiv$ or volunteer$).ti,ab.
56. or/51-55
57. 56 not 38
58. 57 not (39 or 50)
59. 39 or 50 or 58
60. and/30,59

EMBASE Search Strategy (search dates 1980 to August 29, 2006)

1. Probiotic Agent/ or probiotic$.tw.
2. LACTOBACILLUS/ or LACTOBACILLUS ACIDOPHILUS/ or LACTOBACILLUS CASEI/ or LACTOBACILLUS BIFIDUS/ or BIFIDOBACTERIUM/ or lactobacill$.tw. or bifidobacter$.tw. or "b infantis".tw. or "b bifidum".tw. or "b longum".tw. or sarrharomyce$.tw. or "s boulardii".tw. or clostridium butyricum.tw. or clostridium difficile.mp. or streptococcus thermophilus.tw. or enterococcus faecium.mp.
3. (antibiosis or biotherapeutic agent$).tw.
4. "Microbiological Phenomena and Function"/
5. or/1-4
6. exp Antibiotic Agent/ or antibiotic$.tw.
7. ((antimicrobial or anti microbial or antimycrobial or antimycobacteri$ or antibacteri$ or bacteriocid$) adj3 agent$).tw.
8. 6 or 7
9. exp DIARRHEA/ or diarrhea.tw. or diarrhoe$.tw. or diarhe$.tw. or diahoe$.tw. or dysenter$.tw. or gastro enteritis$.tw. or gastroenteriti$.tw.
10. child/ or infant/ or adolescence/ or exp infant, new born/ or exp child, preschool/
11. (child$ or newborn$ or adolescen$ or infan$).tw.
12. (preschool$ or pre-school$).tw.
13. teen$.tw.
14. (kindergarten$ or kindergarden$).tw.
15. elementary school$.tw.

16. secondary school$.tw.
17. nursery school$.tw.
18. high school$.tw.
19. highschool$.tw.
20. youth$.tw.
21. (baby$ or babies$ or preemie$ or premature$).tw.
22. (schoolchild$ or "school child$").tw.
23. (schoolage$ or school age$).tw.
24. toddler$.tw.
25. pubert$.tw.
26. (pre-pubescen$ or prepubescen$ or post-pubescen$ or postpubescen$).tw.
27. (kid or kids or boy$ or girl$).tw.
28. juvenile.tw.
29. or/11-28
30. 10 or 29
31. RANDOMIZED CONTROLLED TRIAL.pt.
32. CONTROLLED CLINICAL TRIAL.pt.
33. RANDOMIZED CONTROLLED TRIALS/
34. RANDOM ALLOCATION/
35. DOUBLE BLIND METHOD/
36. SINGLE-BLIND METHOD/
37. or/31-36
38. ANIMAL/ not HUMAN/
39. 37 not 38
40. CLINICAL TRIAL.pt.
41. exp CLINICAL TRIALS/
42. (clin$ adj25 trial$).ti,ab.
43. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).ti,ab.
44. PLACEBOS/
45. placebo$.ti,ab.
46. random$.ti,ab.
47. RESEARCH DESIGN/
48. or/40-47
49. 48 not 38
50. 49 not 39
51. COMPARATIVE STUDY/
52. exp EVALUATION STUDIES/
53. FOLLOW UP STUDIES/
54. PROSPECTIVE STUDIES/
55. (control$ or prospectiv$ or volunteer$).ti,ab.
56. or/51-55
57. 56 not 38
58. 57 not (39 or 50)
59. 39 or 50 or 58
60. 30 or 59
61. 5 and 8 and 9 and 30 and 60

CENTRAL Search Strategy (search Issue 3, 2006)

1. exp PROBIOTICS/ or probiotic$.tw.
2. LACTOBACILLUS/ or exp LACTOBACILLUS ACIDOPHILUS/ or lactobacillus bifidus.tw. or lactobacill$.tw. or bifidobacter$.tw. or sarrharomyce$.tw. or clostridium butyricum.tw. or streptococci$.tw.
3. exp ANTIBIOTICS/ or antibiotic$.tw.
4. ANTIINFECTIVE AGENTS/ or DRUG RESISTANCE, MICROBIAL/
5. ((antimicrobial or anti microbial or antimycrobial or antimycobacteri$ or antibacteri$ or bacteriocid$) adj3 agent$).tw.
6. exp DIARRHEA/ or diarrhea.tw. or diarrhoe$.tw. or diarhe$.tw. or diahoe$.tw. or dysenter$.tw. or gastro enteritis$.tw. or gastroenteriti$.tw.
7. child/ or infant/ or adolescence/ or exp infant, new born/ or exp child, preschool/
8. (child$ or newborn$ or adolescen$ or infan$).tw.
9. (preschool$ or pre-school$).tw.
10. teen$.tw.
11. (kindergarten$ or kindergarden$).tw.
12. elementary school$.tw.
13. secondary school$.tw.
14. nursery school$.tw.
15. high school$.tw.
16. highschool$.tw.
17. youth$.tw.
18. (baby$ or babies$ or preemie$ or premature$).tw.
19. (schoolchild$ or "school child$").tw.
20. (schoolage$ or school age$).tw.
21. toddler$.tw.
22. pubert$.tw.
23. (pre-pubescen$ or prepubescen$ or post-pubescen$ or postpubescen$).tw.
24. (kid or kids or boy$ or girl$).tw.
25. juvenile.tw.
26. or/8-25
27. 7 or 26
28. random assignment/
29. random sample/
30. crossover design/
31. exp clinical trials/
32. exp comparative studies/
33. "control (research)".mp.
34. control group/
35. factorial design/
36. quasi-experimental studies/
37. nonrandomized trials/
38. placebos/
39. meta analysis/
40. clinical nursing research.mp. or clinical research/
41. community trials/ or experimental studies/ or one-shot case study/

42. community trials/ or experimental studies/ or one-shot case study/ or pretest-posttest design/ or solomon four-group design/ or static group comparison/ or study design/
43. (clinical trial or systematic review).pt.
44. random$.mp.
45. ((singl$ or doubl$ or tripl$ or trebl$) adj10 (blind$ or mask$)).mp.
46. (cross?over or placebo$ or control$ or factorial or sham$).mp.
47. ((clin$ or intervention$ or compar$ or experiment$ or preventive or therapeutic) adj10 trial$).mp.
48. (meta?analy$ or systematic review$).mp.
49. or/28-48
50. convenience sample/
51. exp research, allied health/ or research, medical/ or research, nursing/
52. research question/
53. nursing practice, research-based/
54. research methodology/
55. exp evaluation research/
56. concurrent prospective studies/ or prospective studies/
57. (nursing interventions or research).pt.
58. or/50-57
59. 49 or 58
60. 1 or 2
61. 3 or 4 or 5
62. 6 and 27 and 59 and 60 and 61

CINAHL Search Strategy (search dates 1982 to August 29, 2006)
1. exp PROBIOTICS/ or probiotic$.tw.
2. LACTOBACILLUS/ or exp LACTOBACILLUS ACIDOPHILUS/ or lactobacillus bifidus.tw. or lactobacill$.tw. or bifidobacter$.tw. or "b infantis".tw. or "b bifidum".tw. or "b longum".tw. or sarrharomyce$.tw. or "s boulardii".tw. or clostridium butyricum.tw. or clostridium difficile.mp. or "streptococcus thermophilus".tw. or enterococcus faecium.mp.
3. (antibiosis or biotherapeutic agent$).tw.
4. exp ANTIBIOTICS/ or antibiotic$.tw.
5. ANTIINFECTIVE AGENTS/ or DRUG RESISTANCE, MICROBIAL/
6. ((antimicrobial or anti microbial or antimycrobial or antimycobacteri$ or antibacteri$ or bacteriocid$) adj3 agent$).tw.
7. exp DIARRHEA/ or diarrhea.tw. or diarrhoe$.tw. or diarhe$.tw. or diahoe$.tw. or dysenter$.tw. or gastro enteritis$.tw. or gastroenteriti$.tw.
8. child/ or infant/ or adolescence/ or exp infant, new born/ or exp child, preschool/
9. (child$ or newborn$ or adolescen$ or infan$).tw.
10. (preschool$ or pre-school$).tw.
11. teen$.tw.
12. (kindergarten$ or kindergarden$).tw.
13. elementary school$.tw.
14. secondary school$.tw.
15. nursery school$.tw.

16. high school$.tw.
17. highschool$.tw.
18. youth$.tw.
19. (baby$ or babies$ or preemie$ or premature$).tw.
20. (schoolchild$ or "school child$").tw.
21. (schoolage$ or school age$).tw.
22. toddler$.tw.
23. pubert$.tw.
24. (pre-pubescen$ or prepubescen$ or post-pubescen$ or postpubescen$).tw.
25. (kid or kids or boy$ or girl$).tw.
26. juvenile.tw.
27. or/9-26
28. 8 or 27
29. random assignment/
30. random sample/
31. crossover design/
32. exp clinical trials/
33. exp comparative studies/
34. "control (research)".mp.
35. control group/
36. factorial design/
37. quasi-experimental studies/
38. nonrandomized trials/
39. placebos/
40. meta analysis/
41. clinical nursing research.mp. or clinical research/
42. community trials/ or experimental studies/ or one-shot case study/
43. community trials/ or experimental studies/ or one-shot case study/ or pretest-posttest design/ or solomon four-group design/ or static group comparison/ or study design/
44. (clinical trial or systematic review).pt.
45. random$.mp.
46. ((singl$ or doubl$ or tripl$ or trebl$) adj10 (blind$ or mask$)).mp.
47. (cross?over or placebo$ or control$ or factorial or sham$).mp.
48. ((clin$ or intervention$ or compar$ or experiment$ or preventive or therapeutic) adj10 trial$).mp.
49. (meta?analy$ or systematic review$).mp.
50. or/29-49
51. convenience sample/
52. exp research, allied health/ or research, medical/ or research, nursing/
53. research question/
54. nursing practice, research-based/
55. research methodology/
56. exp evaluation research/
57. concurrent prospective studies/ or prospective studies/
58. (nursing interventions or research).pt.
59. or/51-58

60. 50 or 59
61. 1 or 2 or 3
62. 4 or 5 or 6
63. 7 and 28 and 60 and 61 and 62

AMED Search Strategy (search dates 1985 to August 29, 2006)

We used same search strategy as MEDLINE (databases use same control vocabulary)

DATA EXTRACTION FORM (Probiotics for the Prevention of Pediatric AAD)

FIRST AUTHOR_____ PUBLICATION YEAR _____

STUDY ID_____ REVIEWER_____

STUDY CHARACTERISTICS

Language_____ Country_____

Setting
- [ ] Community
- [ ] Hospital general care
- [ ] Unclear
- [ ] Other specify:_____

Multi-site [ ] Single site [ ]

Source of Funding
- [ ] Industry
- [ ] Government
- [ ] Other:_____
- [ ] Not reported

Peer Review [ ]
- External [ ]
- Internal [ ]
- Unknown [ ]

Non-peer review [ ]

Study's Definition of Diarrhea

Inclusion Criteria	Exclusion Criteria

DESIGN CHARACTERISTICS

Study Design

[ ] Parallel
[ ] Crossover
[ ] Other:_____

Any further details

Blinding

[ ] Assessors
[ ] Clinicians
[ ] Patients
[ ] Other 1:_____
[ ] Other 2:_____
[ ] Unknown

Patient Make-up

[ ] Consecutive patients
[ ] Random sample
[ ] Convenience Sample (by day of week, time, etc.)
[ ] Other (eg. volunteers): _____
[ ] Unknown

PARTICIPANTS

Population

[ ] Clearly stated (at least 2 of age, sex, diagnosis, socio.)
[ ] Partially stated (one of above only)
[ ] Not mentioned

Total Sample Size_____

	Treatment Number	Treatment Number	Control Number	Total Number
Patients enrolled				
Patients completing the trial				
Patients receiving the full course of treatment				
Dropouts/Withdrawls:				
Excluded:				
Percent loss to follow-up:				

Intention-to-Treat: yes/no

Any further details

Baseline Characteristics

	Age	Sex	Other:
Total			
Treatment:			
Treatment:			
Control:			

Diagnosis

	Total	Treatment	Control
Otitis			
Pharyngitis			
Bronchitis			
Dermatologic			
Sinusitis			
Sepsis			
Meningitis			
Other 1:			
Other 2:			
Other 3:			

Reasons for Withdrawals

Measures of Compliance

INTERVENTIONS

PROBIOTIC	Route	Dose	Duration	Notes
Probiotic strain(s):				
Probiotic strain(s):				
Control:				

Co-interventions

	Route	Dose	Duration	Number
Antiobiotic:				
Others 1:				
Others 2:				
Others 3:				
Others 4:				
Others 5:				

OUTCOMES (Using the primary investigators' definition of diarrhea)

	Timepoint	Treatment 1	Treatment 2	Control	P value
Percent suffering from diarrhea					
Mean duration of diarrhea					
Mean stool consistency					

Mean stool frequency					
Other:					

Total Adverse Events_____

	Timepoint	Treatment 1	Treatment 2	Control
1:				
2:				
3:				
4:				
5:				
6:				

Additional Comments

_____
_____
_____
_____
_____

APPENDIX 3-1 (Systematic Review of RCTs of PAD)

MEDLINE Search Strategy (search dates 1966 to March 22, 2007)

1. exp Diarrhea/
2. (diarrh$ or diarh$ or diahoe$).tw.
3. 1 or 2
4. exp child/ or exp infant/ or exp adolescent/
5. (child$ or newborn$ or adolescen$ or infan$ or neonat$ or teen$ or youth or pediatric$ or paediatric$).tw.
6. 4 or 5
7. 3 and 6
8. exp randomized controlled trials/
9. (controlled clinical trial or randomized controlled trial).pt.
10. (random$ or placebo$).ti,ab,sh.
11. ((singl$ or double$ or triple$ or treble$) and (blind$ or mask$)).tw,sh.
12. exp cohort studies/ or exp case-control studies/
13. controlled clinical trial.pt.
14. (cohort$ or longitudinal or prospective or follow-up stud$ or case control or case series).tw.
15. exp epidemiologic methods/
16. limit 15 to yr=1950-1989
17. ((case$ and control$) or (case$ and series)).ti,ab.
18. EVALUATION STUDIES/ or FOLLOW UP STUDIES/ or PROSPECTIVE STUDIES/
19. or/8-14,16-18
20. 7 and 19
21. (diagnosis or epidemiology).fs.
22. exp Diagnosis/
23. exp Epidemiologic Methods/
24. exp Epidemiology/
25. exp population surveillance/
26. "outcome and process assessment (health care)"/ or "outcome assessment (health care)"/
27. exp severity of illness index/
28. exp nursing assessment/
29. (evaluat$ or diagnos$ or prognos$ or assess$ or measur$ or symptom$ or severity or scale$ or instrument$ or screen$ or predict$).tw.
30. (epidemiolog$ or incidence or surveillance).tw.
31. or/21-30
32. 20 and 31
33. exp Bacterial Infections/
34. exp Virus Diseases/
35. Clostridium difficile/
36. exp Rotavirus/
37. exp Intestinal Diseases, Parasitic/
38. exp HIV Infections/

39. exp diarrhea/ci
40. (antibiotic associated diarrhea or aad).tw.
41. exp diarrhea/ and exp antibacterial agents/ae
42. exp acute disease/ or acute.tw.
43. exp Intubation, Gastrointestinal/ae [Adverse Effects]
44. or/33-43
45. 32 and 44
46. exp Anti-Bacterial Agents/ or exp Lactobacillus/ or exp Probiotics/
47. exp Antibiosis/
48. (probiotic$ or prebiotic$ or biotherapeutic$ or antimicrobial$ or anti microbial$ or antimicrobial$ or antimyco$ or antibiotic$ or bacteriocid$ or oral rehydrat$).tw.
49. (lactobacill$ or l acidophilus or l casei or bifodobacter$ or b infantis or b bifidium or b longum or saccharomyce$ or s bouladii or clostridium butyricum or streptococcus thermophilus or enterococcus faecium).tw.
50. exp Vaccination/
51. exp Fluid Therapy/
52. exp Electrolytes/tu [Therapeutic Use]
53. diosmectite.mp.
54. exp Bismuth/tu [Therapeutic Use]
55. exp Kaolin/ or kaopectate.mp.
56. calcium polycarbophil.mp.
57. exp Analgesics, Opioid/
58. exp Antidiarrheals/
59. exp Zinc/tu [Therapeutic Use]
60. exp Vitamin A/tu [Therapeutic Use]
61. nitazoxanide.mp.
62. exp Quinine/tu [Therapeutic Use]
63. breastfeeding.mp.
64. yogurt.mp.
65. exp Azithromycin/tu [Therapeutic Use]
66. exp Erythromycin/tu [Therapeutic Use]
67. exp Ciprofloxacin/tu [Therapeutic Use]
68. racecadotril.mp.
69. or/46-68
70. 45 and 69
71. limit 70 to english language

EMBASE Search Strategy (search dates 1980 to March 22, 2007)
1. exp diarrhea/
2. exp child/
3. exp infant/
4. exp adolescent/
5. (child$ or newborn$ or adolescen$ or teen$ or youth or infan$ or neonat$ or newborn$ or pediatric$ or paediatric$).tw.
6. or/2-5

7. 1 and 6
8. exp randomized controlled trial/
9. (random$ or placebo$).ti,ab,sh.
10. ((singl$ or double$ or triple$ or treble$) and (blind$ or mask$)).tw,sh.
11. controlled clinical trial$.tw,sh.
12. exp cohort analysis/ or exp longitudinal study/ or exp prospective study/ or exp follow-up/
13. exp case control study/
14. (cohort$ or longitudinal or prospective or follow-up stud$).tw.
15. ((case and control) or (case and series)).tw.
16. or/8-15
17. 7 and 16
18. limit 17 to human
19. exp "diagnosis, measurement and analysis"/
20. exp disease severity/ or exp disease course/
21. exp Health Survey/
22. exp epidemiological data/
23. exp epidemiology/
24. (evaluat$ or diagnos$ or prognos$ or assess$ or measur$ or screen$ or instrument$ or scale$ or surveillance or symptom$ or severity or predict$).tw.
25. incidence.tw.
26. or/19-25
27. 18 and 26
28. exp Probiotic Agent/ or exp oral rehydration therapy/ or exp prebiotic agent/
29. LACTOBACILLUS/ or LACTOBACILLUS ACIDOPHILUS/ or LACTOBACILLUS CASEI/ or LACTOBACILLUS BIFIDUS/ or BIFIDOBACTERIUM/ or "Microbiological Phenomena and Function"/ or exp Antibiotic Agent/ or exp antidiarrheal agent/
30. (probiotic$ or prebiotic$ or biotherapeutic$ or antimicrobial$ or anti microbial$ or antimicrobial$ or antimyco$ or antibiotic$ or bacteriocid$ or oral rehydrat$).tw.
31. (lactobacill$ or l acidophilus or l casei or bifodobacter$ or b infantis or b bifidium or b longum or saccharomyce$ or s bouladii or clostridium butyricum or streptococcus thermophilus or enterococcus faecium).tw.
32. (clostridium difficile or c difficile).tw.
33. diosmectite.mp.
34. exp BISMUTH/
35. kaopectate.mp. or exp Kaolin Pectin/
36. exp Polycarbophil Calcium/
37. exp Opiate/
38. exp ZINC/
39. vitamin a.mp. or exp Retinol/
40. exp Oral Rehydration Solution/
41. exp VACCINATION/
42. fluid therapy/
43. Electrolyte/iv, po, pa [Intravenous Drug Administration, Oral Drug Administration, Parenteral Drug Administration]

44. exp Breast Feeding/
45. exp Yoghurt/
46. exp Antidiarrheal Agent/dt [Drug Therapy]
47. exp NITAZOXANIDE/dt [Drug Therapy]
48. exp QUININE/dt [Drug Therapy]
49. exp AZITHROMYCIN/dt [Drug Therapy]
50. exp ERYTHROMYCIN/dt [Drug Therapy]
51. exp CIPROFLOXACIN/dt [Drug Therapy]
52. exp ACETORPHAN/dt [Drug Therapy]
53. or/28-52
54. 27 and 53
55. limit 54 to english language

CENTRAL Search Strategy (Issue 1, 2007)
1. (diarrh$ or diarh$ or diahoe$).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
2. (infan$ or newborn$ or neonat$ or child$ or pediatric$ or paediatric$ or adolescen$ or teen$ or youth$).mp.
3. 1 and 2
4. (rotavir$ or cryptosporidium or viral infection$ or bacterial infection$ or hiv or aids or acquired immun$ or intubation or enteral or c difficile or clostridium difficile or acute or antibiotic associated or aad).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
5. (parasit$ or gastroenter$ or travel$ or gastrointestinal infection$).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
6. (dysenter$ or cholera).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
7. or/4-6
8. 3 and 7
9. (probiotic$ or prebiotic$ or biotherapeutic$ or antimicrobial$ or anti microbial$ or antimicrobial$ or antimyco$ or antibiotic$ or bacteriocid$ or oral rehydrat$).tw.
10. (lactobacill$ or l acidophilus or l casei or bifodobacter$ or b infantis or b bifidium or b longum or saccharomyce$ or s bouladii or clostridium butyricum or streptococcus thermophilus or enterococcus faecium).tw.
11. (diosmectite or bismuth or kaolin or kaopectate or calcium polycarbophil or polycarbophil calcium or opioid$ or opiate$ or codeine or lomotil or loperamide or Imodium or vancomycin or metronidazole).tw.
12. (antidiarrhea$ or vitamin a or nitazoxanide or quinine or yogurt or breastfe$ or azithromycin or erythromycin or ciprofloxacin or racecadotril).mp. [mp=title, original title, abstract, mesh headings, heading words, keyword]
13. or/9-12
14. 8 and 13

Global Health Search Strategy (search dates 1973 to March 22, 2007)
1. (diarrh$ or diarh$ or diahoe$).mp. [mp=abstract, title, original title, broad terms, heading words]
2. (infan$ or newborn$ or neonat$ or child$ or pediatric$ or paediatric$ or adolescen$ or teen$ or youth$).mp.
3. 1 and 2
4. (random$ or placebo$ or blind$ or mask$).mp.
5. (control$ adj3 trial$).mp.
6. (cohort$ or longitudinal or prospective or follow-up stud$ or case control or case series).mp. [mp=abstract, title, original title, broad terms, heading words]
7. or/4-6
8. 3 and 7
9. (rotavir$ or cryptosporidium or viral infection$ or bacterial infection$ or hiv or aids or acquired immun$ or intubation or enteral or c difficile or clostridium difficile or acute or antibiotic associated or aad).mp. [mp=abstract, title, original title, broad terms, heading words]
10. (parasit$ or gastroenter$ or travel$ or gastrointestinal infection$).mp. [mp=abstract, title, original title, broad terms, heading words]
11. (dysenter$ or cholera).mp. [mp=abstract, title, original title, broad terms, heading words]
12. or/9-11
13. (probiotic$ or prebiotic$ or biotherapeutic$ or antimicrobial$ or anti microbial$ or antimicrobial$ or antimyco$ or antibiotic$ or bacteriocid$ or oral rehydrat$).tw.
14. (lactobacill$ or l acidophilus or l casei or bifodobacter$ or b infantis or b bifidium or b longum or saccharomyce$ or s bouladii or clostridium butyricum or streptococcus thermophilus or enterococcus faecium).tw.
15. (vaccin$ or fluid therapy or electrolytes or diosmectite or bismuth or kaolin or kaopectate or calcium polycarbophil or opioid antidiarrhe$ or zinc or vitamin a or nitroxanide or quinine or breastfe$ or yogurt or azithromycin or erythromycin or ciprofloxacin or racecadotril).mp. [mp=abstract, title, original title, broad terms, heading words]
16. or/13-15
17. 8 and 12 and 16
18. limit 17 to english language
19. (diarrh$ or diarh$ or diahoe$).mp. [mp=abstract, title, original title, broad terms, heading words]
20. (infan$ or newborn$ or neonat$ or child$ or pediatric$ or paediatric$ or adolescen$ or teen$ or youth$).mp.
21. 19 and 20
22. (random$ or placebo$ or blind$ or mask$).mp.
23. (control$ adj3 trial$).mp.
24. (cohort$ or longitudinal or prospective or follow-up stud$ or case control or case series).mp. [mp=abstract, title, original title, broad terms, heading words]
25. or/22-24
26. 21 and 25

27. (rotavir$ or cryptosporidium or viral infection$ or bacterial infection$ or hiv or aids or acquired immun$ or intubation or enteral or c difficile or clostridium difficile or acute or antibiotic associated or aad).mp. [mp=abstract, title, original title, broad terms, heading words]
28. (parasit$ or gastroenter$ or travel$ or gastrointestinal infection$).mp. [mp=abstract, title, original title, broad terms, heading words]
29. (dysenter$ or cholera).mp. [mp=abstract, title, original title, broad terms, heading words]
30. or/27-29
31. (probiotic$ or prebiotic$ or biotherapeutic$ or antimicrobial$ or anti microbial$ or antimicrobial$ or antimyco$ or antibiotic$ or bacteriocid$ or oral rehydrat$).tw.
32. (lactobacill$ or l acidophilus or l casei or bifodobacter$ or b infantis or b bifidium or b longum or saccharomyce$ or s bouladii or clostridium butyricum or streptococcus thermophilus or enterococcus faecium).tw.
33. (vaccin$ or fluid therapy or electrolytes or diosmectite or bismuth or kaolin or kaopectate or calcium polycarbophil or opioid or antidiarrhe$ or zinc or vitamin a or nitroxanide or quinine or breastfe$ or yogurt or azithromycin or erythromycin or ciprofloxacin or racecadotril).mp. [mp=abstract, title, original title, broad terms, heading words]
34. or/31-33
35. 26 and 30 and 34
36. limit 35 to english language

Search results by database

Database	**Number of citations**
MEDLINE	1199
EMBASE	1527
CENTRAL	572
Global Health	461
Combined total	3759
Total after de-duplication	2738

DATA EXTRACTION FORM (Definitions & Outcome Measures in RCTs of PAD)

FIRST AUTHOR_____PUBLICATION YEAR _____

STUDY ID_____ ____ REVIEWER_____ DATE_____

POPULATION & INTERVENTIONS

Study setting:
[ ] Community, [ ] Hospital inpatient, [ ] Hosp outpatient, [ ] Unclear

[ ] Other specify_____

Country(s)_____

Treatment or Prophylaxis trial:
[ ] Prophylaxis trial (eg, vaccine trial, probiotics for AAD)
[ ] Treatment trial (eg, oral or IV rehydration)
If so, bacterial/viral/parasite confirmed [ ] Yes, [ ] No, [ ] Unclear

Specify_____

Number of children enrolled_____ Number of children completed_____

Age range of children included_____ [ ] NS

Age categories [ ] Yes, [ ] No If yes, [ ] *a priori*, [ ] post hoc, [ ] NS

If yes, list age categories_____

Interventions (list):

1._____

2._____

3._____

4._____

5._____

CLINICAL OUTCOMES

Did the authors state a primary outcome related to diarrhea: [ ] Yes, [ ] No

Sample size calculation on diarrhea [ ] Yes, [ ] No, [ ] Unclear (eg 20% difference for main outcomes)

If yes, what was the estimated sample size_____

Study's definition of diarrhea (frequency, consistency, duration, volume, etc.) [ ] NS

Study's definition of diarrhea resolution (frequency, consistency, duration, volume, etc.) [ ] NS

What is/are the **primary outcome**(s) (according to SS and/or text). Please check items relevant to "primary outcome(s)" and describe:

[ ] Frequency_____

[ ] Consistency_____

[ ] Severity_____

[ ] Duration_____

[ ] Volume/Output/Weight_____

[ ] Incidence_____

[ ] Other 1_____

[ ] Other 2_____

[ ] Other 3_____

Other measured outcomes (as reported in tables and figures). Please check items relevant to other outcomes, and describe:

[ ] Absenteeism (social, school or work)
[ ] Cramping
[ ] Dehydration
[ ] Pain
[ ] Temperature

[ ] Fever
[ ] Urgency or incontinence
[ ] Vomiting
[ ] Safety/Adverse events
[ ] Microbiological testing

[ ] Other 1_____

[ ] Other 2_____

[ ] Other 3_____

[ ] Other 4_____

Additional comments:
_____
_____
_____
_____

Who monitored for diarrhea [ ] child, [ ] parent, [ ] caretaker, [ ] healthcare worker,
[ ] unclear, [ ] other_____

Do the authors report the measurement properties of diarrheal primary outcome(s): [ ] Yes, [ ] No
Validity [ ] Yes, [ ] No
Reliability [ ] Yes, [ ] No
Responsiveness [ ] Yes, [ ] No

-If yes to any of the above, please describe how this was reported

Should the study be otherwise flagged because it used unique outcome measures (eg, pictures, scale, chart or categories): [ ] Yes, [ ] No

If yes describe:
_____
_____
_____
_____

RCT - QUALITY ASSESSMENT

Part 1 (from Jadad – Controlled Clin Trials 1996; 17:1-12)

Score

1. Was the study described as randomized (this includes the use of words such as randomly, random and randomization)?
   Yes = 1 No = 0 _____

2. Was the study described as double-blind?
   Yes = 1 No = 0 _____

3. Was there a description of withdrawals and drop-outs?
   Yes = 1 No = 0 _____

Additional points: Add 1 point if:

Method to generate the sequence of randomization was described and appropriate (e.g. table of random numbers, computer generated, coin tossing) _____

Method of double-blinding described and appropriate (identical placebo, active placebo, dummy) _____

Point deduction: Subtract 1 point if:

Method of randomization described and it was inappropriate (allocated alternately, according to date of birth, hospital number, etc.) -_____

Method of double-blinding described but it was inappropriate (comparison of tablet vs injection with no double dummy) -_____

OVERALL SCORE (Maximum 5) _____

Part 2 (from Schulz – JAMA 1995; 273:408-12)

Concealment of treatment allocation:

- ☐ Adequate
- ☐ Inadequate
- ☐ Unclear

Adequate:	e.g. central randomization; numbered/coded containers; drugs prepared by pharmacy; serially numbered, opaque, sealed envelopes
Inadequate:	e.g. alternation, use of case record numbers, dates of birth or day of week; open lists
Unclear:	Allocation concealment approach not reported or fits neither above category

APPENDIX 4-1 (Letter of Invitation to Experts in Pediatric Acute Diarrhea)

Dear Dr. (*Insert Name*),

My name is Bradley Johnston and I am a PhD Candidate in the Department of Medicine, University of Alberta. I am developing and validating two new instruments designed to measure the severity of pediatric acute diarrhea and pediatric acute gastroenteritis for research purposes. Using an expert panel review process, I am seeking input from pediatricians and nurses to help validate the International Pediatric Acute Diarrhea Scale (IPADS) and International Pediatric Acute Gastroenteritis Scale (IPAGS). The items for both scales are based on a systematic review of 121 randomized controlled trials of pediatric acute diarrhea and consultation with a small group of experts in pediatrics and measurement. It is hoped that this effort will provide research scientists involved in studies of acute diarrhea and acute gastroenteritis in children standard and valid instruments for measuring these common pediatric illnesses.

You will also be asked at the end of the survey if you would be willing participate in a follow-up conference call. During this call, which will take approximately three hours, the participants will discuss the results of the study and issues that arise from the survey.

The survey has been approved by the Health Research Ethics Board at the University of Alberta and should take less than 20 minutes to complete. We will ask you to return the survey by **Friday, September 26th**, and to participate in a focus group via conference call approximately 4 weeks later (i.e., **Friday, October 24th** or **Thursday October 30th**).

Participants will not be identified by name and the subsequent analysis of the data will involve group results only. All information pertaining to individual participants will remain confidential, including participant's email and IP address. You may withdraw your consent to participate at any time during the study, without prejudice, and your data will be destroyed. There are no foreseen risks or direct benefits to study participation.

If you are willing to participate, **your consent is implied with the completion and return of the survey questionnaire to follow**.

If you have further questions regarding to this project, please contact me at 780-907-5655 or by email at bjohnston@med.ualberta.ca. Should you have questions regarding one's rights as a research participant, please contact the Health Research Ethics Board at 780-492-0302. I hope that you agree to participate in this exciting collaborative project.

Yours sincerely,

Bradley C. Johnston, PhD Candidate
Dr. Sunita Vohra, MD, MSc
Dr. Todd Rogers, PhD

Introductory Letter

Insert Date Here

Dear (*Insert Name Here*),

Thank you for agreeing to participate in this study, which consists of three parts: (i) draft definitions of pediatric acute diarrhea and pediatric acute gastroenteritis (page 2, page 10); (ii) survey of items for the measure of pediatric acute diarrhea and pediatric acute gastroenteritis severity (page 3, page 11); (iii) background questionnaire about yourself and your background (page 15).

Please note that there are definitions for pediatric acute diarrhea (on page 2) and pediatric acute gastroenteritis (on page 10). The six common items follow the definition for pediatric acute diarrhea and the two unique items follow the definition for pediatric acute gastroenteritis.

I would be extremely grateful if you would attempt to complete the review within **10 days** (**Friday, September 26**) of the date you receive this letter. If you have any questions, please do not hesitate to contact me. Your help in validating this instrument is very much appreciated.

The survey questionnaire can be found at:

https://www.epicore.ualberta.ca/ipads_ipags/

Sincerely,

Bradley C. Johnston, PhD Candidate

Section 1 - IPADS

Definition of Pediatric Acute Diarrhea – PLEASE READ BEFORE RESPONDING

The International Pediatric Acute Diarrhea Scale (IPADS) is an instrument intended to measure the severity of pediatric acute diarrhea and its associated symptoms in children (birth to ≤72 mo). Although there are many definitions of acute diarrhea, the following definitions of diarrhea and its resolution are proposed:

Pediatric Acute Diarrhea (definition)
Pediatric acute diarrhea is defined by an increase in the frequency of bowel movements and a change in the consistency of stool (i.e., diminished degree of firmness). Diarrhea may be associated with dehydration, pain or discomfort, restrictions in normal daily activities of the child or caregiver, and fever.

Resolution of Pediatric Acute Diarrhea (definition)
The duration of acute diarrhea typically resolves in less than 14 days and resolution is marked by (a) production of 2 consecutive normal stools or; (b) production of one normal stool followed by 12 hours with no stool production or; (c) no stool production for a period of 12 hours.

1. Is this definition complete?

a) Please indicate the degree to which you agree with this definition for *pediatric acute diarrhea*. Use the following five-point scale on the right and place an X in the box corresponding to your answer.
Feel free to make comments on the definition of pediatric acute diarrhea in the space below.

Comments..... Strongly disagree 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
..... Strongly agree 5. ☐

b) Please indicate the degree to which you agree with the definition for *resolution of pediatric acute diarrhea*. Use the following five-point scale on the right and place an X in the box corresponding to your answer.
Feel free to make comments on the definition of pediatric acute diarrhea in the space below.

Comments..... Strongly disagree 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
..... Strongly agree 5. ☐

To our knowledge, there is no standard or valid outcome measure for evaluating pediatric acute diarrhea (PAD) or pediatric acute gastroenteritis (PAG). This makes it very difficult to compare, contrast, or combine the results of randomized controlled trials. The International Pediatric Acute Diarrhea Scale (IPADS) and International Pediatric Acute Gastroenteritis Scale (IPAGS) have been designed to determine the severity these illnesses and their associated symptoms in children (birth to ≤72 mo). It is anticipated that IPADS and IPAGS will be used in clinical trials to assess the effects of treatment or prophylactic interventions for in both outpatient and inpatient settings, as scored by a healthcare provider and parent.

The IPADS contains six items. The IPAGS contains 8 items. Using a 5-point scale, for each item below you will be asked to indicate:

1) How *relevant* each item is as a measure of the severity of PAD or PAG

2) With respect to each item, how *appropriate* is each accompanying subscale as a measure of PAD or PAG

In addition, you will be asked to rate how *representative* the diagrams of stool consistency are and how *representative* the set of "relevant" items are of the overall construct - pediatric acute diarrhea and pediatric acute gastroenteritis. Please read each statement carefully and then use the following five-point scale to indicate the degree of *relevancy* or *representativeness*:

1 = *Not representative*, 2, 3, 4, 5 = *Very representative*

Important items to note:

1) The first section of the survey focuses on acute diarrhea only (International Pediatric Acute Diarrhea Scale); whereas, the second section focus on acute gastroenteritis (International Pediatric Acute Gastroenteritis Scale).

2) All proposed items are scored one time per day.

- Six of eight items refer to the previous 24 hours; whereas the additional items, i.e., *Duration with loose & unformed or liquid stools*, *Duration of vomiting* refers to the cumulative number of affected days for specific symptoms.

- Two items, *dehydration* and *fever* refer to the child's health at the time of evaluation. One item, *pain/discomfort* is anchored with behaviors that the child may have displayed in the previous 24 hours. This item would be scored by the parent, as they make the best proxy for measuring pediatric pain.

- The intent of the scale is to be used once a day, ideally at the same time of day for each evaluation (i.e. 24 hours after the last evaluation).

For each item, please insert an X in the box [X] to the right of the item to indicate your response. If you wish to add a comment, please do so in the space provided.

1 Stool Consistency Diagrams

Before starting, please indicate how *representative* the diagrams (and accompanying text) below are for yielding an accurate and consistent estimation of the full range of stool consistency seen in children (birth to ≤72 mo).

Hard and Formed	Soft but Formed	Loose and Unformed	Liquid
			
Having a hard or firm texture & retaining a definite shape like a banana, cigar or marbles.	Retaining same general shape in the collection bag; does not spread all over the bottom of the bag or has a texture that appears like peanut butter.	Lacking any shape of its own; spreads over the bottom of the collection bag; having a texture that appears like hot cereal.	Like water.

Adopted from Stool Consistency Classification System, Bliss *et al*. J Wound Ostomy Contin Nurs 2001

Comments: ..... Not representative 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very representative 5. ☐

2 Stool Frequency and Consistency/24 hrs (0-10 scale)

a) Regardless of your answer to # 1 above, please rate the relevancy of the item below.

Stool frequency in preceding 24 hours				
Stool consistency	0 stools	1 – 2 stools	3 – 4 stools	≥ 5 stools
Loose & unformed	0	1	5	9
Liquid	0	2	6	10

Constructed based on previous work by: Whelan *et al.* Euro J Clin Nutr 2004; Bliss *et al.* J Wound Ostomy Contin Nurs 2001

Comments: ..... Not relevant 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very relevant 5. ☐

b) Please rate the appropriateness of the accompanying scoring scale for measuring stool frequency and consistency.

Comments: ..... Not appropriate 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very appropriate 5. ☐

3 Number of Days with Loose/Unformed Stools or Liquid Stools (0-10 scale)

a) Please rate the relevancy of the following item.

Number of Days (in the previous 13 days)				
Stool Consistency	0 days	1 – 2 days	3 - 4 days	≥5 days
Loose & Unformed Stools	0	1	5	9
Liquid Stools	0	2	6	10

Comments: ..... Not relevant 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very relevant 5. ☐

b) Please rate the appropriateness of the scoring scale for measuring the number of days with loose/unformed stools or liquid stools.

Comments: ..... Not appropriate 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very appropriate 5. ☐

4 Level of Dehydration (0-10 scale)

a) Please rate the relevancy of the following item.

Sign or Symptom	**Score = 0 each**	**Score = 1.5 each**	**Score = 2.5 each**
General Appearance	Normal	Thirsty, restless or lethargic but irritable when touched	Drowsy, limp, cold, sweaty, +/- comatose
Eyes	Normal	Slightly sunken	Very sunken
Mucus membranes (tongue)	Moist	"Sticky"	Dry
Tears†	Tears	Decreased tears	Absent tears

†By history or examination

Adapted from Friedman *et al.* J Pediatrics 2004; Goldman *et al*. Pediatrics 2008

Comments: ..... Not relevant 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very relevant 5. ☐

b) Please rate the appropriateness of the scoring scale for measuring the level of dehydration.

Comments: ..... Not appropriate 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very appropriate 5. ☐

5 Pain and Discomfort (0-10 scale)

a) Please rate the relevancy of the following item.

Category	**Score = 0**	**Score = 1**	**Score = 2**
Face	No particular expression or smile	Occasional grimace or frown, withdrawn, disinterested	Frequent or constant quivering chin, clenched jaw
Legs	Normal position or relaxed	Uneasy, restless, tense	Kicking, or legs drawn up
Activity	Lying quietly, normal position, moves easily	Squirming, shifting back and forth, tense	Arched, rigid or jerking
Cry	No cry (awake or asleep)	Moans or whimpers; occasional complaint	Crying steadily, screams or sobs, frequent complaints
Consolability	Content, relaxed	Reassured by occasional touching, hugging or being talked to, distractible	Difficult to console or comfort

Adopted from Merkel et al. Pediatr Nurs 1997 (Face Legs Activity Cry Consolability Scale)

Comments: ..... Not relevant 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very relevant 5. ☐

b) Please rate the appropriateness of the scoring scale for measuring pain and discomfort.

Comments: ..... Not appropriate 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very appropriate 5. ☐

6 Restrictions of Normal Daily Activities/24 hours (0-10 scale)

a) Please rate the relevancy of the following item.

Score = 0	Score = 2	Score = 4	Score = 6	Score = 8	Score = 10
Activities not disturbed	Child less playful/social	Child or caregiver sleep or daily activities disturbed	Child or caregiver unable to attend to homemaking duties, daycare/school, or work	Visit to healthcare practitioner	Admitted to hospital

Comments: ..... Not relevant 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very relevant 5. ☐

b) Please rate the appropriateness of the scoring scale for measuring restrictions of normal daily activities.

Comments: ..... Not appropriate 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very appropriate 5. ☐

7 Fever (0-10 Scale)

a) Please rate the relevancy of the following item.

Temperature (Axillary)				
	34.7°C to 37.3 °C	37.4 to 38.1 °C	38.3 to 38.2 °C	≥39.0 °C
Yes	0	2	6	10

Comments: ..... Not relevant 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very relevant 5. ☐

b) Please rate the appropriateness of the scoring scale for measuring fever.

Comments: ..... Not appropriate 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very appropriate 5. ☐

8 Thank you for completing the questions on the relevancy of each of the individual items. Overall, do the items you ranked as *relevant* (4 or 5) *represent* pediatric acute diarrhea (page 2).

Comments: .....
..... Not representative 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very representative 5. ☐

9 If you indicated that the items you rated as 4 or 5 were not representative of the pediatric acute diarrhea as defined, what variable(s) would you recommend be **added**?
If so, please indicate the variable and nature of the item you recommend be used to collect the data for measuring the severity of pediatric acute diarrhea in the space below.

Variable/item:.....
.....
.....
.....
.....

Section 2 - IPAGS

Definition of Pediatric Acute Gastroenteritis - PLEASE READ BEFORE RESPONDING

The International Pediatric Acute Gastroenteritis Scale (IPAGS) is an instrument intended to measure the severity of pediatric acute gastroenteritis and its associated symptoms in children (birth to ≤72 mo). Although there are many definitions of acute gastroenteritis, the following definition of acute gastroenteritis are proposed:

Pediatric Acute Gastroenteritis (definition)
Pediatric acute gastroenteritis is defined by an increase in the frequency of bowel movements, and a change in the consistency of stool (i.e., diminished degree of firmness) and vomiting. Gastroenteritis may be associated dehydration, pain or discomfort, restrictions in normal daily activities of the child or caregiver and fever.

Resolution of Pediatric Acute Gastroenteritis (definition)
The duration of acute gastroenteritis typically resolves in less than 14 days and resolution is marked by (a) production of 2 consecutive normal stools and absence of vomiting or; (b) production of one normal stool followed by 12 hours with no stool production or vomiting or; (c) no stool production or vomiting for a period of 12 hours.

1 Is this definition complete?

a) Please indicate the degree to which you agree with this definition of acute gastroenteritis. Use the following five-point scale on the right and place an X in the box corresponding to your answer.

Comments.....	Strongly disagree 1.	☐
.....	2.	☐
.....	3.	☐
.....	4.	☐
.....	Strongly agree 5.	☐

b) Please indicate the degree to which you agree with the definition for resolution of acute gastroenteritis. Use the following five-point scale on the right and place an X in the box corresponding to your answer.

Comments.....	Strongly disagree 1.	☐
.....	2.	☐
.....	3.	☐
.....	4.	☐
.....	Strongly agree 5.	☐

2 Which of the items you just rated would you include in an instrument developed to measure pediatric acute gastroenteritis as defined on the previous page? Please place an X in the box to the right of each item you would include.

1. Stool Frequency and Consistency ☐
2. Number of Days with Loose/Unformed Stools or Liquid Stools ☐
3. Level of Dehydration ☐
4. Pain and Discomfort ☐
5. Restrictions in Normal Daily Activities ☐
6. Fever ☐

Now consider the following three additional items for measuring pediatric acute gastroenteritis.

3 Vomiting Frequency/24 hours

a) Please rate the relevancy of the following item.

Episodes of Vomiting in Preceding 24 hours				
Vomiting	0 episodes	1 – 2 episodes	3 - 4 episodes	≥5 episodes
Yes	0	2	6	10

Comments: ..... Not relevant 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very relevant 5. ☐

b) Please rate the appropriateness of the scoring scale for measuring vomiting frequency.

Comments: ..... Not appropriate 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very appropriate 5. ☐

4 Duration of Vomiting

a) Please rate the relevancy of the following item.

Number of Days with Vomiting (in the previous 13 days)				
Vomiting	0 days	1 – 2 days	3 - 4 days	≥5 days
Yes	0	2	6	10

Comments: ..... Not relevant 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very relevant 5. ☐

b) Please rate the appropriateness of the scoring scale for measuring the number of days with vomiting.

Comments: ..... Not appropriate 1. ☐
..... 2. ☐
..... 3. ☐
..... 4. ☐
Very appropriate 5. ☐

5 Overall, do the items you ranked as *relevant* (4 or 5) in section 1 and 2 *represent* pediatric acute gastroenteritis (page 10).

Comments: .....
..... Not representative 1. ☐
..... 2. ☐
..... 3. ☐
4. ☐
Very representative 5. ☐

6 If you indicated that the items you rated as 4 or 5 were not representative of the pediatric acute gastroenteritis as defined, what variable(s) would you recommend be:

a) Deleted? (please insert an X in the boxes corresponding to the item(s) you recommend deleting):

1. Stool Frequency and Consistency ☐
2. Number of Days with Loose/Unformed Stools or Liquid Stools ☐
3. Level of Dehydration ☐
4. Pain and Discomfort ☐
5. Restrictions in Normal Daily Activities ☐
6. Fever ☐
7. Number of Vomiting Episodes ☐
8. Number of Days with Vomiting ☐

b) <u>Added</u>? (Please indicate the variable and nature of the item you recommend be used to collect the data for measuring the severity of <u>pediatric acute gastroenteritis</u> in the space below).

Variable/item:.....

.....

.....

.....

7 Additional comments:

……..……..

.....

.....

.....

8 A final review round will include a conference call with panel members. Would you be willing to participate in a conference call to discuss the results of this survey (target date of **Friday, October 24th** or **Thursday, October 30th**)?

Yes ☐
No ☐

If yes, please provide your preferred contact information (i.e., phone, email) and when is the best time to reach you (including the time zone you reside in)

Name:…….

Phone number and extension:……

Email:…….

Please provide your availability on October 24th and/or 30th (in Eastern Standard Time)?

……

APPENDIX 4-2: Background questionnaire

1. What is your profession?
[ ] Pediatrician
[ ] Nurse
[ ] Other_____

2. Do you have a pediatric specialty or subspecialty?

[ ] Yes (please continue)
[ ] No (please go to question 3)

[ ] Emergency
[ ] Gastroenterology
[ ] General Pediatrics
[ ] Other_____

3. Which of the following do you consider your self primarily?
[ ] Clinician
[ ] Clinician-scientist
[ ] Researcher

If you are a "clinician-scientist" or a "researcher", which type of research do you do primarily?
[ ] Clinical trials
[ ] Research methodology
[ ] Epidemiology
[ ] Other_____

4. How many years have you been in practice (excluding residency)?
[ ] Less than 5 years
[ ] 5 to 10 years
[ ] 11 to 15 years
[ ] 16 to 20 years
[ ] Great than 20 years

5. What type of setting/institution do you practice in primarily?
[ ] Community
[ ] Hospital
[ ] Academic/University
[ ] Other_____

6. What country do you work in?
[ ] Australia
[ ] Bangladesh
[ ] Brazil
[ ] Canada
[ ] Finland
[ ] India
[ ] Israel
[ ] Italy
[ ] Peru
[ ] Poland
[ ] United Kingdom
[ ] United States
[ ] Other_____

7. How many hours per week, on average, do you provide clinical care?

Please fill in the number_____

8. What is the number of children (birth to ≤72 mo), on average, you see per week?

Please fill in the number_____

9. How many of these, on average, would you say present with acute diarrhea or acute gastroenteritis?

Please fill in the number_____

10. Have you had children?
[ ] No
[ ] Yes

11. If, yes how many?

Please fill in the number_____

12. If you have children, has your child(ren) suffered from acute diarrhea or acute gastroenteritis?
[ ] No
[ ] Yes

13. Have you participated in randomized controlled trials of pediatric acute diarrhea?
[ ] No
[ ] Yes

14. If yes, how many studies?

Please fill in the number_____

15. Have you participated in randomized controlled trials of pediatric acute gastroenteritis?
[ ] No
[ ] Yes

16. If yes, how many studies?

Please fill in the number_____

17. What was your role in these studies? Regarding your role in the study(s) you've been involved, please fill in a number beside each of the roles. For example, if you were a co-investigator on two RCTs of acute diarrhea, please insert a "2" in the co-investigator brackets.
[ ] Principal Investigator
[ ] Senior Investigator
[ ] Co-investigator

18. Have you participated in studies that evaluate reliability or validity?
[ ] No
[ ] Yes

19. If yes, how many studies?

Please fill in the number_____

20. What was your role in these studies? Regarding your role in the study(s) you've been involved, please fill in a number beside each of the roles. For example, if you were a co-investigator on two studies involving the measurement (i.e., validity or reliability), please insert a "2" in the co-investigator brackets.
[ ] Principal Investigator
[ ] Senior Investigator
[ ] Co-investigator

21. What is your age?
[ ] 20 to 29 years
[ ] 30 to 39 years
[ ] 40 to 49 years
[ ] 50 to 59 years
[ ] 60 to 69 years
[ ] 70 to 79 years

22. What is your gender?
[ ] Male
[ ] Female

Thank you for taking time to complete our survey. I will be sure to send you a copy of the final manuscript.

Reminder Letter to Experts in Pediatric Acute Diarrhea

Dear expert panel member,

If you have already responded to my survey, please ignore this email. If not, I would like to remind you to please participate in developing a new tool to measure the severity of pediatric acute diarrhea and pediatric acute gastroenteritis. This is the last component of my PhD project and as you might imagine, I hope to finish soon!

Please note that your completion of the online survey implies your voluntary consent to participation. There is no financial compensation for participation. Abstention or withdrawal at any time will not in any way be penalized. The Human Research Ethics Board at the University of Alberta Health Sciences Faculties has approved this study.

The survey questionnaire can be found at:

https://www.epicore.ualberta.ca/ipads_ipags/

Yours sincerely,

Bradley Johnston, PhD Candidate

APPENDIX 4-3: Structured Questionnaires sent to Outlying Experts

Contact with Expert 1 (Panel A)

Section/Item; Score; Comments (from initial survey)

- *Definition of Diarrhea; 4.00; No diarrhea for 24 or 48 hours*
- *Definition of Diarrhea; 4.00; The diarrhea free interval should be longer*
- *Diarrhea Frequency Scale; 1.00; Too complicated, especially in a community setting*
- *Duration of Diarrhea Scale; 1.00; It is unclear for which setting the appropriateness refers to. I am responding primarily for use in community-based trials*
- *Dehydration Scale; 1.00; Better to use what is recommended by WHO to maintain consistency*

Questions:

1. Based on feedback from yourself and others (international experts in pediatric acute diarrhea), I have collapsed the definitions for AD and AG and increased the diarrhea free interval period to 24 hours (please see definitions and revisions in blue text below). Do you have any major concerns with the new definition?

Pleae note: Overall, your ratings for the definition of diarrhea where low. It was not made clear to survey participants that the definition was a "constitutive" definition, used to generate the items (e.g., diarrhea consistency, dehydration etc.) for the overall scale. The subscales for each item were then used to operationalize (or quantify) the definition.

Response:

- *Would still prefer the diarrea free interval to be 48 hours or more*
- *Am unsure whether the resolution definition can be applied in setting where literacy rates are low and information on type of stools passed is obtained through recall*.

2. Based on ratings and comments from yourself and others, I have collapsed the scales for AD and AG. Do you have any major concerns with the scale (below)?

Response:

No

3. Based on comments from yourself and others I have simplified the following two subscales (Diarrhea Frequency and Duration of Diarrhea. Do you have any major concerns with the subscales (below)?

Response:
Perhaps, there should be additional cut offs with
- *≥7 (used in several trials in the past) and*
- *≥14 (persistent diarrhea)*

4. Based on comments from yourself and others, I have added two categories (i.e., capillary refill and skin turgor) to the Dehydration subscale (below). Do you have any major concerns with this?

Response:
Hope this is consistent with the recommendations by World Health Organization (diagnosis of none, some and severe dehydration) as management guidelines are based on the same.

5. Based on your comments, the final correspondance with 10 international experts in RCTs of Pediatric Acute Diarrhea will ask experts to rate the useability of the final overall scale (to be determined) by parents (community setting), nurses and doctors (inpatient setting). Would such a question help decipher the useability of the final scale in research studies in different settings?

Response:
Yes

Contact with Expert 2 (Panel A)

Section/Item; Score; Comments (from initial survey)

- *Gastroenteritis Definition; 5.00; I don't think there is a need for separate gastroenteritis scale. Vomiting may or not be part of diarrhea and is simply a symptom not really different in severity.*

- *Diarrhea Diarrhea Scale; 1.00; Loose/unformed and liquid can be combined - they often occur on the same day*

- *Dehydration Scale; 4.00; Would add skin turgor*

- *Activities of Daily Living; 3.00; In developing countries most severe diarrhea is in very young where scores 2-6 are difficult*

  Please note: (2 = child less playful/social; 4 = child or caregiver sleep or daily activities disturbed; 6 = child or caregiver unable to attend to homemaking duties, daycare/school, or work)

Questions:

1. Based on feedback from yourself and other I have collapsed the definitions for AD and AG (please see definitions and revisions in blue text below). Do you have any major concerns with the new definitions?

> Please note: Overall, your ratings for the definition of diarrhea where low. It was not made clear to survey participants that the definition was a "constitutive" definition, used to generate the items (e.g., diarrhea consistency, dehydration etc.) for the overall scale. The subscales for each item were then used to operationalize (or quantify) the definition.

Response:

See below

2. Based on comments from yourself and others (international experts in PAD), I have collapsed the scales for AD and AG. Do you have any major concerns with the scale (below)?

Response:

See below

3. Based on comments from yourself and others, I have added two categories (i.e., capillary refill and skin turgor) to the Dehydration subscale (below). Do you have any major concerns with this?

Response:
See below

4. Based on your feedback, I have revised the "Restrictions in Activities of Daily Living" categories (below). In particular, I have dropped "child less playful/social" and revised the wording on the other categories. Do you have any major concerns with this?

Response:
Expert 2 had one overall response: *I agree with the changes*

Contact with Expert 5 (Panel A)

Section/Item; Score; Comments (from initial survey)

- *Gastroenteritis Definition; 5.00; The presence of vomiting does not necessarily characterize a gastric involvement. The whole definition is flawed.*

- *Additional comments: Again, I don't think one can make the differentiation between acute diarrhea and acute gastroenteritis. The term "gastroenteritis" - although admittedly widely used - should be deleted, as the presence of gastric and small intestinal inflammation (as implied in the term) is far from constant in the episodes of acute diarrhea in children. Examples: Rotavirus, Norovirus, Cholera, E.coli, Campylobacter, Yersinia (no gastric involvement); Cholera (no enteritis)*

Questions:
1. Judging from your comments, you seem to indicate that AD and AG should be seperated. Is this correct?

Response:
Yes

2. I have collapsed the definitions for AD and AG. Do you have any major concerns with the new definition (below)?

Response:
As I mentioned, there is confusion in the literature and medical community regarding the term gastroenteritis. I understand if you use the term gastroenterology in your scale as it is so commonly used/recognized, but it would help to clarify the confusion with this term at the beginning of the publication that results from this work. For example, cholera results in acute diarrhea but does not result in inflamation of the stomach or small intestine. The pathophysiology of this should be well-documented in the literature.

Regarding the definition, Expert 5 suggested the changes underlined:

<u>Pediatric Acute Diarrhea & Gastroenteritis (definition)</u>
Pediatric acute diarrhea is defined by an increase in the frequency of bowel movements, and a change in the consistency of stool (i.e., diminished degree of firmness), <u>and is commonly, but not necessarily, associated with</u> vomiting. Diarrhea may be associated <u>with</u> dehydration, <u>fever</u> and restrictions in normal daily activities of the child or caregiver.

<u>Resolution of Pediatric Acute Diarrhea & Gastroenteritis (definition)</u>
<u>Acute diarrhea</u> typically lasts less than 7 days and not longer than 14 days (Guarino et al. 2008) and resolution is marked by (a) production of 2 consecutive normal stools and absence of vomiting or; (b) production of one normal stool followed by 12 hours with no stool production or vomiting or; (c) no stool production or vomiting for a period of 24 hours.

3. I have collapsed the scales for AD and AG. Do you have any major concerns with the scale (below)?

Response:
It looks like a robust document.

4. Based on feedback from other panel A members, I have added two categories (i.e., capillary refill and skin turgor) to the Dehydration subscale (below). Do you have any major concerns with this?

Response:
I fully endorse the the addition of skin turgor and capillary refill.

APPENDIX 4-4: Survey 2

Thank you for taking the time to participate in survey #2.

Overall, the ratings from the panel members on the definition of diarrhea were low. This may have been because it was not made clear that the definition was a "constitutive" definition used to generate the items (e.g., *Diarrhea Frequency, Dehydration, Fever,* etc.) to include in the scale. These items were to operationalize the definition.

Based on feedback from the group of nineteen international experts in pediatric acute diarrhea as well as findings in the literature on randomized controlled trials and/or measurement of pediatric acute diarrhea (e.g., risk of dehydration or hospitalization), I have made the eight revisions listed below.

1. Collapsed the definitions for *Acute Diarrhea* and *Acute Gastroenteritis* and made some minor revisions
2. Collapsed the scales for *Acute Diarrhea* and *Acute Gastroenteritis*
3. Collapsed the stool consistency categories "*Loose & Unformed* and *Liquid Stool(s)*" into one category
4. Revised the *Dehydration* categories (i.e., added *Capillary Refill* and *Skin Turgor*)
5. Removed the *Pain* subscale
6. Revised the *Restrictions in Activities of Daily Living* categories
7. Removed scaling values (i.e., 0 to 10) as these will be developed empirically
8. Revised the response options for *Diarrhea Frequency* and *Vomiting Frequency*

Please note that the instrument (IPADG) is proposed for use in randomized controlled trials (e.g., prevention trials involving vaccines or treatment trials involving ORS) by doctors, nurses, research assistants and associates, and by parents/guardians in a community setting given appropriate training (e.g., capillary refill, skin turgor, axillary temperature).

Please review the revisions to the definitions and IPADG Scale, which are underlined in italics below. Then, would you please respond to the three questions that follow the Scale (as it will be formatted for empirical field research). These questions are on the useability of the IPADG scale. If you would like to add some comments, space for comments is provided after these three questions

At the very end, a summary of the ratings made by the expert panel is provided for your information. Names have been removed to respect confidentiality.

Again, many thanks for your time and the assístance you have provided me.

Sincerely,

Bradley Johnston, PhD Candidate

Definitions

Pediatric Acute Diarrhea & Gastroenteritis (definition)

For children up to 5 years of age, acute diarrhea is marked by an increase in the frequency of bowel movements above normal for the individual and a change in the consistency of stool (i.e., loose and unformed or liquid stool), and is commonly, but not necessarily, associated with vomiting. Diarrhea may also be associated with fever, restrictions in normal daily activities of the child and dehydration.

Resolution of Pediatric Acute Diarrhea & Gastroenteritis (definition)

For children up to 5 years of age, acute diarrhea typically lasts less than 7 days and not longer than 14 days and resolution is marked by (a) production of 2 consecutive normal stools (i.e., soft and formed or hard and formed stool) and absence of vomiting or; (b) production of one normal stool followed by 12 hours with no stool production or vomiting or; (c) normal stool production (or no stool production) and no vomiting for a period of 24 hours.

International Pediatric Acute Diarrhea & Gastroenteritis (IPADG) Scale

Instructions

First please review the *Stool Consistency Classification System* below (Bliss *et al.* 2001). Then imagine responding to each item by placing a check in the appropriate box. In particular, imagine the clinical characteristics of a typical child, up to 5 years of age, with acute diarrhea.

Please note that "*Loose & Unformed or Liquid Stool(s)"* = diarrheal stool.

International Pediatric Acute Diarrhea and Gastroenteritis Scale (Version 2)

1) Stool Consistency Classification System			
Hard and Formed	**Soft but Formed**	**Loose & Unformed**	**Liquid**
			
Having a hard or firm texture an retaining a definite shape like a banana, cigar or marbles	Retains its general shape; has a texture that appears like *butter*	Lacking any shape of its own; having a texture that appears like hot cereal	Like water

2) Diarrheal Frequency				
	Diarrhea Stools in Preceding 24 hours			
Stool Consistency	0 stools	1 – *4* stools	*5 – 8* stools	*≥9* stools
Loose & Unformed Or Liquid Stool(s)	□	□	□	□

3) Duration of Diarrhea				
	Number of Days with Diarrhea (in previous 13 days)			
Stool Consistency	0 days	1 – 2 days	3 - 4 days	≥5 days
Loose & Unformed Or Liquid Stool(s)	□	□	□	□

4) Vomiting Frequency				
	Episodes of Vomiting in Preceding 24 hours			
Vomiting	0 episodes	*1 episode*	*2 episodes*	*≥3 episodes*
Yes	□	□	□	□

5) Duration of Vomiting				
	Number of Days with Vomiting (in the previous 13 days)			
Vomiting	0 days	1 – 2 days	3 - 4 days	≥5 days
Yes	□	□	□	□

6) Fever				
	Temperature			
Axillary	34.7 to 37.3°C	37.4 to 38.1°C	38.3 to 38.2°C	≥39.0°C
Yes	□	□	□	□

7) Normal Daily Activities				
Restrictions	□	□	□	□
Child's normal daily activities (e.g., *eating,* sleeping, playing, daycare or school)	*Normal*	*Disturbed*	*Unable to participate*	*Admitted to hospital*

8) Dehydration			
	Signs & Symptoms		
General *Behavior*	□ Normal	□ Thirsty, restless or letharg but irritable when touched	□ Drowsy, limp, cold, sweaty, +/ comatose
Eyes	□ Normal	□ Slightly sunken	□ Very sunken
Mucus Membranes (tongue)	□ Moist	□ "Sticky"	□ Dry
Tears	□ Tears	□ Decreased tears	□ Absent tears
Skin Turgor	□ *Immediate*	□ *Slow (≤ 2 sec)*	□ *Very slow (> 2 sec)*
Capillary Refill	□ *< 1.5 sec*	□ *1.5 to 3 sec*	□ *> 3 sec*

Please answer the following three questions on the IPADG scale.

1) Is IPADG appropriate for use by parents in a *community* setting?

Please mark yes or no:
[ ] Yes
[ ] No

2) Is the IPADG appropriate for use by nurses and doctors in an outpatient setting?

Please mark yes or no:
[ ] Yes
[ ] No

3) Is the IPADG appropriate for use by nurses and doctors in an inpatient setting?

Please mark yes or no:
[ ] Yes
[ ] No

4) Please provide any comments you may have in the box below: